# TABLE OF CONTENTS

Technical Papers .................................................................................................................................................................... iv
Welcome from the General Co-Chairs .................................................................................................................................. xxi
IEEE IFCS-EFTF 2015 Organizing Committee ................................................................................................................ xxiii
IFCS-EFTF 2015 Joint Technical Program Committee ................................................................................................... xxv
Special Thanks .................................................................................................................................................................... xxix
Exhibitors ........................................................................................................................................................................... xxx
IFCS 2015 Awards ............................................................................................................................................................ xxxiii
EFTF 2015 Awards ............................................................................................................................................................ xxxiv
Student Paper Competition ................................................................................................................................................ xxxv
Future Symposia ............................................................................................................................................................... xxxvii
Tutorials ............................................................................................................................................................................ xxxix
Entrepreneurs Forum ........................................................................................................................................................... xli
IEEE Women in Engineering ................................................................................................................................................ xl
Time & Frequency and Fundamental Physics .................................................................................................................... xlii
Plenary Session Invited Talk ............................................................................................................................................... xliii
Monday, April 13, 2015

Session: A1L-A: Non Linear Phenomena & Mechanical Signal Processors

**Parametric Excitation in Geometrically Optimized AlN Contour Mode Resonators** ..................................................... 1
Ruochen Lu, Anming Gao, Songbin Gong
University of Illinois at Urbana Champaign, United States

Third Order Intermodulation Distortion in Capacitive-Gap Transduced Micromechanical Filters ........................................... 5
Jalal Naghsh Nilchi, Ruonan Liu, Scott Li, Mehmet Akgul, Tristan Rocheleau, Clark Nguyen
University of California, Berkeley, United States

**Nonlinear Acceleration Sensitivity of Quartz Resonators** ............................................................................................. 11
Jianfeng Chen, Yook-Kong Yong, Randall Kubena, Deborah Kirby, David Chang
1HRL Laboratories, LLC, United States; 2Rutgers University, United States

**Multiple SAW Resonance Sensing Through One Communication Channel with Multiple Phase Detectors** ............... 17
Yoshinori Takizawa, Takayuki Shibata, Shinji Kashiwada, Yasuo Yamamoto, Masayoshi Esashi, Shuji Tanaka
1DENSO Corporation, Japan; 2Tohoku University, Japan

Session: A1L-B: Vapor Cell Properties

**Imaging the Static Magnetic Field Distribution in a Vapor Cell Atomic Clock** ............................................................ 21
Christoph Affolderbach, Guan-Xiang Du, Thejesh Bandi, Andrew Horsley, Philipp Treutlein, Gaetano Miletta
1Universität Basel, Switzerland; 2Université de Neuchâtel, Switzerland

**87Rb Isoclinic Point Thermometry** ................................................................................................................................... 25
Nathan Wells, Travis Driskell, James Camparo
Aerospace Corporation, United States

**Spectroscopy and Hyperfine Clock Frequency Shift Measurements in Cs Vapor Cells Coated with Octadecyltrichlorosilanes (OTS)** ...................................................................................................................................... 33
Moustafa Abdel Hafiz, Vincent Maurice, Ravinder Chutani, Nicolas Passilly, Christophe Gorecki, Stéphane Guérandel, Emeric De Clercq, Rodolphe Boudot
1FEMTO-st Institute, France; 2Observatoire de Paris, France

**Buffer Gas Consumption in Rubidium Discharge Lamps** ............................................................................................. 37
Bernardo Jaduszliwer, Michael Huang, James Camparo
Aerospace Corporation, United States

Session: A1L-C: Applications of Optical Clocks

**mSTAR: Testing Special Relativity in Space Using High Performance Optical Frequency References** ..................... 47
Thilo Schuld, Shailendhar Saraf, Alberto Stochino, Klaus Döringhoff, Sasha Buchman, Grant D. Cutler, John Lipa, Si Tan, John Hanson, Belgaem Jaroux, Claus Braxmaier, Norman Gülrebeck, Sven Herrmann, Claus Lämmerza
1Airbus DS GmbH, Germany; 2German Aerospace Center, Germany; 3Humboldt-Universität zu Berlin, Germany; 4King Abdulaziz City for Science and Technology, Saudi Arabia; 5King Abdulaziz City for Science and Technology / Stanford University, Saudi Arabia; 6NASA Ames Research Center, United States; 7Stanford University, United States; 8Universität Bremen, Germany

**Geometrical Scale-Factor Stabilization of Square Cavity Ring Laser Gyroscopes** .......................................................... 51
Jacopo Belfi, Angela Di Virgilio, Nicolò Beverini, Giorgio Carelli, Enrico Maccioni, Andreino Simonelli, Rosa Santagata
1Istituto Nazionale di Fisica Nucleare, Italy; 2Università di Pisa, Italy; 3Università di Siena and Istituto Nazionale di Fisica Nucleare, Italy
Session: A2L-A: Small Scale Oscillators

Möbius Metamaterial Topology: Applications in Resonators and Tunable Oscillator Circuits ................................................. 56
Ajay Poddar², Ulrich Rohde¹
¹Brandenburgische Technische Universität / Synergy Microwave Corporation, Germany; ²Synergy Microwave Corporation, United States

Frequency Signal Source's PN (Phase Noise) Measurements: Challenges and Uncertainty ........................................... 62
Ulrich Rohde¹, Ajay Poddar², Enrico Rubiola², Marius Alexandru Silaghi⁴
¹Brandenburgische Technische Universität / Synergy Microwave Corporation, Germany; ²FEMTO-st Institute, France; ³Synergy Microwave Corporation, United States; ⁴Universitatea din Oradea, Romania

Model for Acoustic Locking of Spin Torque Oscillator ........................................................................................................ 68
Tanay Gosavi², Sunil Bhave¹
¹Analog Devices Inc, United States; ²Cornell University, United States

Piezoelectrically-Acutated Opto-Acoustic Oscillator ........................................................................................................ 72
Siddhartha Ghosh, Jeronimo Segovia-Fernandez, Gianluca Piazza
Carnegie Mellon University, United States

UHF SiGe Push-Pull VCO MEMS Oscillators .................................................................................................................. 76
Yeong Yoon, Harris Moyer, Deborah Kirby, Randall Kubena, Richard Joyce, Ross Bowen, Hung Nguyen, David Chang
HRL Laboratories, LLC, United States

Ultra-Low Phase Noise Frequency Synthesis Chains for High-Performance Vapor Cell Atomic Clocks ................. 81
Bruno François¹, Rodolphe Boudot¹, Claudio Eligio Calosso², Jean-Marie Danet³
¹FEMTO-st Institute, France; ²Istituto Nazionale di Ricerca Metrologica, Italy; ³Observatoire de Paris, Italy

Session: A2L-B: Sensors & Precision Measurements

Design of a Novel Length Extension Vibratory Gyroscope ................................................................................................. 84
Gobong Choi, Yook-Kong Yong
Rutgers University, United States

Remote Atomic Vapor Magneto meter with Sub-pT Resolution Operating at Ambient Temperature ..................... 90
Janet Lou¹, Fredrik Fatemi², Geoffrey Cranch²
¹Sotera Defense Solutions, Inc., United States; ²U. S. Naval Research Laboratory, United States

Session: A3P-D: Materials for Resonators

UHF Acoustic Attenuation and Quality Parameter Limits in the Diamond Based HBAR .............................................. 94
Arseniy Telichko¹, Boris Sorokin², Gennady Kvashnin³
¹Moscow Institute of Physics and Technology, Russia; ²Moscow Institute of Physics and Technology / Technological Institute for Superhard and Novel Carbon, Russia; ³Technological Institute for Superhard and Novel Carbon Materials, Russia

How to Qualify LGT Crystal for Acoustic Devices? ........................................................................................................ 100
Maroua Allani², Xavier Vacheret³, Alexandre Clairet³, Thomas Baron³, Jean-Jacques Boy⁴, C. Reibel⁵, O. Cambon⁶, Jean-Marc Lesage⁶, Olivier Bel⁸, Hugues Cabane⁶, C. Pecheyrnan⁷
¹Cristal Innov / Université Lyon1, France; ²DGA - Ministère de la Défense, France; ³FEMTO-st Institute, France; ⁴RAKON, France; ⁵Université de Montpellier, France; ⁶Université de Pau, France

Langasite Family Crystals as Promising Materials for Microacoustic Devices at Cryogenic Temperatures .......... 106
Andrey Sotnikov⁷, Elena Smirnova¹, Hagen Schmidt³, Manfred Weinhacht³, Jens Götze⁴, Sergey Sakharov⁵
¹A.F. Ioffe Physical-Technical Institute St. Petersburg, Russia; ²FOMOS-MATERIALS, OAO, Russia; ³Leibniz-Institut für Festkörper- und Werkstoffforschung, Germany; ⁴Technische Universität Bergakademie Freiberg, Germany

As-Doped Si's Complex Permittivity and its Effects on Heating Curve at 2.45 GHz Frequency ................................. 111
Siddharth Varadan, George Pan, Zhao Zhao, Terry Alford
Arizona State University, United States
Sputtered Al(l-x)ScxN Thin Films with High Areal Uniformity for Mass Production ................................................................. 117
Valeriy Felmetsger¹, Mikhail Mikhot¹, Mario DeMiguel-Ramos², Marta Clement², Jimena Olivares³, Teona Mirea³, Enrique Iborra³
¹OEM Group, United States; ²Universidad Politécnica de Madrid, Spain

Session: A3P-E: Oscillators, Synthesizers, Noise & Circuit Techniques I

Micro OXCO EWOS-0513: a 20 Years Space Odyssey Up to 67P/Churyumov-Gerasimenko .................................................. 121
Philippe Guillemot¹, Gilles Cibiel¹, Yves Richard², Jean-Marie Tarot², Guy Richard²
¹CNES - French Space Agency, France; ²Syrlinks, France

Noise Modeling Methodology of an Integrated Circuit for Quartz Crystal Oscillator ......................................................... 125
Nikolay Vorobyev², Joel Imbaud², Thomas Baron², Gilles Cibiel¹, Serge Galliou²
¹CNES - French Space Agency, France; ²FEMTO-st Institute, France

The Prediction, Simulation and Verification of the Phase Noise in Low-Phase-Noise Crystal Oscillator .............................. 129
Xianhe Huang, Junjie Jiao, Fuyu Sun, Wei Fu
University of Electronic Science and Technology of China, China

The Border Effect in Frequency Signal Processing and the Phase Measurement with Arbitrary Frequency Relationship .......... 133
Wei Zhou, Lina Bai, Zhiqi Li, Faxi Chen, Xiaotian Cao, Yadong Duan, Xuyang Zhou, Longfei Xu
Xidian University, China

A 250nm CMOS Low Phase Noise Differential VCO Circuit Without Varactors ................................................................. 136
Anatoly Kosykh, Konstantin Murasov, Alexandr Lepetaev, Sergey Zavyalov
Omsk State Technical University, Russia

Novel Gyroscopic Mounting for Crystal Oscillator (Payload) Applied in High Dynamic Host Vehicle (Platform) to Improve its Output Stability .......................................................... 139
Maryam Abedi, Tian Jin
Beihang University, China

Frequency Performance of the New Horizons Ultra-Stable Oscillators: Nine Years of Continuous in-Flight Monitoring .................. 145
Robert Jensen, Gregory Weaver
Johns Hopkins Applied Physics Laboratory, United States

A CMOS LC-Based Frequency Reference with ±40ppm Stability from -40°C to 105°C ......................................................... 151
Si-Ware Systems, Egypt

Effects of Pressure and Bias Voltage on the Phase Noise of CMOS-MEMS Oscillators .................................................... 155
Wan-Cheng Chiu, Ming-Huang Li, Chao-Yu Chen, Sheng-Shian Li
National Tsing Hua University, Taiwan

1/F Noise of Quartz Resonators: Measurements, Modelization and Comparison Studies .................................................. 158
Fabrice Sthal¹, Michel Deve⁵, Joel Imbaud², Roger Bourquin², Ahmed Bakir³, Cedric Vuillemín², Santunu Ghosh², Philippe Abbé², David Vernier², Gilles Cibiel¹
¹CNES - French Space Agency, France; ²FEMTO-st Institute, France

Session: A3P-F: Microwave Standards I

Estimation of the Light Shift in Ramsey-Coherent Population Trapping .......................................................... 162
Yuichiro Yano², Shigeyoshi Goka¹, Masatoshi Kajita¹
¹National Institute of Information and Communications Technology, Japan; ²Tokyo Metropolitan University, Japan

CPT Pulse Excitation Method Based on VCSEL Current Modulation for Miniature Atomic Clocks .......................... 167
Takumi Ide, Shigeyoshi Goka, Yuichiro Yano
Tokyo Metropolitan University, Japan
Preliminary Results of a Cs Vapor Cell CPT Clock Using Push-Pull Optical Pumping .............................................. 171
Moustafa Abdel Hafiz, Rodolphe Boudot
FEMTO-st Institute, France

Alkali Metal Source Tablet for Vapor Cells of Atomic Magnetometers ................................................................. 174
Kazuhiro Ban1, Akira Terao1, Natsuhiko Mizutani1, Kazuya Tsujimoto2, Yoshikazu Hirai2, Tetsuo Kobayashi2, Osamu Tabata2
1Canon Inc., Japan; 2Kyoto University, Japan

Alkali Metal Consumption by Discharge Lamps Fabricated from GE-180 Aluminosilicate Glass ......................... 180
Charles Klimcak, Michael Huang, James Camparo
Aerospace Corporation, United States

Mercury Lamp Studies in Support of Trapped Ion Frequency Standards .............................................................. 188
Lin Yi, Eric Burt, Robert Tjoelker
Jet Propulsion Laboratory / California Institute of Technology, United States

Session: A3P-G: Sensors III

Multimode SiC Trampoline Resonators Manipulate Microspheres to Create Chladni Figures ............................... 193
Hao Jia, Hao Tang, Philip Feng
Case Western Reserve University, United States

Calibrating Temperature Coefficient of Frequency (TCf) and Thermal Expansion Coefficient (Alpha) of MoS2 Nanomechanical Resonators ................................................................. 198
Rui Yang, Zenghui Wang, Philip Feng
Case Western Reserve University, United States

Performance Evaluation of CMOS-MEMS Thermal-Piezoresistive Resonators in Ambient Pressure for Sensor Applications ................................................................................................................................. 202
Jung-Hao Chang, Cheng-Syun Li, Cheng-Chi Chen, Sheng-Shian Li
National Tsing Hua University, Taiwan

Comparison of Acoustic Wave Pressure Sensors for TPMS Applications ............................................................... 205
Manohar Nagaraju1, Suresh Sriraman1, Andrew Lingley2, John Larson III1, Brian Otis2, Richard Ruby1
1Avago Technologies, United States; 2University of Washington, United States

Micromechanical Piezoelectric-on-Silicon BAW Resonators for Sensing in Liquid Environments ......................... 209
Abhinav Prasad, Ashwin Seshia, Jerome Charmet
University of Cambridge, United Kingdom

Stress Sensitivity Coefficients of HBAR .................................................................................................................. 214
Thomas Baron2, Valérie Petiri2, Gilles Martin2, Guillaume Combe2, Alexandre Clairet3, Bernard Dulmet2, Jean-Marc Lesage1, Thierry Laroche3, Sylvain Ballandras3
1DGA - Ministère de la Défense, France; 2FEMTO-st Institute, France; 3frec-n-sys SAS, France

A 400µW Differential FBAR Sensor Interface IC with Digital Readout ................................................................. 218
Manohar Nagaraju1, Suresh Sriraman1, Andrew Lingley2, Reed Parker2, Richard Ruby1, Brian Otis2
1Avago Technologies, United States; 2University of Washington, United States

Dual-Mode NEMS Self-Oscillator for Mass Sensing ............................................................................................... 222
Guillaume Gourlat, Marc Sansa, Guillaume Jourdan, Patrick Villard, Gilles Sicard, Sebastien Hentz
Commissariat à l’énergie atomique et aux énergies alternatives, France

Session: A3P-H: Timekeeping, Time & Frequency Transfer, GNSS Applications I

Uncertainty Evaluation of 2013 TL METODE Link Calibration Tour ................................................................. 226
Shinn-Yen Lin, Yi-Jiun Huang, Wen-Hung Tseng
Chunghwa Telecom Co., Ltd., Taiwan
Link Calibration or Receiver Calibration for Accurate Time Transfer ................................................................. 230
Zhiheng Jiang
International Bureau of Weights and Measures, France

The Performance Evaluation of the BD One-Way Time Service ........................................................................... 236
Wei Li, Wei Guang, Zhe Gao, Jihai Zhang, Yongliang Xu, Yajing Wei
National Time Service Center, CAS, China

Techniques of Antenna Cable Delay Measurement for GPS Time Transfer .......................................................... 239
Daniele Rovera¹, Michel Abgrall¹, Pierre Uhrich¹, Marco Siccardi²
¹Observatoire de Paris, France; ²SKK Electronics, Italy

Relative Calibration of Galileo Receivers Within the Time Validation Facility (TVF) ........................................ 245
Ricardo Píriz², Daniel Rodríguez², Pedro Rodalán², Alexander Mudrak¹, Andreas Bauch³, Julia Leute³, P. Pánek³,
Alexander Kuna³
¹European Space Agency / European Space Research and Technology Centre, Netherlands; ²GMV Innovating Solutions, Spain; ³Institute of Photonics and Electronics, Academy of Sciences CR, v.v.i., Czech Rep.; ⁴Physikalisch-Technische Bundesanstalt, Ger

A New Modem for Two Way Satellite Time and Frequency Transfer ................................................................. 250
Shengkang Zhang, Xueyun Wang, Haifeng Wang, Hongbo Wang, Yuan Yuan, Keming Feng
Beijing Institute of Radio Metrology and Measurement, China

SASO Time Scale and Measurement Capability ......................................................................................................... 254
Khalid Aldawood
Saudi Standards, Metrology and Quality Organization, Saudi Arabia

Stability Analysis of the French Timescale UTC(OP) ..................................................................................................... 257
Michel Abgrall, Sébastien Bize, Baptiste Chupin, Jocelyne Guéna, Philippe Laurent, Peter Rosenbusch, Pierre Uhrich,
Daniele Rovera
Observatoire de Paris, France

Preliminary Step for a UTC(It) Steering Algorithm Based on the ITCsF2 Primary Frequency Standard Measurements ..................................................................................................................... 260
Giovanna Signorile, Patrizia Tavella, Davide Calonico, Filippo Levi, Giovanni A. Costanzo, Giancarlo Cerretto, Roberto
Costa, Elena Cantoni, Ilaria Sesia
Istituto Nazionale di Ricerca Metrologica, Italy

Acquisition Method of Loran-C Signal Based on Matched Filter ........................................................................ 265
Yuanyuan Gao¹, Yu Hua¹, Yuanhong Cao², Haifeng Jiang¹
¹National Time Service Center, CAS, China; ²Sichuan Spaceon Time & Frequency Tech.Co., Ltd, China

Verification of Time Telegrams in Long Wave Radio Systems .................................................................................. 270
Matthias Schneider, Christoph Ruland
Universität Siegen, Germany

Two-Way Coherent Frequency Transfer in a Commercial DWDM Communication Network in Sweden ............... 276
Sven-Christian Ebenhag¹, Martin Zelan², Per Olof Hedekvist³, Magnus Karlsson¹, Börje Josefsson³
¹Chalmers University of Technology, Sweden; ²SP Technical Research Institute of Sweden, Sweden; ³Swedish University
Computer Network, Sweden

Frequency Distribution in Delay-Stabilized Optical DWDM Network Over the Distance of 3000 km ...................... 280
Lukasz Sliwczynski¹, Przemyslaw Krehlik¹, Marcin Lipinski¹, Krzysztof Turza², Artur Binczewski²
¹AGH University of Science and Technology, Poland; ²Poznań Supercomputing and Networking Center, Poland

The Research Progress of Two Way Time Synchronization with Fiber Based on Spread Spectrum Signal .......... 284
Xiangwei Zhu¹, Hang Gong², Guangfu Sun², Kun Liang¹
¹National Institute of Metrology, China; ²National University of Defense Technology, China
High Precise Time-Synchronization Based on Ultra-Short Pulse
Fan Shi, Shengkang Zhang, Huaiying Shang, Hongbo Wang, Haifeng Wang, Hang Yi, Zhenggang Ding, Feng Nian, Keming Feng
Beijing Institute of Radio Metrology and Measurement, China

Study on Autonomous and Distributed Time Synchronization Method for Formation UAVs
Tao Liu, Yonghui Hu, Yu Hua, Haifeng Jiang
National Time Service Center, CAS, China

Analysis of System Time Performance in BeiDou Satellite Navigation System
Jun Lu¹, Ye Ren², Xiaohui Li², Ya Liu², Shougang Zhang²
¹Beijing Institute of Tracking and Telecommunication Technology, China; ²National Time Service Center, CAS, China

Session: A3P-J: Optical Clocks

Ytterbium Optical Lattice Clock at INRIM
Marco Pizzocaro¹, Filippo Bregolin¹, Gianmaria Milani¹, Benjamin Rauf¹, Pierre Thoumany¹, Giovanni Antonio Costanzo², Filippo Levi¹, Davide Calonico¹
¹Istituto Nazionale di Ricerca Metrologica, Italy; ²Politecnico di Torino, Italy

Two Independent Strontium Optical Lattice Clocks for Practical Realization of the Meter and Secondary Representation of the Second
Michal Zawada², Marcin Bober², Piotr Morzynski², Agata Cygan², Daniel Lisak², Piotr Maslowski², Mateusz Prymaczek², Piotr Wcislo², Piotr Ablewski², Mariusz Piwinski², Szymon Wójtewicz², Katarzyna Bielska², Dobroslawa Bartoszek-Bobe
¹Jagiellonian University, Poland; ²Nicolaus Copernicus University, Poland; ³University of Warsaw, Poland

Session: A3P-K: Student Contest

A Magnetometer Based on Coherent Population Beating
Li Liu, Yigen Wang, Xiaona Zhao, Yuxin Zhuang, Zhong Wang
Peking University, China
Tuesday, April 14, 2015

Session: B1L-A: Sensors I

Monitoring the Adhesion Process of Tendon Stem Cells Using Shear-Horizontal Surface Acoustic Wave Sensors ................................................................. 310
Huiyan Wu, Hongfei Zu, Qing-Ming Wang, Guangyi Zhao, James H-C. Wang
University of Pittsburgh, United States

A Wireless Temperature Sensor Powered by a Piezoelectric Resonant Energy Harvesting System .................. 316
Peng Wang1, Robert Gray2, Zenghui Wang1, Philip Feng1
1Case Western Reserve University, United States; 2Case Western Reserve University / Hawken School, United States

Session: B1L-B: GNSS Development

The Time Validation Facility (TVF): an All-New Key Element of the Galileo Operational Phase .................. 320
Ricardo Píriz2, Daniel Rodríguez2, Pedro Roldán1, Alexander Mudrak1, Andreas Bauch5, Franziska Riedel5, Egle Staliuniene5, Francisco Javier Galindo6, Héctor Esteban5, Ilaria Sesia3, Giancarlo Cerretto5, Kenneth Jaldehag7, Carsten Rieck7, Pierre Uhrich4, Daniele Rovera1
1European Space Agency / European Space Research and Technology Centre, Netherlands; 2GMV Innovating Solutions, Spain; 3Istituto Nazionale di Ricerca Metrologica, Italy; 4Observatoire de Paris, France; 5Physikalisch-Technische Bundesanstalt, Germany; 6Presentación Real Instituto y Observatorio de la Armada, Spain; 7SP Technical Research Institute of Sweden, Sweden

Session: B3L-A: Phase Noise

Front-End Receiver: Recent and Emerging Trend .................................................................................. 326
Ulrich Rohde1, Ajay Poddar3, Enrico Rubiola2, Marius Alexandru Silaghi4
1Brandenburgische Technische Universität / Synergy Microwave Corporation, Germany; 2FEMTO-st Institute, France; 3Synergy Microwave Corporation, United States; 4Universitatea din Oradea, Romania

Oscillator Phase Noise: a 50-Year Retrospective .................................................................................. 332
David Leeson
Stanford University, United States

Least-Square Fit, Omega Counters, and Quadratic Variance .................................................................. 338
Francois Vernotte2, Michel Lenczner1, Pierre-Yves Bourgeois1, Enrico Rubiola1
1FEMTO-st Institute, France; 2Observatoire de Besançon / University of Franche-Comté, France

SDR and Self-Focusing Radar Techniques for milliHerz Measurement of Multi-Component Phase Noise Spectra ........................................................................................................ 343
Michael Underhill
Underhill Research Ltd, United Kingdom

Characterization of a Set of Cryocooled Sapphire Oscillators at the 10-16 Level with the Three-Cornered Hat Method .................................................................................................................. 343
Christophe Fluhr, Serge Grop, Timothée Accadia, Ahmed Bakir, Yann Kersalé, Enrico Rubiola, Vincent Giordano, Benoît Dubois
FEMTO-st Institute, France

Session: B3L-B: Atomic Laser Stabilization

Generating Entanglement Between Atomic Spins with Low-Noise Probing of an Optical Cavity ............... 351
Kevin Cox1, Joshua Weiner1, Graham Greve1, James Thompson2
1University of Colorado Boulder, United States; 2University of Colorado Boulder and NIST, United States

Laser Stabilization on Velocity Dependent Nonlinear Dispersion of Sr Atoms in an Optical Cavity .......... 357
Bjarke Takashi Rajl Christensen2, Stefan Alaric Schäffer2, Martin Romme Henriksen2, Philip Grabow Westergaard1, Jun Ye3, Jan Westenkaer Thomsen2
1Danish Fundamental Metrology, Denmark; 2Niels Bohr Institute, University of Copenhagen, Denmark; 3University of Colorado Boulder, United States
Ten Years of Active Optical Frequency Standards ................................................................. 363
Duo Pan\textsuperscript{2}, Wei Zhuang\textsuperscript{1}, Xiaobo Xue\textsuperscript{2}, Xiaogang Zhang\textsuperscript{2}, Mo Chen\textsuperscript{2}, Zhichao Xu\textsuperscript{2}, Jingbiao Chen\textsuperscript{2}
\textsuperscript{1}National Institute of Metrology, China; \textsuperscript{2}Peking University, China

Session: B3L-C: Ground & Space Time Scales

Precise Cascade Synchronization of Two Digitally Tuned Space Clocks to UTC (GPS) ........................................ 369
He Wang, Gebriel Iyanu
Aerospace Corporation, United States

Robust Clock Ensemble for Time and Frequency Reference System ......................................................... 374
Qinghua Wang, Fabien Droz, Pascal Rochat
Orolia Switzerland SA, Switzerland

A Status Report on Time Scale Generation in PTB .................................................................................. 379
Andreas Bauch\textsuperscript{2}, Egle Staliuniene\textsuperscript{2}, Gihan Gomah\textsuperscript{1}
\textsuperscript{1}National Institute for Standards, Egypt; \textsuperscript{2}Physikalisch-Technische Bundesanstalt, Germany

Makkah Timescale Generation and Measurement Capability .............................................................. 384
Yaseen M. Almleaky, Alaa Almleaky, Hamzah Almleaky, Samy Khadem-Al-Charieh
King Abdullah Centre for Crescent Observation & Astronomy, Saudi Arabia

Session: B4P-D: Acoustic Microresonators

Balanced Low-Loss 2-IDT Double Mode SAW Filter with Narrowed Passband and Improved Selectivity ........ 388
Sergei Doberstein
Omskiy Nauchno Issledovatelskiy Institut Priborostroeniya, Russia

Switchable and Tunable Resonators with Barium Strontium Titanate on GaN/Sapphire Substrates ............. 392
Thottam Kalkur\textsuperscript{3}, Milad Hmeda\textsuperscript{2}, Almonir Mansour\textsuperscript{2}, Pamir Alpay\textsuperscript{4}, Nick Sbockey\textsuperscript{1}, Gary Tompa\textsuperscript{2}
\textsuperscript{1}Structured Material Ind., United States; \textsuperscript{2}University of Colorado Boulder, United States; \textsuperscript{3}University of Colorado Colorado Springs, United States; \textsuperscript{4}University of Connecticut, United States

Resonant Transformation of Acoustic Waves Observed for the Diamond Based HBAR .............................. 396
Gennady Kvashnin\textsuperscript{3}, Boris Sorokin\textsuperscript{2}, Arseniy Telichko\textsuperscript{1}
\textsuperscript{1}Moscow Institute of Physics and Technology, Russia; \textsuperscript{2}Moscow Institute of Physics and Technology / Technological Institute for Superhard and Novel Carbon, Russia; \textsuperscript{3}Technological Institute for Superhard and Novel Carbon Materials, Russia

An Analysis of Thickness-Shear Vibrations of an Annular Plate with the Mindlin Plate Equations ............. 402
Ji Wang\textsuperscript{1}, Hui Chen\textsuperscript{1}, Tingfeng Ma\textsuperscript{1}, Jianke Du\textsuperscript{1}, Lijun Yi\textsuperscript{1}, Yook-Kong Yong\textsuperscript{2}
\textsuperscript{1}Ningbo University, China; \textsuperscript{2}Rutgers University, United States

Thickness-Shear Vibration Frequencies of an Infinite Plate with a Generalized Material Property Grading Along the Thickness ........................................................................................................ 406
Ji Wang, Wenliang Zhang, Dejin Huang, Tingfeng Ma, Jianke Du, Lijun Yi
Ningbo University, China

Wideband Ladder Filters Fully Covering Digital TV Band Based on Shear Horizontal Plate Wave ................. 412
Michio Kadota, Shuji Tanaka
Tohoku University, Japan

Enhancement of Effective Electromechanical Coupling Factor by Mass Loading in Layered SAW Device Structures .......................................................................................................................... 416
Gongbin Tang\textsuperscript{1}, Tao Han\textsuperscript{3}, Akihiko Teshigahara\textsuperscript{2}, Takao Iwaki\textsuperscript{2}, Ken-Ya Hashimoto\textsuperscript{1}
\textsuperscript{1}Chiba University, Japan; \textsuperscript{2}DENSO Corporation, Japan; \textsuperscript{3}Shanghai Jiao Tong University, China

Second Order Temperature Compensated Piezoelectrically Driven 23 MHz Heavily Doped Silicon Resonators with +/-10 ppm Temperature Stability ................................................................. 420
Antti Jaakkola, Panu Pekko, James Dekker, Mika Prunnila, Tuomas Pensala
VTT Technical Research Centre of Finland, Finland
Highly Tuneable X-Band Bragg Resonator - Initial Results ................................................................. 423
Pratik Deshpande, Simon Bale, Mark Hough, Jeremy Everard
University of York, United Kingdom

The Effect of Contour Concentricity on the Acceleration Sensitivity of Quartz Crystal Resonators .......... 427
Peter Morley
Vectron International, United States

Anchor Loss Suppression Using Butterfly-Shaped Plates for AlN Lamb Wave Resonators ...................... 432
Jie Zou¹, Chih-Ming Lin¹, Alber Pisano²
¹University of California, Berkeley, United States; ²University of California, San Diego, United States

Session: B4P-E: Oscillators, Synthesizers, Noise & Circuit Techniques II

Ultra-Low Noise All Fiber Mode-Locked Laser ..................................................................................... 436
Yaolin Zhang, Quansheng Ren, Shuangyou Zhang, Dong Hou, Jianye Zhao
Peking University, China

Digitally Temperature Compensated SAW Oscillator Based on the New Excitation Circuit ............... 439
Alexei Liashuk, Sergey Zavyalov, Alexandr Lepetaev, Anatoly Kosykh, Igor Khomenko
Omsk State Technical University, Russia

Phase Group Characteristics and Phase Coincidence Detection Based Phase Noise Measurement Method ...... 443
Shaofeng Dong, Wei Zhou, Wei Hu, Jinsong Zhan, Hongbo Qin
Xidian University, China

Precise Measurement of Complicated Frequency Signals ...................................................................... 445
Lina Bai, Meina Xuan, Yuzhen Jin, Bo Ye, Zhenjian Cui, Wei Zhou
Xidian University, China

Single-Bit-Output All-Digital Frequency Synthesis Using Multi-Step Look-Ahead Bandpass Sigma-Delta Modulator-Like Quantization Processing ............................................................................. 448
Charis Basetas, Paul Sotiriadis
National Technical University of Athens, Greece

Hardware Implementation Aspects of Multi-Step Look-Ahead Sigma-Delta Modulation-Like Architectures for All-Digital Frequency Synthesis Applications .................................................. 452
Charis Basetas, Anthimos Kanteres, Paul Sotiriadis
National Technical University of Athens, Greece

Session: B4P-F: Microwave Standards II

Compact Clocks for Industrial Applications: the EMRP Project IND 55 MClocks ........................................ 456
Salvatore Micalizio², Filippo Levi², Aldo Godone², Claudio Eligio Calosso², Bruno François¹, Stéphane Guérandel¹, David Holleville⁴, Emeric De Clercq⁴, Luigi De Sarlo⁴, Peter Yun⁴, Jean-Marie Danet⁴, Mehdi Langlois⁴, Rodolphe Boudo
¹FEMTO-st Institute, France; ²Istituto Nazionale di Ricerca Metrologica, Italy; ³Muquans, France; ⁴Observatoire de Paris, France; ⁵TÜBİTAK National Metrology Institute, Turkey; ⁶Université de Neuchâtel, Switzerland

Advances of Chip-Scale Atomic Clock in Peking University .................................................................... 462
Jianye Zhao, Yaolin Zhang, Haoyuan Lu, Dong Hou, Shuangyou Zhang, Zhong Wang
Peking University, China

An Atomic Frequency Micrometer Based on the Coherent Population Beating Phenomenon ................ 465
Zhong Wang, Jianye Zhao, Xiaona Zhao, Li Liu, Yuxin Zhuang, Dawei Li
Peking University, China

Digital Servo System Based on FPGA for Optically Pumped Magnetometer ........................................... 471
Sheng Zhou², Chang Liu², Yanhui Wang², Daoweng Zhang¹
¹North China Engineering Co. LTD, China; ²Peking University, China
Measuring Buffer-Gas Pressure in Sealed Glass Cells ................................................................. 474
Travis Driskell, Michael Huang, James Camparo
Aerospace Corporation, United States

Majorana Atomic Transition Research in H-Maser’s Magnetic State Selection Region .......................... 480
Aleynikov Mikhail
FGUP VNIIIFTRI, Russia

Noise Investigation on Optical Detection in a Cesium Beam Clock with Magnetic State Selection ........... 483
Chang Liu, Sheng Zhou, Yanhui Wang
Peking University, China

The Effect of Bend on the Ramsey Cavity .......................................................................................... 487
Fuyu Sun, Xianhe Huang
University of Electronic Science and Technology of China, China

Design of the New NIM6 Fountain with Collecting Atoms from a 3D MOT Loading Optical Molasses ........ 492
Fang Fang, Weiliang Chen, Nian Feng Liu, Kun Liu, Rui Suo, Tianchu Li
National Institute of Metrology, China

Advances in the Atomic Fountain Clock at SIOM .............................................................................. 495
Yuanbo Du, Rong Wei, Richang Dong, Fan Zou, Yuzhu Wang
Shanghai Institute of Optics and Fine Mechanics, China

Session: B4P-G: Sensors IV

Comparison of Frequency Estimators for Interrogation of Wireless Resonant SAW Sensors ................. 498
Victor Kalinin
Transense Technologies PLC, United Kingdom

Acoustic Power Gain Induced by 2D Electron Drifting ........................................................................ 504
Lei Shao, Kevin Pipe
University of Michigan, United States

Modelling and Control of a Travelling Wave in a Finite Beam, Using Multi-Modal Approach and Vector Control Method .................................................................................................................. 509
Sofiane Ghenna, Frédéric Giraud, Christophe Giraud-Audine, Michel Amberg, Betty Lemaire-Semail
Université Lille1, France

Measurement and Analysis of a Circular Wedge Acoustic Waveguide Using a PZT Sensor .................... 515
Tai-Ho Yu
National United University, Taiwan

Characterization and Temperature Sensor Application of Ca₃TaGa₃Si₂O₁₄ Crystals .............................. 518
Hongfei Zu¹, Huiyan Wu², Qing-Ming Wang³, Quanming Lin¹, Yanqing Zheng¹
¹Shanghai Institute of Ceramics, China; ²University of Pittsburgh, United States

Improvement in Tracking Loop Threshold of High Dynamic GNSS Receiver by Installation of Crystal Oscillator on Gyroscopic Mounting .......................................................................................... 522
Maryam Abedi, Tian Jin
Beihang University, China

Feasibility Study of Proximity Sensing by Using a Conventional Airborne Transducer ......................... 528
Ken Yamada, Shu Agatsuma
Tohoku-Gakuin University, Japan

Interrogation of Orthogonal Frequency Coded SAW Sensors Using the USRP ..................................... 530
James Humphries, Mark Gallagher, Daniel Gallagher, Arthur Weeks, Donald Malocha
University of Central Florida, United States
SH-SAW -- Based Sensor for Heavy Metal Ion Detection .......................................................... 536
Zeinab Ramshani, Binu B. Narakathu, Avuthu S. G. Reddy, Massood Z. Atashbar, Jared Wabeke, Sherine Obare
Western Michigan University, United States

Session: B4P-H: Timekeeping, Time & Frequency Transfer, GNSS Applications II

The Study of BeiDou Timing Receiver Delay Calibration ......................................................... 541
Hongbo Wang¹, Hang Yi¹, Shengkang Zhang¹, Hailfeng Wang¹, Fan Shi¹, Huaiying Shang¹, Yujie Yang¹, Jun Ge¹, Zhiqi Li²
¹Beijing Institute of Radio Metrology and Measurement, China; ²Xidian University, China

Developing of One Time Link Calibrator with GNSS at NIM ............................................... 545
Kun Liang², Aimin Zhang², Zhiquian Yang², Weibo Wang², Hang Yang¹
¹Beijing JiaoTong University, China; ²National Institute of Metrology, China

Discovery of Persistent Ionospheric Frequency Shifts of a Few Herz and Impact
on Time and Frequency Transfer .......................................................................................... 549
Michael Underhill
Underhill Research Ltd, United Kingdom

Research on Time and Frequency Transfer Based on BeiDou Common View ......................... 553
Hang Yi¹, Hongbo Wang¹, Shengkang Zhang¹, Hailfeng Wang¹, Fan Shi¹, Huaiying Shang¹, Jun Ge¹, Yujie Yang¹, Zhiqi Li²
¹Beijing Institute of Radio Metrology and Measurement, China; ²Xidian University, China

Preparing ACES-PHARAO Data Analysis ............................................................................. 557
Frédéric Meynadier, Pacôme Delva, Christophe Le Poncin-Laffite, Christine Guerlin, Philippe Laurent, Peter Wolf
Observatoire de Paris, France

Investigating the Correlation Between Hydrogen-Maser Clocks in the Same Place ................. 562
Chao Gao², Bo Wang², Xi Zhu², Tianchu Li¹, Lijun Wang²
¹National Institute of Metrology, China; ²Tsinghua University, China

LMJ Timing and Fiducial System: Overview of the Global Architecture and Performances ...... 565
Vincent Drouet, Michel Prat, Pierre Raybaut, Damien Sainte-Beuve
CEA-DAM, France

Practical Limitations of NTP Time Transfer ........................................................................... 570
Andrew N. Novick, Michael A. Lombardi
National Institute of Standards and Technology, United States

Precise Three-Channel Integrated Time Counter ..................................................................... 575
Ryszard Szplet, Pawel Kwiatkowski, Zbigniew Jachna, Krzysztof Różyć
Military University of Technology, Poland

A Fiber Link for the Remote Comparison of Optical Clocks and Geodesy Experiments .......... 579
Cecilia Clivati, Davide Calonico, Matteo Frittelli, Alberto Mura, Filippo Levi
Istituto Nazionale di Ricerca Metrologica, Italy

OPTIME - the System Grows - a New 330 km Line ................................................................. 583
Łukasz Bučzek¹, Jacek Kołodziej⁵, Przemysław Krehlik⁵, Marcin Lipinski¹, Łukasz Sliwczynski¹, Piotr Dunst², Dariusz
Lemanski³, Jerzy Nawrocki², Pawel Nogas³, Albin Czubla³, Artur Binczewski⁶, Wojbor Bogacki³, Piotr Osta³powicz⁶, Maciej
Sroinski⁴, Krzysztof Turza⁴, Waldemar Adamowicz⁵, Jacek Igaison⁵, Tadeusz Pawsza⁷, Janusz Pieczerek⁷, Michal
Zawada⁷
¹AGH University of Science and Technology, Poland; ²Astrogeodynamic Observatory, Poland; ³Central Office of
Measures, Poland; ⁴Nicolaus Copernicus University, Poland; ⁵Orange Polska S.A., Poland; ⁶Poznan Supercomputing and
Networking Center, Poland

Design of the Optical Fiber Transmission Link in a Femtosecond-Precision,
Fiber-Optic Timing Synchronization System ........................................................................... 587
Huaiying Shang, Shengkang Zhang, Fan Shi, Hongbo Wang, Hailfeng Wang, Hang Yi, Feng Nian, Keming Feng
Beijing Institute of Radio Metrology and Measurement, China
The Method of Determination of GEO Satellite Precise Clock Bias During Maneuvering ........................................ 591
Meifang Wu, Pei Wei, Xuhai Yang, Shougang Zhang
National Time Service Center, CAS, China

Session: B4P-J: Combs & Stable Lasers

Comparison of Different Carrier-Envelope Frequency Stabilization Methods for a High Performance DPSSL Frequency Comb .......................................................................................................................... 594
Stefan Kundermann, Steve Lecomte
Centre Suisse d’Électronique et de Microtechnique, Switzerland

Development of an Erbium-Fiber-Laser-Based Optical Frequency Comb at NTSC .................................................. 599
Yanyan Zhang¹, Lulu Yan¹, Songtao Fan¹, Long Zhang¹, Wenyu Zhao¹, Wenge Guo², Shougang Zhang¹, Haifeng Jiang³
¹National Time Service Center, CAS, China; ²Xi’an Shiyou University, China

High Spectral Purity Laser Characterization with a Self-Heterodyne Frequency Discriminator ............................. 602
Olivier Llopis, Zeina Abdallah, Vincent Auroux, Arnaud Fernandez
LAAS-CNRS, France

External Cavity Diode Laser with Long-Term Frequency Stabilization Based on Mode Boundary Detection ...... 606
Zhouchang Xu, Kaikai Huang, Xuanhui Guo
Zhejiang University, China

Faraday Anomalous Dispersion Optical Filter at 461nm Utilizing a Strontium Hollow Cathode Lamp .................. 611
Duo Pan, Xiaobo Xue, Xiang Peng, Jingbiao Chen, Hong Guo, Bin Luo
Peking University, China

A Cavityless Laser Using Cesium Cell with 459 nm Laser Pumping .......................................................................... 614
Xiaobo Xue, Duo Pan, Jingbiao Chen
Peking University, China

Active Optical Frequency Standard Based on Narrow Bandwidth Faraday Atomic Filter ........................................ 618
Xiaogang Zhang², Jingbiao Chen², Wei Zhuang¹
¹National Institute of Metrology, China; ²Peking University, China

All-Fiber Implementation of Modulation Transfer Spectroscopy for 4He Atoms ...................................................... 622
Wei Gong, Xiang Peng, Wenhao Li, Teng Wu, Haidong Wang, Jingbiao Chen, Hong Guo
Peking University, China

Large Waist Cavity for Ultra-Narrow Transition Spectroscopy ............................................................................ 625
Stefan Alaric Schäffer, Sigrid Skovbo Adersen, Bjarke Takashi Røjle Christensen, Jan Westenkær Thomsen
Niels Bohr Institute, University of Copenhagen, Denmark
Wednesday, April 15, 2015

Session: C1L-A: Sensors II

Resonant Infrared Detector Based on a Piezoelectric Fishnet Metasurface ............................................................. 630
Yu Hui, Zhenyun Qian, Matteo Rinaldi
Northeastern University, United States

NSPUDT Using C-Axis Tilted ScAlN Thin Film ............................................................. 633
Abhay Kochhar, Yasuo Yamamoto, Akihiko Teshigahara, Ken-Ya Hashimoto, Shuji Tanaka, Masayoshi Esashi
1Chiba University, Japan; 2DENSO Corporation, Japan; 3Tohoku University, Japan

Session: C1L-B: Space Clocks

GNSS RAFS Latest Improvements ................................................................................................................................. 637
Fabien Droz, Pascal Rochat, Sébastien Boillat, Batiste Scheidegger
Orolia Switzerland SA, Switzerland

Session: C1L-C: GNSS Time & Frequency Transfer

Use of Two Traveling GPS Receivers for a Relative Calibration Campaign Among European Laboratories ........ 643
Pierre Uhrich, Daniela Rovera, Baptiste Chupin, Francisco Javier Galindo, Héctor Esteban, Kenneth Jaldehag, Carsten Rieck, Andreas Bauch, Thomas Polewka, Giancarlo Cerretto, Gianluca Fantino, Ricardo Piriz
1GMV Innovating Solutions, Spain; 2Istituto Nazionale di Ricerca Metrologica, Italy; 3Observatoire de Paris, France; 4Physikalisch-Technische Bundesanstalt, Germany; 5Presentation Real Instituto y Observatorio de la Armada, Spain; 6SP Technical Research Institute of Sweden, Sweden

GPS Time Link Calibrations in the Frame of EURAMET Project 1156 ................................................................. 649
Héctor Esteban, Francisco Javier Galindo, Andreas Bauch, Thomas Polewka, Giancarlo Cerretto, Roberto Costa, Peter Whibberley, Pierre Uhrich, Baptiste Chupin, Zhiheng Jiang
1International Bureau of Weights and Measures, France; 2Istituto Nazionale di Ricerca Metrologica, Italy; 3National Physical Laboratory, United Kingdom; 4Observatoire de Paris, France; 5Physikalisch-Technische Bundesanstalt, Germany; 6Presentation Real Instituto y Observatorio de la Armada, Spain

Comparison of Two Continuous GPS Carrier-Phase Time Transfer Techniques ...................................................... 655
Jian Yao, Ivan Skakun, Zhiheng Jiang, Judah Levine
1Central Research Institute of Machine Building, Russia; 2International Bureau of Weights and Measures, France; 3National Institute of Standards and Technology / University of Colorado at Boulder, United States

Correction for Code-Phase Clock Bias in PPP .................................................................................................................. 662
Pascale Defraine, Jean-Marie Sleewaegen
1Royal Observatory of Belgium, Belgium; 2Septentrio Satellite Navigation NV, Belgium

Session: C2L-A: Digital Signal Processing

All Digital Frequency Synthesis Based on New Sigma-Delta Modulation Architectures ........................................ 667
Paul Sotiriadis
National Technical University of Athens, Greece

Noise in High-Speed Digital-to-Analog Converters ...................................................................................................... 672
Pierre-Yves Bourgeois, Takeshi Imaike, Gwenhaël Goavec-Merou, Enrico Rubiola
1FEMTO-st Institute, France; 2Nihon University, Japan

Simple Method for ADC Characterization Under the Frame of Digital PM and AM Noise Measurement .............. 676
Andrea Carolina Cárdenas-Olaya, Enrico Rubiola, Jean Michel Friedt, Massimo Ortolano, Salvatore Micalizio, Claudio Eligio Calosso
1FEMTO-st Institute, France; 2Istituto Nazionale di Ricerca Metrologica, Italy; 3Politecnico di Torino, Italy
6/12-Channel Synchronous Digital Phasemeter for Ultrastable Signal Characterization and Use

Massimo Caligaris1, Giovanni A. Costanzo1, Claudio Eligio Calosso1
1Istituto Nazionale di Ricerca Metrologica, Italy; 2Politecnico di Torino, Italy

Session: C2L-C: Emerging Time Dissemination Techniques

Time Signals Converging Within Cyber-Physical Systems

Marc Weiss1, Sundeep Chandhoke2, Hugh Melvin3
1National Institute of Standards and Technology, United States; 2National Instruments, United States; 3National University of Ireland, Galway, Ireland

Ns-Level Time Transfer Over a Microwave Link Using the PTP-WR Protocol

Mathieu Rico, Jean-Pierre Aubry, Cyril Botteron, Pierre-André Farine
École Polytechnique Fédérale de Lausanne, Switzerland

Precise UTC Dissemination Through Future Telecom Synchronization Networks

Wen-Hung Tseng, Sammy Siu, Shinn-Yan Lin, Chia-Shu Liao
Chunghwa Telecom Co., Ltd., Taiwan

Session: C3L-A: Aluminum Nitride MEMS Resonators

Gap Reduction Based Frequency Tuning for AlN Capacitive-Piezoelectric Resonators

Robert Schneider, Thura Lin Naing, Tristan Rocheleau, Clark Nguyen
University of California, Berkeley, United States

Switchable 2-Port Aluminum Nitride MEMS Resonator Using Monolithically Integrated 3.6 THz Cut-Off Frequency Phase-Change Switches

Gwendolyn Hummel, Matteo Rinaldi
Northeastern University, United States

Analysis of the Impact of Release Area on the Quality Factor of Contour-Mode Resonators by Laser Doppler Vibrometry

Brian Gibson1, Kamala Qalandar2, Kimberly Turner2, Cristian Cassella1, Gianluca Piazza1
1Carnegie Mellon University, United States; 2University California Santa Barbara, United States

Session: C3L-B: Cavity Laser Stabilization

Accurate Removal of Ram from FM Laser Beams

John Hall, Wei Zhang, Jun Ye
University of Colorado Boulder, United States

Session: C3L-C: Microwave Time & Frequency Transfer

Carrier Phase and Pseudorange Disagreement as Revealed by Precise Point Positioning Solutions

Demetrios Matsakis2, Zhiheng Jiang1, Wenjun Wu1
1International Bureau of Weights and Measures, China; 2United States Naval Observatory, United States

Long-Term Uncertainty in Time Transfer Using GPS and TWSTFT Techniques

Victor Zhang1, Thomas Parker1, Jian Yao2
1National Institute of Standards and Technology, United States; 2National Institute of Standards and Technology / University of Colorado at Boulder, United States

Session: C4L-A: Temperature Effects and Frequency Tuning in Resonators

Quality Factors of Quartz Crystal Resonators Operating at 4 Kelvins

Serge Galliou1, Philippe Abbé1, Maxim Goryachev2, Michael Tobar2, Roger Bourquin1
1FEMTO-st Institute, France; 2University of Western Australia, Australia
Session: C4L-B: Atomic Fountains

Bias Corrections in Primary Frequency Standards ................................................................. 733
Thomas Parker, Thomas Heavner, Steven Jefferts
National Institute of Standards and Technology, United States

Session: C4L-C: Optical Time Transfer in Telecommunication Networks

Transmission of a Frequency Channel Through a Long-Haul Optical Fiber Communications Link ............. 736
Curtis Menyuk
University of Maryland Baltimore County, United States

Preliminary Time Transfer Through Optical Fiber at NIM ........................................................... 742
Kun Liang², Aimin Zhang², Zhiqiang Yang², Weiliang Chen², Weibo Wang², Long Bai¹, Guitao Fu¹
¹Beijing Satellite Navigation Center, China; ²National Institute of Metrology, China

Actively and Passively Compensated RF Frequency Disseminations on Branching Fiber Network ............ 747
Bo Wang, Xi Zhu, Yu Bai, Chao Gao, Lijun Wang
Tsinghua University, China
Thursday, April 16

Session: D1L-A: Photonic Microwave Signal Generation

**Comparison of Self-ILPLL Forced Oscillators** ................................................................. 749
Tianchi Sun², Li Zhang², Kevin Receveur¹, Afshin Daryoush¹, Ajay Poddar³, Ulrich Rohde¹
¹Brandenburgische Technische Universität / Synergy Microwave Corporation, United States; ²Drexel University, United States; ³Synergy Microwave Corporation, United States

Session: D1L-B: Ion Microwave Clocks

**Miniature Trapped-Ion Frequency Standard with 171Yb+** ........................................................... 752
Peter D.D. Schwindt⁵, Yuan-Yu Jau⁵, Heather Partner⁴, Darwin Serkland⁴, Aaron Ison⁴, Andrew McCants⁵, Edward Winrow⁵, John Prestage¹, James Kellogg¹, Nan Yu¹, Dan Boschens², Igor Kosvins², David Mailloux², David Scherer², Craig Nelson³, Archita Hat³, David A. Howe³
¹Jet Propulsion Laboratory / California Institute of Technology, United States; ²Microsemi Inc., United States; ³National Institute of Standards and Technology, United States; ⁴Physikalisch-Technische Bundesanstalt / Sandia National Laboratories, United States; ⁵Sandia National Laboratories, United States

**Towards a High-Performance Microwave Frequency Standard Based on 113Cd+ Ions** ................... 758
Jianwei Zhang, Kai Miao, Lijun Wang, Xiaolin Sun, Lijun Wang
Tsinghua University, China

Session: D1L-C: Fiber Optic Time Transfer Technology

**ELSTAB - Electronically Stabilized Time and Frequency Distribution Over Optical Fiber - an Overview** ............................... 761
Przemyslaw Krehlik, Lukasz Sliwczynski
AGH University of Science and Technology, Poland

**Comparison of Forward- and Backward-Propagating Optical-Fiber-Induced Noise for Application to Optical Fiber Frequency Transfer** ................................................................. 765
James Cahill², Olukayode Okusaga¹, Weimin Zhou¹, Curtis Menyuk², Gary Carter²
¹U. S. Army Research Laboratory, United States; ²University of Maryland Baltimore County, United States

**The Optical Fiber Link LIFT for Radioastronomy** ........................................................................ 769
Cecilia Clivati², Roberto Ambrosini¹, Giovanni A. Costanzo², Matteo Frittelli², Filippo Levi², Alberto Mura², Federico Perini¹, Mauro Roma¹, Massimo Zucco², Davide Calonico²
¹Istituto Nazionale di Astrofisica, Italy; ²Istituto Nazionale di Ricerca Metrologica, Italy

**A Round-Trip Fiber-Optic Time Transfer System Using Bidirectional TDM Transmission** .................. 773
Guiling Wu, Liang Hu, Hao Zhang, Jianping Chen
Shanghai Jiao Tong University, China

Session: D2L-A: Materials for Acoustic Resonators

**Manufacturability of Highly Doped Aluminum Nitride Films** .......................................................... 777
Sergey Mishin, Yury Oshmyansky
Advanced Modular Systems, Inc, United States

**Observation of Strong Temperature Hysteresis in Molybdenum Disulfide (MoS2) Vibrating Nanomechanical Resonators** ................................................................. 783
Zenghui Wang, Rui Yang, Arnob Islam, Philip Feng
Case Western Reserve University, United States

**New Capacitive Micro-Acoustic Resonators Machined in Single-Crystal Silicon Stacked Structures** .......... 787
Nesrine Belkadi, Thomas Baron, Bernard Dulmet, Laurent Robert, Etienne Herth, Florent Bernard
FEMTO-st Institute, France

**Evaluation of Elastic Properties of SiO2 Thin Films by Ultrasonic Microscopy** .................................. 793
Kensuke Sakamoto¹, Tatsuya Omori¹, Jun-Ichi Kushibiki¹, Matsuda Matsuda², Ken-Ya Hashimoto²
¹Chiba University, Japan; ²Taiyo Yuden Ltd., Japan
Session: D2L-B: Vapor Cell Clocks

Study on Double-Modulation Coherent Population Trapping Resonance ................................................................. 797
Peter Yun, Sinda Mejri, Francois Tricot, David Holleville, Emeric De Clercq, Stéphane Guérandel
Observatoire de Paris, France

Compact and High-Performance Rb Clock Based on Pulsed Optical Pumping for Industrial Application ........... 800
Songbai Kang, Mohammadreza Gharavipour, Florian Gruet, Christoph Affolderbach, Gaetano Mileti
Université de Neuchâtel, Switzerland

Session: D2L-C: Free Space Optical Links

Identification and Calibration of Ground System Biases in Ground to Space Laser Time Transfer ...................... 804
Ivan Prochazka1, Josef Blazej1, Jan Kodet2
1Czech Technical University in Prague, Czech Rep.; 2Czech Technical University in Prague & Technical University Munich, Czech Rep.

Characterization of an Ultra Stable Quartz Oscillator Thanks to Time Transfer by Laser Link (T2L2, Jason-2) ... 808
Alexandre Belli2, Pierre Exertier2, Etienne Samain2, Clément Courde2, François Vernotte3, Albert Auriol1, Christian Jayles1
1CNES - French Space Agency, France; 2Géoazur, France; 3Observatoire de Besançon / University of Franche-Comté, France

Author Index...................................................................................................................................................................... 813
WELCOME FROM THE GENERAL CO-CHAIRS

We welcome your attendance to the 2015 Joint Conference of the IEEE International Frequency Control Symposium and the European Frequency and Time Forum. In each previous joint symposium, our community has received an enhanced experience from the extended technology outreach, exposure to international scope, and interaction with a greater majority of the world’s leading experts in time and frequency. We expect this year’s conference to perform the same enrichment to your professional and academic development.

We hope that you enjoy our sessions’ venue within the spacious accommodation of the Colorado Convention Center (CCC). We also encourage you to take advantage of our location within the Denver 16th Street Mall area. We are sure you will find a satisfying exploration of the large variety of restaurants, shopping, and nightlife activities through the convenience of the free MallRide shuttle bus. Please also take notice of the other highly active visitor regions within a few blocks of the CCC, including the Denver Performing Arts Complex, Golden Triangle Museum District, and Art District on Santa Fe.

There are two main social events for this conference. On Monday evening, April 13th, a Conference Banquet will be conducted at the Wings over the Rockies Air and Space Museum. Attendees will enjoy dinner and dancing within the truly historic hangar of the former Lowry Air Force Base. At the “Wings” be prepared to take advantage of flight simulators, antique aircraft displays and a fascinating collection of avionic and ground support equipment, such as early radio, telemetry and radar systems – exhibits we believe of high interest to you and your guests. On Tuesday evening, April 14th, we will give special attention to our exhibitors, who will host the Exhibitors’ Reception in the Four Seasons Ballroom. The Exhibitors’ Reception will be collocated with the exhibit booths to arrange dedicated time for extended conversations with our exhibitors.

We are honored that Dr. Jian-yu Lu, President of the IEEE Ultrasonics, Ferroelectrics, and Frequency Control Society (IEEE UFFC) is attending the conference. Dr. Lu has organized a breakfast reception for all registered students on Tuesday, April 14th from 7:30 to 8:30 AM in the Four Season’s Ballroom of the CCC. At the breakfast reception, Dr. Lu will offer additional recognition to our Student Paper awardees, who will be first announced at Monday night’s Welcome Reception by the Academic Chair Dr. Clemens Ruppel.

The conference is proud to have Dr. Patricia Rankin as the guest speaker for the IEEE Women in Engineering (WIE) lunch to be held on Tuesday, April 14 from 12:15 to 1:15 PM. Prof. Rankin is a recognized author and leader in the area of faculty development and the advancement of women in science and engineering. The WIE lunch is complimentary to all women attendees active in the technical areas of the conference.

We recognize many organizations and sponsors whose contributions and participation elevate our conference to a truly collegial event. These include NIST-Boulder Time and Frequency Division, JILA of the Colorado University, CCTF, and the IEEE.

Again welcome to Denver and best wishes for a productive and enjoyable conference!

Ekkehard Peik and Gregory Weaver
General Co-Chairs for the 2015 Joint IEEE IFCS-EFTF
It is our pleasure to welcome you to the 2015 Joint IFCS/EFTF in the Mile-High City. Since the first joint conference in Besancon, France in 1999, this now biennial joint conference has been a great success with increased attendance and high quality of the papers. We continue the tradition of success with the volunteer leadership and tremendous dedication of the General Co-Chairs, Gregory Weaver and Ekkehard Peik, and the competence of the other Conference Committee Chairs. We are in debt to the TPC Vice-Chairs of the six topical groups and the TPC members for their efforts and time in reviewing the abstracts and overall support to build up an attractive conference program. We thank everybody for their time and dedication. Most of all, foremost contribution is from all of the authors who present latest and exciting results in their lectures, posters and manuscripts.

This year, Judah Levine, of the National Institute of Standards and Technology (NIST), will give the plenary talk entitled "Don't tell me how to build a clock; just tell me how to find out what time it is" on distributing time and frequency information.

The regular technical sessions occurring over three and a half days will consist of more than 160 lectures and 210 posters, including sixteen invited lectures. The 30 regular oral sessions will be split in 3 parallel tracks. The posters will remain during the whole conference presented by their authors, on Monday or Tuesday afternoon. Twenty-five student finalists representing all six topical groups were selected among more than 85 abstracts. The winners will be awarded during the Conference Banquet on Monday night.

On Tuesday, Wan-Thai Hsu, Micrel Inc will moderate the Entrepreneurs Forum entitled "A Users’ Perspective and Expectation in a Connected Society". In parallel, we will have an exceptional oral session with three invited talks related to the topic: "Time & Frequency and Fundamental Physics".

On Sunday, we will have a day dedicated to tutorials, with a total of twelve world-renowned expert lecturers. The exhibition will be animated by the leading manufacturers and suppliers of frequency control products and equipment from around the world.

Finally, we will have unique opportunities thanks to free services offered by our colleagues from NIST and JILA for the benefit of this community: 1) Cross-Spectrum L(ƒ) Workshop on Wednesday evening; 2) NIST and JILA tour in Thursday afternoon. Their minimal registration fees for these space-limited events will cover the room, food, and transportation. We strongly recommend to all of the attendees to take advantage of these unique opportunities.

We are confident that all of us will enjoy this conference in every aspect and ultimately satisfy our brains' demand to learn!

Yoonkee Kim and Gaetano Mileti
2015 Joint IFCS/EFTF JTPC Co-Chairs
IEEE IFCS-EFTF 2015 ORGANIZING COMMITTEE

General Co-Chair:
Gregory Weaver
JHU Applied Physics Laboratory
gregory.weaver@jhuapl.edu

General Co-Chair:
Ekkehard Peik
Physikalisch-Technische Bundesanstalt
Ekkehard.Peik@ptb.de

Technical Program Committee Co-Chair:
Yoonkee Kim
USARMY CERDEC
yoonkee.kim.civ@mail.mil

Technical Program Committee Co-Chair:
Gaetano Mileti
Université de Neuchâtel
gaetano.mileti@unine.ch

Finance Chair:
Debra Coler
OEwaves
Debra.Coler@oewaves.com

Academic Chair:
Clemens Ruppel
EPCOS AG
clemens.ruppel@epcos.com

Editorial Chair:
Aaron Partridge
SiTime
ap@sitime.com

Tutorial Co-Chair:
Gianluca Piazza
Carnegie Mellon University
piazza@ece.cmu.edu
**Tutorial Co-Chair:**
Jeremy Everard
University of York
jeremy.everard@york.ac.uk

**Conference Management:**
Conference Catalysts, LLC
laurenp@conferencecatalysts.com

**Exhibits Management:**
Sue Kingston
skingston1514@gmail.com
IFCS-EFTF 2015 JOINT TECHNICAL PROGRAM COMMITTEE

Group 1: Materials, Filters, and Resonators

Emmanuel Defay, CEA, France
Gianluca Piazza, Carnegie Mellon University, USA

Reza Abdolvand, University of Central Florida, USA
Sarah Bedair, US Army Research Labs, USA
Bernard Dulmet, FEMTO-ST, France
Marc Faucher, IEMN, France
Mina Rais-Zadeh, University of Michigan, USA
Randy Kubena, HRL Laboratories, USA
Jan Kuypers, Qorvo, USA
Sheng-Shian Li, National Tsing Hua University, Taiwan
Olivier Le Traon, Onera, France
Bernd W. Neubig, Advanced Crystal Products, Germany
Clark Nguyen, University of California at Berkeley, USA
Derek (Rick) Puccio, Quartzdyne, USA
Alexandre Reinhardt, CEA-LETI, France
Ashwin A. Seshia, University of Cambridge, United Kingdom
Dan Stevens, Consultant, USA
Shuji Tanaka, Tohoku University, Japan
Ventsislav Yantshev, Uppsala University, Sweden
Ji Wang, Ningbo University, China
Dana Weinstein, MIT, USA
Yook-Kong Yong, Rutgers University, USA

Group 2: Oscillators, Synthesizers, Noise, and Circuit Techniques

Fabrice Sthal, FEMTO-ST, France
Jean-Pierre Aubry, Aubry Conseil, China

Martin Bloch, Frequency Electronic Inc, USA
Claudio Calosso, INRIM, Italy
Gilles Cibiel, CNES, France
Michael Driscoll, Consultant, USA
Jeremy Everard, University of York, United Kingdom
Marvin Frerking, Innovative Technology Products, USA
Serge Gailliou, FEMTO-ST, France
Patrick Green, Northrop Grumman, USA
David Howe, NIST, USA
Xianhe Huang, Chengdu University, China
Wan-Thai Hsu, Micrel Inc, USA
Eugene Ivanov, University of Western Australia, Australia
Eun Sok Kim, USC, USA
Olivier Lopis, LASS, France
Takeo Oita, The University of Tokyo, Japan
Stephen Parker, The University of Western Australia, Australia
Ajay Poddar, Synergy Microwave Corporation, USA
Enrico Rubiola, FEMTO-ST, France
Kia Hock Tan, Universiti Tunku Abdul Rahman, Malaysia
Steve Tanner, IMT-EPFL, China
Michael Tobar, The University of Western Australia, Australia
Mike Underhill, Underhill Research, United Kingdom
Wei Zhou, Xidian University, China
Warren Walls, US Naval Observatory, USA
Group 3: Microwave Frequency Standards

Elizabeth Donley, NIST, USA
Krzysztof Szymaniec, NPL, United Kingdom
Patrick Berthoud, Oscilloquartz SA, USA
Eric Burt, Jet Propulsion Laboratory, USA
Christopher Ekstrom, U.S. Naval Observatory, USA
Fang Fang, National Institute of Metrology (NIM), China
Qingha Wang, Spectratime, Switzerland
Kurt Gibble, Penn State, USA
Yuko Hanado, NICT, Japan
Motohiro Kumagai, NICT, Japan
John Kitching, NIST, USA
Svenja Knappe, NIST, USA
Taeg Yong Kwon, KRISS, Korea
Arnaud Landragin, SYRTE, France
Steve Lecomte, CSEM, China
Filippo Levi, INRIM, Italy
Tianchu Li, NIM, China
Robert Lutwak, DARPA, USA
Lute Maleki, OEwaves, USA
Louis Marmet, INMS, NRC Canada, Canada
Salvatore Micalizio, INRIM, Italy
Fritz Riehle, Physikalisch-Technische Bundesanstalt, Germany
Peter Rosenbusch, SYRTE, France
Christophe Salomon, LKB/ENS, France
Robert Tjoelker, Jet Propulsion Laboratory, USA
Stefan Weyers, PTB, Germany

Group 4: Sensors and Transducers

Svenja Knappe, NIST, USA
Leonhard Reindl, University of Freiburg, Germany

Jeff Andle, Synergistic Science, USA
Sylvain Ballandras, FEMTO-ST, France
Alfred Binder, CTR AG, Austria
Mark Cheng, Wayne State University, USA
Weileun Fang, National Tsing-Hua University, Taiwan
Philip Feng, Case Western Reserve University, USA
Jean-Michel Fried, FEMTO, France
Jackie Hines, Applied Sensor Research, USA
Fritze Holger, TU Clausthal Germany, Germany
Fabien Josse, Marquette University, USA
Diethelm Johannsmann, TU Clausthal Germany, Germany
Shigeru Kurosawa, Adv. Industrial Science & Tech, Japan
Olivier Le Traon, ONERA, France
Ryszard Lec, Drexel University, USA
Ralf Lucklum, Otto-von-Guericke-University, Germany
Donald Malocha, University of Central Florida, USA
Glen McHale, Nottingham Trent University, United Kingdom
Paul Muralt, EPFL, Switzerland
Mauricio Pereira da Cunha, University of Maine, USA
Víctor Plessky, GVR Trade SA, Switzerland
Matteo Rinaldi, Northeastern University, USA
Clemens Ruppel, EPCOS AG, Germany
Ashwin Seshia, University of Cambridge, United Kingdom
Isao Shimoyama, University of Tokyo, Japan
Guillermo Villanueva, Ecole Polytechnique Federal de Lausanne (EPFL), Switzerland
Group 5: Timekeeping, Time and Frequency Transfer, GNSS and Applications
Pascale Defraigne, Royal Observatory of Belgium, Belgium
Pierre Waller, ESA, Netherlands
Andreas Bauch, PTB, Germany
Laurent-Guy Bernier, METAS Swiss Federal Office of Metrology, China
Javier de Vicente, ESA-ESOC, Germany
Jérôme Delporte, CNES, France
Gesine Grosche, PTB, Germany
Philippe Guillemot, CNES, France
Miho Fujieda, NICT, Japan
Jorg Hahn, ESA, Netherlands
Per Olof Hedekvist, SP Technical Research Institute of Sweden, Sweden
Michito Imae, AIST, Japan
Jan Johansson, SP Technical Research Institute of Sweden, Sweden
Judah Levine, National Institute of Standards and Technology, USA
Huang-Tien Lin, Chunghwa Telecom Co., Ltd., Taiwan
Shinn-Yan Lin, Chunghwa Telecom Co., Ltd., Taiwan
Xiao Chun Lu, NTSC, China, China
Demetrios Matsakis, U.S. Naval Observatory, USA
Dirk Piester, PTB, Germany
Vitaly Pal'chikov, VNIIFTR,
Ed Powers, U.S. Naval Observatory, USA
Wolfgang Schaefer, TimeTech GmbH, Germany
Amitava Sen Gupta, NPLI India, India
Samuel Stein, Symmetricom, Inc, USA
Patrizia Tavella, INRIM, Italy
Philip Tuckey, LNE-SYRTE, Paris Observatory, France
Pierre Uhrich, SYRTE, France
Bruce Warrington, National Measurement Institute, Australia
Aimin Zhang, NIM, China
Victor Zhang, NIST, USA
Michael Wouters, National Measurement Institute, Australia
Group 6: Optical Frequency Standards:

Sébastien Bize, SYRTE-Observatoire de Paris, France
Andrew Ludlow, NIST, USA

Luigi Cacciapuotti, European Space Agency, Italy
Davide Calonico, INRIM, Italy
James Chou, NIST, USA
Roman Ciurylo, Nicolaus Copernicus University, Poland
Andrei Derevianko, University of Nevada, USA
Pierre Dube, NRC, Canada
Kelvin Gao, Wuhan Institute of Physics and Mathematics/The Chinese Academy of Science, China
Feng-Lei Hong, NMIJ, Japan
Kazumoto Hosaka, NMIJ, Japan
Tetsuya Ido, NICT, Japan
Jason Jones, University of Arizona, USA
Nikolai Kolachevsky, Lebedev Institute of the Russian Academy, Russia
Yann Le Coq, SYRTE, France
Steve Lecomte, CSEM, Switzerland
Dave Leibrandt, NIST, USA
Christian Lisdat, PTB, Germany
Andre Luiten, University of Adelaide, Australia
Long-Sheng Ma, East China Normal University, China
Helen Margolis, NPL, United Kingdom
Mikko Merimaa, MIKES, Finland
John McFerran, University of West Australia, Australia
Ekkehard Peik, PTB, Germany
Ivan Prochazka, Electronics, ELT Prague, Czech Republic
Marianna Safronova, University of Delaware, USA
Thomas Südmeyer, Université de Neuchâtel, Switzerland
Alexey Taichenachev, Inst. of Laser Physics, Russia
Jun Ye, JILA, University of Colorado, USA
SPECIAL THANKS

The 2015 Joint Conference of the IEEE International Frequency Control Symposium & European Frequency and Time Forum is possible with help from:

SPONSORS

IEEE

EFTF

UFFC

PATRONS

NIST
National Institute of Standards and Technology
U.S. Department of Commerce
EXHIBITORS

Advanced Modular Systems
http://www.amssb.com/
Booth 18

Berkeley Nucleonics Corporation
http://www.berkeleynucleonics.com/
Booth 8

Clepsydra Time Systems
http://www.clepsydratime.com/
Booth 17

Crystalline Mirror Solutions
http://www.cristallinemirrors.com/
Booth 21

Frequency Electronics
http://www.frequencyelectronics.com/
Booth 14

GuideTech
http://www.guidetech.com/
Booth 22

Holzworth Instrumentation Inc.
http://www.holzworth.com/
Booth 1

IEEE
https://www.ieee.org/
Lobby

Journal of Micromechanics and
Microengineering
http://iopscience.iop.org/0960-1317/
Booth 12

M Squared Lasers
http://www.m2lasers.com/
Booth 26

Menlo Systems
http://www.menlosystems.com/
Booth 20

Microsemi Corporation
http://www.microsemi.com/
Booth 24
MicroSystems
http://microsystems.de/
Booth 25

MUQUANS
http://www.muquans.com/
Booth 6

NEL Frequency Controls, Inc.
http://www.nelfc.com/
Booth 15

Noise XT
http://noisext.com/
Booth 23

Onefive GmbH
http://www.onefive.com/
Booth 16

Oscilloquartz
http://www.oscilloquartz.com/
Booth 2

Piktime Systems
http://piktime.com/?sl=EN
Booth 17

Saunders & Associates
http://www.saunders-assoc.com/
Booth 4

scia Systems GmbH
http://www.scia-systems.com/
Booth 9

Chengdu Spaceon Electronics
http://www.elecspn.com/
Booth 19

SpectraDynamics, Inc.
http://www.spectradynamics.com/
Booth 10

Spectratime
http://www.spectratime.com/
Booth 7

Stable Laser Systems
http://www.stablelasers.com/
Booth 27
The 2015 C.B. Sawyer Award

Dr. Wan-Thai Hsu, Micrel Inc., USA

“Co-founding Discera, Inc., and pioneering the development and commercialization of MEMS oscillators”

The 2015 W.G. Cady Award

Dr. Ajay K. Poddar, Synergy Microwave Corporation, USA

“For the analysis, design, and development of a host of frequency control products exhibiting state-of-the-art performance, including the development of extremely low noise crystal oscillator circuitry”

The 2015 I.I. Rabi Award

Dr. Ulrich L. Rohde, Brandenberg University of Technology, Germany

“For intellectual leadership, selection and measurement of resonator structures for implementation in high performance frequency sources, essential to the determination of atomic resonance”

IFCS 2016 Award Nominations

Nominations are now open for the 2016 IFCS Awards. Nominations should be sent to the IFCS Awards Chair at gregory.weaver@jhuapl.edu. Information is also available at the Registration Desk, and on the IEEE Frequency Control Symposium website at: http://www.ieee-uffc.org/frequency-control/awards.asp.
EFTF 2015 AWARDS

The following three awards of the European Frequency and Time Forum will be presented at IFCS-EFTF 2015 to recipients selected by the Executive Committee of the EFTF:

The European Frequency and Time Award recognises outstanding contributions in all fields covered by the EFTF.

European Frequency and Time Award 2015

Giorgio Santarelli, Laboratoire Photonique, Numérique et Nanosciences, Universités de Bordeaux, France

“For his outstanding contributions to the development of ultra-low phase noise methods and systems for microwave and optical atomic clocks, of fiber-based optical frequency combs and of optical fiber links”

The EFTF Young Scientist Award is conferred in recognition of a personal contribution that has demonstrated a high degree of initiative and creativity and led to already established or easily foreseeable outstanding advances in the field of time and frequency metrology. The award honours a person under the age of 40 at the date of the opening session of the EFTF conference.

EFTF Young Scientist Award 2015

Franklyn J. Quinlan, NIST, Boulder, USA

“For seminal contributions to the understanding of fundamental noise processes in the photo detection of short optical pulses and the realization of very pure microwave signals using femtosecond frequency combs”

The Marcel Ecabert Award of the EFTF is a lifetime award and honours the excellent achievements of the recipient or an institution in the field of time and frequency. It is named after the late Marcel Ecabert, founding member of the EFTF and member of its Executive Committee.

Marcel Ecabert Award 2015

Bernd Neubig, AXTAL GmbH, Lobbach, Germany

“For almost 40 years of scientific advances in piezoelectric frequency control devices, his business acumen and his efforts in the international standardization of piezoelectric devices”

The awards are sponsored by the Société Française des Microtechniques et de Chronométrie.
This year, students have been encouraged to enter their papers in a Student Paper Competition. From the nearly 80 papers submitted to the competition, 3 or 4 from each topical group have been selected as finalists. From these finalists one winner will be chosen for each group. Judging of the winners will be based on:

(1) clarity of student's presentation  
(2) depth of student's knowledge  
(3) degree of the student's contribution to the project  
(4) relevancy of the work to the field

The finalists’ papers will be on display in a specific location in the Poster Session area. Judging will be held during the first Poster Session on Monday. Winners will be announced at the banquet on Monday evening at 19:00. The prefix number indicates the poster position.

**Student Paper Competition Finalists**

**Group 1: Materials, Resonators, & Resonator Circuits**

(200) Robert Schneider, University of California, Berkeley, USA  
"Gap Reduction Based Frequency Tuning for AlN Capacitive-Piezoelectric Resonators"

(201) Zhenyun Qian, Northeastern University, USA  
"GHz Range Graphene-Aluminum Nitride Nano Plate Resonators"

(202) Gwendolyn Hummel, Northeastern University, USA  
"Switchable 2-Port Aluminum Nitride MEMS Resonator Using Monolithically Integrated 3.6THz Cut-Off Frequency Phase-Change Switches"

(203) Mario Demiguel-Ramos, ETSI de Telecomunicación Universidad Politécnica de Madrid, Spain  
"Tungsten Oxide As High Acoustic Impedance Material for Fully Insulating Acoustic Reflectors"

**Group 2: Oscillators, Synthesizers, Noise, & Circuit Techniques**

(204) Tanay Gosavi, Cornell University, USA  
"Model for Acoustic Locking of Spin Torque Oscillators"

(205) Sidhartha Ghosh, Carnegie Mellon University, USA  
"Piezoelectrically-Actuated Opto-Acoustic Oscillator"

(206) Tianchi Sun, Drexel University, USA,  
"Comparison of Self-ILPLL Forced Oscillators"

(207) Guillaume William Bres-Saix, FEMTO-ST, France/National Institute of Standards and Technology, USA  
"A Zynq Based Digital Phase and Amplitude Measurement System"
Group 3: Microwave Frequency Standards

(208) Andrew Horsley, University of Basel, Switzerland
“Sub-100 µm Resolution Imaging of dc and Microwave Magnetic Fields Using Atomic Vapor Cells”

(209) Liu Li, Peking University, China
“A Magnetometer Based on Coherent Population Beating”

(210) Aaron Bennett, The Pennsylvania State University, USA
“Precision Measurements of Quantum Scattering Phase Shifts Through Feshbach Resonances”

(211) Greg Hoth, NIST, USA
“A Compact Atom Interferometer Based on an Expanding Ball of Atoms”

Group 4: Sensors & Transducers

(212) Yu Hui, Northeastern University, USA
“Resonant Infrared Detector Based on a Piezoelectric Fishnet Metasurface”

(213) Taimur Aftab, University of Freiburg, Germany
“A Novel Microwave Reflector-Antenna As a Resonant Wireless Passive Mechanical Sensor”

(214) Peng Zeng, Wayne State University, USA
“Wireless Monitoring of Eye Intraocular Pressure Using Transparent Graphene LC Sensors”

(215) Huiyan Wu, University of Pittsburgh, Pittsburgh, PA/USA
“Monitoring Tendon Stem Cell (TSC) Adhesion by Shear-Horizontal Surface Acoustic Wave Sensors”

Group 5: Timekeeping, Time and Frequency Transfer, GNSS Applications

(216) Chaoqun Ma, East China Normal University, China
Delivery of Optical Frequency via 34 km Urban Fiber Link with a Linewidth Broadening of 1 mHz”

(217) James Cahill, University of Maryland, USA
“Comparison of Forward- and Backward-Propagating Optical-Fiber-Induced Noise for Application to Optical Fiber Frequency Transfer”

(218) Wei Huang, Royal Observatory of Belgium, Belgium
“CGTTS results with Beidou using the R2CGTTS”

Group 6: Optical Frequency Standards and Applications

(219) Kwangyun Jung, Korea Advanced Institute of Science and Technology (KAIST), South Korea
“Repetition Rate Stabilization of an Optical Frequency Comb to a High Quality Factor Optical Fiber Delay Line”

(220) Bjarke Takashi Rojle Christensen, University of Copenhagen, Denmark
“Laser Stabilization on Velocity Dependent Nonlinear Dispersion of Sr Atoms in an Optical Cavity”

(221) Joe Becker, NIST, Boulder, USA
“Toward Chip Integrated Ultra-Low-Noise Lasing Using a Microrod Resonator”

(222) Giacomo Bolognesi, INRIM, Istituto Nazionale di Ricerca Metrologica, Italy
“Yellow Solid State Laser for Yb Atomic Clock Application”
FUTURE SYMPOSIA

2016 IEEE International Frequency Control Symposium
May 9 – 12, 2016 | The Roosevelt Hotel | New Orleans, Louisiana, USA

FIRST CALL FOR PAPERS

ABSTRACT SUBMISSION DEADLINE: JANUARY 15, 2016

The IFCS has chosen the Roosevelt Hotel in New Orleans, LA, USA as the venue for our 2016 symposium. The Roosevelt Hotel is located near the French Quarter and is within walking distance of the city’s major attractions, offering pedestrian commerce, restaurants and nightlife.

General Chair:
Lute Maleki
OEwaves
lute.maleki@oewaves.com

Technical Program Chair:
Elizabeth Donley
NIST
elizabeth.donley@nist.gov

Finance Chair:
Debra Coler
OEwaves
debra.coler@oewaves.com

Editorial Chair:
Aaron Partridge
SiTime
ap@sitime.com

Tutorial Chair:
Gianluca Piazza
Carnegie Mellon University
piazza@ece.cmu.edu

Abstracts will be collected through a web-based submission tool. Authors are invited to submit abstracts of recent and original work of interest to the frequency control communities in the following topics:

Group 1: Materials, Resonators, & Resonator Circuits
Group 2: Oscillators, Synthesizers, Noise, & Circuit Techniques
Group 3: Microwave Frequency Standards
Group 4: Sensors & Transducers
Group 5: Timekeeping, Time and Frequency Transfer, GNSS Applications
Group 6: Optical Frequency Standards and Applications
We are very pleased to announce that the next European Frequency and Time Forum (EFTF) will be held at the University of York in the UK. The venue enables extensive technical and social interaction as the presentations, exhibitions, catering and accommodation are all within a few 100 metres. The venue is also very close to the beautiful historic city of York (10 minutes by bus).

EFTF, as you know, is an international conference and exhibition, providing information on the latest advances and trends of scientific research and industrial development in the fields of Frequency and Time. We look forward to seeing you in April 2016.

Contact: Jeremy Everard, jeremy.everard@york.ac.uk

www.eftf.org
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 – 10:30</td>
<td><strong>MEMS Oscillators</strong>&lt;br&gt;Specifications, Applications, and System Perspectives&lt;br&gt;Dr. Aaron Partridge, SiTime</td>
</tr>
<tr>
<td></td>
<td><strong>Frequency References and Phase Noise</strong>&lt;br&gt;The Pound-Drever-Hall frequency control, and applications&lt;br&gt;Dr. E. Rubiola, Femto-ST</td>
</tr>
<tr>
<td></td>
<td><strong>M/NEMS, BAW, SAW, Quartz Resonators</strong>&lt;br&gt;Fundamentals, Analysis, Design and Performance of Crystal Resonators and Oscillators&lt;br&gt;Dr. John Vig, Consultant &amp; Prof. Yook-Kong Yong, Rutgers</td>
</tr>
<tr>
<td>10:30 – 10:50</td>
<td><strong>Break – 100’s Corridor</strong>&lt;br&gt;</td>
</tr>
<tr>
<td>10:50 – 12:50</td>
<td><strong>Analytical models for co-design of micromachined resonators and electronic circuits in MEMS oscillators</strong>&lt;br&gt;Prof. Ashwin Seshia, Cambridge</td>
</tr>
<tr>
<td></td>
<td><strong>We have met the enemy, and it is vibration</strong>&lt;br&gt;Dr. Mike Driscoll, Consultant</td>
</tr>
<tr>
<td></td>
<td><strong>Fundamentals of RF Acoustic Resonators and Their Characterization</strong>&lt;br&gt;Prof. Kenja Hashimoto, Chiba University</td>
</tr>
<tr>
<td>12:50 – 13:50</td>
<td><strong>Lunch – Four Seasons Ballroom 3</strong>&lt;br&gt;</td>
</tr>
<tr>
<td>13:50 – 15:50</td>
<td><strong>MEMS-Based Oscillators</strong>&lt;br&gt;Prof. Clark Nguyen, UC Berkeley</td>
</tr>
<tr>
<td></td>
<td><strong>Introduction to Atomic Clocks</strong>&lt;br&gt;Dr. Robert Lutwak, DARPA</td>
</tr>
<tr>
<td></td>
<td><strong>CMOS-MEMS Resonator Technology for Signal Processing and Sensing</strong>&lt;br&gt;Prof. S.S. Li, NTHU</td>
</tr>
<tr>
<td>15:50 – 16:10</td>
<td><strong>Break – 100’s Corridor</strong>&lt;br&gt;</td>
</tr>
<tr>
<td>16:10 – 18:10</td>
<td><strong>Piezoelectric MEMS Oscillators</strong>&lt;br&gt;Dr. Troy Olsson, DARPA</td>
</tr>
<tr>
<td></td>
<td><strong>Photonic Oscillators</strong>&lt;br&gt;Dr. Lute Maleki, OEwaves</td>
</tr>
<tr>
<td></td>
<td><strong>Piezoelectric Resonant MEMS Devices for Radio Frequency Communication and Sensing Applications</strong>&lt;br&gt;Prof. M. Rinaldi, North Eastern</td>
</tr>
</tbody>
</table>
ENTREPRENEURS FORUM
Tuesday, April 14, 10:50 - 12:30, Room 103/105

The IEEE Frequency Control Society and European Frequency Time Forum is full of entrepreneurial spirit. Many of our society members are the catalysts behind the state-of-the-art research that has led to new commercial products. Their products are making major impacts on our daily lives. This year we have the great honor of inviting four highly successful entrepreneurs to share their valuable experiences through their ventures in the frequency and timing area. This is a must-attend event for those of you interested in engaging with these leading entrepreneurs in the field.

Moderator
Wan-Thai Hsu
Micrel Inc.

Lute Maleki
Founder, President and CEO of OEwaves

Aaron Partridge
Co-founder, Chief Scientist of SiTime

Ulrich Rohde
Chairman of Synergy Microwave Corp.

Richard Ruby
Pioneer in FBAR Technologies

IEEE WOMEN IN ENGINEERING
Tuesday, April 14, 12:15 – 13:15, Room 201

Lunch Speaker: Patricia Rankin, Associate Vice Chancellor for Research, University of Colorado Boulder

Women active in the technical areas of the IEEE IFCS/EFTF conference are invited to attend a complimentary lunch and networking event organized by the women of the UFFC Society on Tuesday, April 14 from 12:15 to 13:15. Patricia Rankin, Associate Vice Chancellor for Research at the University of Colorado, Boulder is the featured speaker. Prof. Rankin is a recognized author and leader in the area of faculty development and the further advancement of women in science and engineering.
TIME & FREQUENCY AND FUNDAMENTAL PHYSICS

Tuesday April 14, 10:50–12:30, Room 102/104/106

Chair: Gaetano Mileti, LTF, University of Neuchâtel, Switzerland

Andrei Derevianko
University of Nevada, USA
Dark Matter Search with Atomic Clocks and GPS

William Shillue
National Radio Astronomy Observatory, Charlottesville, USA
Stable Time and Frequency Transfer in the Atacama Large Millimeter Array

Ernst-Maria Rasel
Leibniz Universität Hannover, Germany
Quantum Tests of the Einstein Equivalence Principle on Ground and in Space
I will discuss methods that are used to distribute time and frequency information from the perspective of the end-user. I will compare the accuracy that can be realized by means of several different methods including one-way, common-view, and two-way. I will demonstrate the capability of each of these methods with specific examples, and I will discuss how the time and frequency data that they produce can be used to control and synchronize a local clock or frequency standard. I will briefly discuss the legal and technical requirements that are important for commercial and financial institutions or, for any situation where certifying and validating the performance of a time system is an important consideration.

Judah is a physicist in the Time and Frequency Division of the National Institute of Standards and Technology (NIST) in Boulder, Colorado, and is a Fellow of JILA, the world-class joint research institute between NIST and the University of Colorado. Judah is also a NIST Fellow, a distinction reserved for the top 1% of NIST scientists. Judah received the IEEE IFCS I. I. Rabi award in 2013.

Judah’s research and metrology leadership spans many areas over a scientific career of more than 40 years, including spectroscopy and geodesy. Judah is probably best known for his scientific leadership in three areas:
- The development, operation, statistics and steering of complex time scales comprising ensembles of atomic clocks for the best possible realization of precision time of day.
- Pioneering and improving precision time and frequency transfer through telecommunications and GNSS satellites, including two-way satellite time and frequency transfer, GNSS common-view transfer, and GNSS carrier-phase transfer.
- Pioneering new methods of network time transfer, including the NIST Internet Time Service, automatically synchronizing clocks in computers and networked devices.

Judah’s work has direct impacts on the United States technology infrastructure and its economy. Judah’s Internet Time Service (ITS) now provides more than 10 billion automated synchronizations per day, and is built into commercial operating systems such as Windows, Mac, and commercial Linux. In the United States, ITS is the primary source for legally required time stamping of electronic financial transactions, including hundreds of billions of dollars of trade each day on the New York Stock Exchange, NASDAQ, and American Stock Exchange.
Parametric Excitation in Geometrically Optimized AlN Contour Mode Resonators

Ruochen Lu, Anming Gao, and Songbin Gong
Department of Electrical and Computer Engineering
University of Illinois at Urbana–Champaign
Urbana, IL, USA 61801
Email: {rlu10, agao4, songbin}@illinois.edu

Abstract—This work reports the first observation of parametric excitation in geometrically optimized Aluminum Nitride (AlN) contour mode resonators (CMRs). The concept of parametric excited AlN CMRs harnesses the fact that the resonant frequencies of extensional mode vibrations along transverse and longitudinal directions can both be determined by resonator dimensions. Therefore, by geometrically optimizing lateral dimensions, dual resonances can be engineered at $f_0$ and $2f_0$ respectively for inputting parametric excitation and outputting fundamental oscillations. In operation, the parametric excitation amplifies an orthogonal oscillation at $f_0$ by periodically modulating the stiffness constants of AlN piezoelectric thin film via straining the structure. The experimental results have shown quality factor ($Q$) enhancement from 50 to 2708 for a parametrically excited resonance. Upon further scaling and optimizations, it is anticipated that this type of devices will lead to the development of GHz low noise frequency sources and nano-electro-mechanical logic.

Keywords—parametric excitation; aluminum nitride; contour mode resonators; quality factor enhancement

I. INTRODUCTION

Recently, a wide spectrum of applications, including computing [1], sensing [2], and noise squeezing in frequency sources [3], has been proposed leveraging parametrically excited resonances in nano and micro scale devices. These demonstrations, implemented with either electrostatic or piezoelectric actuations [4], all fundamentally exploit the phenomenon of parametric amplification [5], namely the excitation of a parametric resonance with a pump signal that serves to time-varyingly modulate a system parameter, often stiffness constants of the structure, at twice the parametric resonance. Despite the exciting potential these devices have demonstrated, they mostly employ flexural acoustic modes that require vacuum ambience for achieving high $Q$ and lack the scalability to higher frequencies. Alternative resonator technologies, such as aluminum nitride (AlN) contour mode resonators (CMRs), have not been explored for parametric amplification-based applications.

Piezoelectric CMRs operate in laterally modes of expansion and contraction, of which the resonances are determined by the lateral dimension. Their displacement at resonance is predominantly in-plane and thus less susceptible to air-damping. Due to their great frequency salability (up to 10 GHz) and high $Q$ in dry air ambiance (a few thousands), piezoelectric CMRs have been vastly researched in the past decade in linear regime for applications including timing [6] and radio frequency (RF) front-end filtering [7]. Studies on their nonlinear behaviors just start to emerge [8][9], but primarily focus on the adverse effects of non-linearity (e.g. limited power handling and intermodulation) other than the exploitation of the nonlinear process in parametric amplification.

Inspired by the Melde’s experiment, which was first used in 1859 to observe the parametric amplification in mechanical systems, we present in this work geometrically optimized AlN contour mode resonators that have enabled the first observation of parametric excitation in this class of devices. Similarly to the Melde’s experiment (Fig. 1(a)), in which a string periodically vibrates at its fundamental frequency $f_0$ owing to an orthogonal excitation applied via a tuning fork vibrating at $2f_0$, an AlN CMR (Fig. 1(b)) is designed to support orthogonal vibrations with resonances at $f_0$ and $2f_0$. Measurements of the fabricated devices have shown the parametric amplification phenomenon via exhibiting high quality factor ($Q$) enhancement at $f_0$. Upon further scaling and optimizations, we anticipate that this type of devices will lead to the development of GHz low noise frequency sources and nano-electro-mechanical logic.

Fig. 1 (a) Typical example of parametric excitation-Melde’s experiment: the resonance of the string is modulated by the vibration of a tuning fork. (b) Mocked up view of the prototype parametric resonator. The dimensions of the prototypes are listed in Table 1.
Fig. 2 (a) Simulated admittance response of the prototype resonator with multiple resonances and corresponding mode shapes: (b) Longitudinal-direction second order extensional mode with a resonance at 36.2 MHz. (c) Transverse-direction fundamental extensional mode with a resonance at 72.4 MHz.

II. DEVICE DESIGN

As seen in Fig. 1 (b), the design of the parametrically excited resonator consists of a suspended piezoelectric thin film (AlN) with two top electrodes (Al) and one bottom electrode (Pt). Combined with the grounded bottom electrode, the two top electrodes form two ports of the designed resonator for inputting the parametric excitation ($2f_0$) and outputting the fundamental oscillations ($f_0$). Owing to the fact the resonances of CMRs can be defined by dimensions in design and subsequently by lithography in fabrication, the geometry of the designed resonator is first set to be 200 µm by 50 µm. The geometry leads to a fundamental resonance for the transverse direction extensional mode that is twice of the second order resonance for the longitudinal direction extensional mode [10].

The estimation of the resonances is formulated by [11]:

$$f_n = \frac{n}{2W_p} \sqrt{\frac{E_{eq}}{\rho_{eq}}}$$  \hspace{1cm} (1)

where $f_n$ is resonant frequency for the $n^{th}$ order overtone in a non-dispersive structure, $W_p$ is the pitch width, and $E_{eq}$ and $\rho_{eq}$ are the equivalent Young’s modulus and density of the comprising film stack.

However, the exact resonant frequencies of longitudinal and transverse contour modes also have weak dependence on the device aspect ratio. To ensure that the designed resonators could feature dual resonances at $f_0$ and $2f_0$, 3D finite-element analysis (FEA) were conducted using the fastPZE module of CoventorWare. The width (staring from 50 µm) and anchor length were varied in FEA and three designs were selected based on if dual resonances at $f_0$ and $2f_0$ can be attained. The key parameters of these three designs are compared in Table I. As an example, the simulated admittance response for device 1 is plotted in Fig. 2 with the identified resonances and mode shapes. As expected, two resonances designed for coupling the parametric pump and input signals are recognized. The simulated mode shapes have confirmed that the device could be excited into orthogonal extensional mode vibrations at 36.2 MHz and 72.4 MHz. The concurrent implementation of both resonances is critical for efficiently coupling the parametric excitation and amplifying the pre-existing oscillation. Anticipating variations in the fabrication process, successfully achieving dual-resonances at $f_0$ and $2f_0$ solely based on simulation results can be challenging. Therefore, an empirical approach had to be adopted. Multiple devices with incremental variations in width and anchor size from the designs in Table I were also included to ensure the demonstration of resonators with a longitudinal extensional mode resonance that is twice the resonant frequency of the transverse one.

The concept of prototype device can be interpreted as a reproduction of Melde’s experiment on chip-level and micro scale. In operation, the resonator is first excited into transverse extensional mode vibration by a single tone signal at $2f_0$ and

<table>
<thead>
<tr>
<th>Design Parameter</th>
<th>Symbol</th>
<th>Device 1</th>
<th>Device 2</th>
<th>Device 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resonator Width (µm)</td>
<td>$W_R$</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Resonator Length (µm)</td>
<td>$L_R$</td>
<td>55</td>
<td>56.5</td>
<td>54.5</td>
</tr>
<tr>
<td>Electrode Length (µm)</td>
<td>$L_E$</td>
<td>45</td>
<td>46.5</td>
<td>44.5</td>
</tr>
<tr>
<td>Electrode Width (µm)</td>
<td>$W_E$</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Anchor Length (µm)</td>
<td>$L_A$</td>
<td>20</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Anchor Width (µm)</td>
<td>$W_A$</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Top Electrode Thickness (nm)</td>
<td>$T_{Top}$</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Bottom Electrode Thickness (nm)</td>
<td>$T_{Bot}$</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>AlN Thin Film Thickness (nm)</td>
<td>$T_{AlN}$</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>

Fig. 3 (a) Fabrication process for the contour mode AlN parametric resonators, and SEM images of the fabricated resonator: (b) perspective view and (c) top view.
subsequently experiences parametrically amplified resonant vibration at $f_0$ in the longitudinal direction. The pump signal at $2f_0$ amplifies the fundamental oscillation by periodically modulating the stiffness constants of piezoelectric thin film. The amplification results in a larger admittance near the parametrically excited resonance and give rise to the enhancement in the measured $Q$.

III. EXPERIMENTAL RESULTS

In order to investigate the parametric excitation in the designed CMRs and understand underlying physics, devices specified in Table I were fabricated on a 4-inch high resistivity Si wafer with a process similar to the one described in the previous publication [11]. The fabrication flow chart and SEM images of the fabricated device are shown in Fig. 3. Due to the residue stress in AlN thin film, the released structure exhibits upward warping.

The admittance measurements of devices were first done in linear regime to identify the dual resonances and thus determine the frequency of the parametric pump. As seen in Fig. 5, the measured admittance response displays resonances at 37.35 MHz and 74.75 MHz for orthogonal direction extensional modes, with slight deviation from the simulation results. This is probably caused by the residue stress and film warping in released AlN.

The parametric amplification process was then studied by observing the quality factor enhancement of the resonance at $f_0$ through the measured admittance response. The experimental setup is shown in Fig. 4. An excitation signal around 74.75 MHz was applied to the parametric excitation port at different power levels (9–17 dBm) using an Agilent N5181A RF analog signal generator as the admittance response was being measured at the input port with an Agilent 5230A PNA-L network analyzer. 0V DC bias was employed in the measurement. The parametric pump signal was expected to amplify the vibration initialized by the VNA around 37.35 MHz during the admittance measurement.

The $Q$ enhancement induced by parametric excitation was investigated concerning the excitation power and pump frequency. As seen in Fig. 6, when a parametric pump (74 MHz) of slightly less than twice the longitudinal resonance (37.35 MHz) is applied and a power threshold of 14 dBm is reached, the parametric excitation phenomenon is observed. Further increasing the pump power causes more pronounced $Q$ enhancement, which maximizes at 2708 with a pump power level of 17dBm. No higher power was applied in the measurement due to the power limit of the RF signal source.

At the power level of 17dBm, varying the pump frequency near twice the longitudinal direction resonance can also have an impact on the parametric excitation. As seen in Fig. 7(a), if the pump signal deviates from the optimal frequency at which parametric excitation is most pronounced, the $Q$ enhancement of the parametric resonance in Device 2 quickly diminishes. Similar trend is observed for Device 3, which features a slightly larger but also limited pump frequency range for supporting parametric excitation. More interestingly, as the pump signal increases, the parametrically enhanced resonance shifts upward until the pump frequency falls out of the aforementioned range. As the pump frequency approaches boundaries of this limited frequency range, the parametric resonance also lessens gradually. The observed dynamics is expected for the phenomenon of parametric excitation, although a quantitative model is required in the future to fully explain the results. Analytical models for accurately describing the nonlinear behaviors of the parametrically excited resonators are currently under development.
IV. CONCLUSION

In this work, the first observation of parametric excitation in a geometrically optimized AlN contour mode resonator is reported. The dynamics of parametric amplification was experimentally studied and Q factor enhancement as large as x50 is reported. This class of parametric resonators has shown potentials in enabling the development of GHz low noise frequency sources and nano-electro-mechanical logic.

ACKNOWLEDGMENT

The authors would like to thank DARPA for funding support under the Young Faculty Award (YFA) and the staff in Micro and Nanotechnology Laboratory cleanroom for help with the fabrication process. The authors would also like to thank Cristian Cassella and Prof. Gianluca Piazza at Carnegie Mellon University for the AlN thin film deposition.

REFERENCES

Abstract—The third-order intermodulation distortion in properly terminated high-order bridged clamped-clamped beam (CC-beam) channel-select micromechanical filters has been measured for the first time and found to be appreciably higher than seen on unterminated stand-alone CC-beams. In particular, a three-resonator bridged 8-MHz filter with 140nm gaps posts a measured $IIP_3$ of 25dBm and 36dBm for two-tone offsets of 200kHz and 400kHz, respectively, which are much larger than the -11.7dBm for a stand-alone 9.2MHz CC-beam with 200kHz two-tone offset. For the case where the required minimum signal-to-noise ratio is 10dB, the $IIP_3$ of 36dBm translates to an impressive spurious-free dynamic range of 102.3dB. The result matches well the prediction of a new model for non-linearity that incorporates not only parallel-plate capacitor nonlinearity, but also the influence of the filter structure, where $Q$-loading by the termination reduces the degree of out-of-channel tone suppression, but also reduces the amplitude of resonator motion, which then improves the $IIP_3$. The full model incorporates these phenomena, plus dependencies on dc-bias, gap spacing, and electrode area.

Keywords—Third-order intermodulation, micromechanical resonators, micromechanical filters, channel selection, capacitive-gap transducer.

I. INTRODUCTION

Channel-select filters like those of this work, capable of rejecting all interferer signals relax the dynamic range requirements on subsequent stages, e.g., the LNA and the mixer, thereby greatly reduce power consumption. However, the degree of interferer suppression depends strongly on the linearity of the filter, which if not sufficiently linear, can also generate intermodulation spurs even after rejecting interferers.

A recent 223-MHz capacitive-gap transduced disk micromechanical filter [1] proved that with electrode-to-resonator gaps on the order of 39nm, its transduction was sufficient to achieve a channel-select response with a small percent bandwidth of 0.09% and in-band insertion loss of 2.7dB, as well as a 50dB of out-of-channel rejection, all while maintaining a termination resistance around 500Ω which is easily $L$-network-matched to conventional antennas. However, use of such devices in RF communication circuits has so far been deferred due to skepticism on the linearity of these filters, which must satisfy the very strict specifications of present-day wireless standards. For example, the European GSM standard for mobile communications [2] requires a minimum total $IIP_3$ of -18dBm in the receiver path to insure adequate suppression of interferer signals from adjacent channels.

Previous works [3, 4, 5] on this issue were limited to two-tone measurements on single resonators, as opposed to multiple resonator filters. Pursuant to determining the linearity of high-order capacitive-gap micromechanical filters, this paper represents a complete analytical formulation for the $IIP_3$ of such devices and then verifies the formulation via experimental measurement on 3rd- and 4th-order 8-MHz clamped-clamped beam filters, shown in Fig. 1. With $IIP_3$ on the order of 36dBm for 400-
components arise, inserting (2) into (1), harmonics as well as intermodulation signals at frequencies not present in the original input. Specifically, not only a scaled version of the input, but also spurious a nonlinear transfer function of the form in (1), the output in-

desired signal at frequency

\[ \xi(\omega_0 t) + \xi(\omega_1 t) + \xi(\omega_2 t) \]

where \( x_1 \) and \( x_2 \) add to the desired signal at frequency \( \omega_0 \). When this signal passes through a nonlinear transfer function of the form in (1), the output includes not only a scaled version of the input, but also spurious signals at frequencies not present in the original input. Specifically, inserting (2) into (1), harmonics as well as intermodulation components arise

\[
y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \tag{1}
\]

where \( x \) and \( y \) are input and output and \( a_s \) describe system behavior. For the case of the third order intermodulation distortion, the input of consequence takes the form

\[
x = S_0 \cos(\omega_0 t) + S_1 \cos(\omega_1 t) + S_2 \cos(\omega_2 t) \tag{2}
\]

where two interferer signals at frequencies \( \omega_1 \) and \( \omega_2 \) add to the desired signal at frequency \( \omega_0 \). When this signal passes through a nonlinear transfer function of the form in (1), the output includes not only a scaled version of the input, but also spurious signals at frequencies not present in the original input. Specifically, inserting (2) into (1), harmonics as well as intermodulation components arise

\[
y = \cdots + a_1 S_0 \cos(\omega_0 t) + \frac{1}{2} a_2 S_0^2 \cos(2\omega_0 t) + \frac{3}{4} a_3 S_0^3 \cos(2\omega_1 - \omega_2) t \cdots \tag{3}
\]

The harmonic terms occupy much higher frequencies, so can be filtered out easily. However, intermodulation between two equispaced input tones can produce output terms at frequencies near the desired signal. In particular, third-order nonlinearity can generate an output component exactly at the desired frequency if one tone is twice as far from \( \omega_0 \) as the other one, as shown in Fig. 2 and Eq. (3). This third-order intermodulation component, \( IM_3 \), can directly impact the signal-to-noise ratio of the desired channel and eventually mask the channel completely, if the system nonlinearity is too large. Hence, IM distortion must be con-

strained below a minimum acceptable value.

The third-order input intercept point (IIP3) is defined as the input power level at which the extrapolated intermodulation component has the same power as the fundamental output. In general, a larger IIP3 indicates smaller nonlinearity (or better linearity) in a given system and hence smaller intermodulation component generation, which is a design goal of communication systems.

As shown in Fig. 2, while a micromechanical channel-select filter attenuates out-of-channel interferer signals, its nonlinearity can still result in intermodulation components at its output that corrupt the desired signal, which motivates the necessity to de-

sign for maximum filter IIP3. Note that interferer signals for a channel-select filter are outside the filter passband, unlike the case for a band-select filter, for which in-band interferers must be considered.

B. Capacitive-Gap Transducer Nonlinearity

Nonlinearity in either resonator stiffness [7] or capacitive-
gap transduction [4] is often the most important contributors to filter nonlinearity. The former becomes significant when large displacement induces non-negligible internal strain in the reso-

nator, which manifests as a stiffness nonlinearity. Since inter-

ferer signals are out-of-channel for a channel-select filter, the induced displacement is generally very small, so the stiffness nonlinearity is negligible. Therefore, transducer nonlinearity which translates to input force nonlinearity generates the inter-

modulation term of a channel-select filter.

As shown in Fig. 3, the transducers in the micromechanical filters of Fig. 1 are simplified to a parallel-plate capacitor com-

posed of a gap spacing of \( d_0 \) separating two polysilicon electrodes: one fixed, the other suspended by an effective stiffness \( k_r \). The transducer takes as input a dc voltage \( V_r \) applied to one electrode and an ac excitation voltage \( v_i \) applied to the other. Since no dc current flows through the capacitor, there is no dc power consumption. The free electrode moves in response to force \( F_{tot} \) generated by the input voltage combination following a biquad frequency response of Fig. 3, where the electrode ef-

effective stiffness \( k_r \) and quality factor \( Q \) determine the maximum displacement according to

\[
\frac{X}{F}(\omega) = \frac{1}{k_r} \cdot \theta(\omega) \tag{4}
\]
where \( \omega_0 \) is the resonance frequency from the ratio of \( k \) and effective mass \( m_e \).

The actuation force \( F_{\text{tot}} \) that drives the resonator to move a displacement \( x \) derives from the co-energy in the capacitor [4]:

\[
F_{\text{tot}} = \frac{1}{2} (V_p - V_i)^2 \frac{\partial C}{\partial x} \left( 1 - \frac{x}{d_0} \right)^{11} \frac{x}{d_0} \left( 1 + \frac{4x^2}{d_0^2} + \frac{4x^3}{d_0^3} + \ldots \right)
\]

where \( C_0 \) is the parallel-plate capacitance at rest, and \( d_0 \) is the electrode-to-resonator gap spacing, neglecting beam bending under applied dc-bias voltage \( V_P \). For purposes of IIP3 determination, the input voltage comprises the sum of two out-of-band signals:

\[
v_i = V_1 \cos(\omega_1 t) + V_2 \cos(\omega_2 t)
\]

The \( F_{\text{tot}} \) expression (6) has a second-order dependence on input voltage \( v_i \), which combined with the nonlinear nature of the change in the capacitance per unit displacement \( \partial C/\partial x \), results in higher order nonlinearities. In particular, out-of-band interferer signals at \( \omega_1 \) and \( \omega_2 \), equispaced from \( \omega_0 \) and from each other (\( \omega_1 = \omega_0 - \omega_2 \)), induce small displacements at \( \omega_1 \) and \( \omega_2 \), respectively. Nonlinear interaction between these displacements and with the input voltages produces a displacement spur at \( \omega_0 \) via third-order intermodulation terms, as shown in Fig. 4.

Here, the total resulting resonator displacement takes the form:

\[
x = X_0 \cos(\omega_0 t + \phi_0) + X_1 \cos(\omega_1 t + \phi_1) + X_2 \cos(\omega_2 t + \phi_2)
\]

where the \( \phi_0 \) and \( \phi_1 \) are the phases between the input and output voltage \( v_i \), and the \( X_0 \) is the displacement amplitudes. The equation of motion governs the relation between \( F_{\text{tot}} \) and \( x \):

\[
m_r \ddot{x} + b_r \dot{x} + k_r x = F_{\text{tot}}
\]

Neglecting nonlinear force terms, approximate expressions for the out-of-band \( X_1 \) and \( X_2 \) magnitudes take the form

\[
X_1 = \frac{V_p V_i C_0}{k_r} \Theta(\omega_1)
\]

\[
X_2 = \frac{V_p V_i C_0}{k_r} \Theta(\omega_2)
\]

If \( \omega_1 \) and \( \omega_2 \) are sufficiently distant from the center frequency \( \omega_0 \), the approximate phases of the displacement components become

\[
\phi_0 = 90^\circ, \phi_1 = \phi_2 = 0 \text{ or } 180^\circ
\]

Substituting (10)-(12) into (8) and using this \( x \) in (6) yields the input force intermodulation term (13) shown at the bottom of the next page, where \( \varepsilon_0 \) is the permittivity in vacuum, \( d_0 = W_r W_e \) is the electrode-to-resonator overlap area, \( \Theta(\omega_1) \) and \( \Theta(\omega_2) \) are [\( \Theta(\omega) \)]

The third-order input intercept point (IIP3) is defined as the input power for which the third-order modulation IM3 term is equal to the fundamental term. Assuming the nonlinearity associated with current generation at the output transducer is negligible compared to those associated with force generation at the input transducer, equating \( F_{\text{fund}} \) (14) to the fundamental force \( F_{\text{fund}} \) (14) when exciting the resonator at \( \omega_0 \), and then solving for \( V_i \), yields the expression for input voltage \( V_{\text{IIP3}} \), as given in (15).

\[V_{\text{IIP3}} = \frac{V_i}{d_0^2}
\]

At the corresponding input power, \( I_{\text{IIP3}} \), then takes the form

\[I_{\text{IIP3}} = \frac{V_{\text{IIP3}}}{2R_T}
\]

where \( R_T \) is the total resistance in the system, seen at the input, dominated by the resonator motional resistance.

A careful examination of equation (15) shows that \( V_{\text{IIP3}} \) is dependent on not only material properties, but also resonator and electrode geometry. In particular, \( V_{\text{IIP3}} \) increases with decreases in \( V_P \) and \( A_0 \) and with increases in \( d_0 \) and \( k_r \). Since these parameters also affect the motional resistance \( R_e \) of micromechanical structure, there is a trade-off between the \( V_{\text{IIP3}} \) and the motional resistance of a capacitive-gap transducer.

C. Filter Considerations in IIP3 Calculation

Although the parallel-plate approximation is still valid for a capacitive-gap micromechanical filter, the \( I_{\text{IIP3}} \) expression requires two major modifications to account for the following:

1) Total Resistance:

Proper termination is essential to minimize in-channel ripple and attain a flat passband. Before termination, the resonator \( Q \)'s are too large and the filter passband consists of distinct peaks, as
shown in Fig. 5(a). The termination resistance needed to flatten the passband of a filter with center frequency \( f_0 \), bandwidth \( B \), and small insertion loss is [8]:

\[
R_T = \left( \frac{Q_{\text{fltr}}}{Q_{\text{fltr}}} - 1 \right) R_x \equiv \frac{Q}{Q_{\text{fltr}}} R_x
\]

(17)

where \( Q_{\text{fltr}} = f_0 / B \). In contrast to the single resonator case, these termination resistors \( R_q \) dominate the resistance in the system, so \( R_T \sim 2R_q \) in equation (16) when calculating the \( I_{\text{IP3}} \) of filters.

2) High-Order Mechanical System:

The equation of motion (9) describes only single resonators, whereas the micromechanical filters of Fig. 1 comprise several resonators linked by coupling beams. Solving the complete mechanical system yields transfer functions listed in filter cook books, such as [8]. For example, the transfer function of a second-order filter coupled by quarter-wavelength coupling beams takes the form

\[
\Theta(\omega) = \frac{X}{F} (\omega) = \frac{1}{k_r} \frac{\Theta(\omega)}{1 + \frac{P_{\text{bw}}}{Q_{\text{fltr}}}} (18)
\]

(19)

where \( P_{\text{bw}} \) is the filter percent bandwidth. Note that terminating the filter with \( R_q \) loads the quality factors of its end resonators, which is why \( Q_{\text{fltr}} \) appears in (19). The loading effect reduces resonator displacement at resonance \( (X_i \propto Q/k_r) \) and reduces out-of-band rejection leading to larger out-of-band displacements \( X_1 \) and \( X_2 \). However, compared with a similarly terminated stand-alone resonator, the high-order transfer function of a multi-resonator filter shown in Fig. 5 provides a larger out-of-band attenuation that reduces out-of-band displacement, decreasing \( I_{\text{IP3}} \), hence increasing \( I_{\text{IP3}} \).

III. COMPLETE FORMULATION FOR \( I_{\text{IP3}} \)

While the parallel-plate capacitor approximation provides an analytical solution and design insight for the effect of third-order nonlinearities, it neglects phenomena such as beam bending due to dc-bias voltage and location-dependent effective stiffness. This can introduce errors in the \( I_{\text{IP3}} \) calculation.

\( V_Y \)-induced beam bending results in \( C_0 \) and \( d_0 \) of (10) and (11) that are not constant, but rather functions of location on the \( y \)-axis given by \( d(y) \) in [5], neglecting variation along the \( z \)-axis. On the other hand, the effective resonator stiffness is also location dependent and changes according to the mode shape at the point of interest. For similar reasons, \( X_1 \) and \( X_2 \) vary along the beam length (the \( y \)-axis in Fig. 1(b)), approaching zero near the anchors. In general, with knowledge of the peak displacement (at the beam midpoint) governed by the resonator lumped model, displacements at other beam locations follow from the resonator mode shape. To correctly determine the total actuation force \( F_{\text{act}} \) intermodulation force components \( dF_{\text{IM3}} \) in infinitesimal regions \( dy \) at locations \( y \) should be integrated over the entire beam length.

Since effective stiffness increases dramatically moving away from the beam center, displacement is a strong function of location. This phenomenon can result in a \( V_{\text{IP3}} \) value twice as large as that derived using a simple parallel-plate approximation. Including these modifications is essential to better explain the experimental results.

IV. EXPERIMENTAL RESULTS

Third- and fourth-order clamped-clamped beam bridged micromechanical filters [9] like those of Fig. 1 were designed and fabricated to test the efficacy of the filter \( I_{\text{IP3}} \) formulations. Besides the quarter-wavelength coupling beams connecting adjacent resonators to form the basic filter transfer functions, quarter- and third-quarter-wavelength bridging beams also couple the first and last non-adjacent resonators [9] to provide parallel paths for mechanical signals from input to output. With proper

\[
F_{\text{IM3}} = V_i^2 \left[ \frac{V_P^3 (\epsilon \rho A_0)^3}{k_r^2 d_0} \left( 2|\theta_1|^2 |\theta_2| + \frac{9}{4} |\theta_1||\theta_2|^2 + |\theta_1|^3 + \frac{3}{8} |\theta_1|^2 |\theta_2|^2 \right) \right]
\]

+ \( V_i^3 \left[ \frac{1}{4 k_r} \left( 2|\theta_1| + |\theta_2| \right) + \frac{3 V_P^2 (\epsilon \rho A_0)^3}{2 k_r^2 d_0} \left( |\theta_1| + 2|\theta_2| \right) |\theta_1|^2 |\theta_2| + \frac{3 V_P^2 (\epsilon \rho A_0)^3}{2 k_r^2 d_0} |\theta_1|^3 + \frac{3}{8} |\theta_1|^2 |\theta_2|^2 \right) \]

(13)

\[
V_{\text{IP3}} = V_i^2 \left[ \frac{V_P^3 (\epsilon \rho A_0)^3}{k_r^2 d_0} \left( 2|\theta_1|^2 |\theta_2| + \frac{9}{4} |\theta_1||\theta_2|^2 + |\theta_1|^3 + \frac{3}{8} |\theta_1|^2 |\theta_2|^2 \right) \right]
\]

+ \( V_i^3 \left[ \frac{1}{4 k_r} \left( 2|\theta_1| + |\theta_2| \right) + \frac{3 V_P^2 (\epsilon \rho A_0)^3}{2 k_r^2 d_0} \left( |\theta_1| + 2|\theta_2| \right) |\theta_1|^2 |\theta_2| + \frac{3}{8} |\theta_1|^3 + \frac{3}{8} |\theta_1|^2 |\theta_2|^2 \right) \]

(15)
design, these parallel paths are out of phase with the primary paths so insert a zero in the filter transfer function, effectively adding a notch in the filter frequency response and sharpening the passband-to-stopband rolloff. Each resonator in the filters of Fig. 1 has a tuning electrode to compensate any deviation in the resonator center frequency due to fabrication tolerances, which then enables near perfect (flat) passbands.

A. Fabrication

Filters were fabricated using a previously described vertical gap surface-micromachining process [5], summarized by the process cross-sections in Fig. 6, with some modifications to incorporate a damascene process to enable a thick, low resistance interconnect layer. Fabrication starts with deposition of 2µm-thick silicon dioxide and 400nm-thick silicon nitride on the silicon substrate to electrically isolate different interconnects. A wet dip in hydrofluoric acid then releases devices, leaving free standing structures, such as shown in Fig. 6(d). A wet dip in hydrofluoric acid then releases devices, leaving free standing structures, such as shown in Fig. 6(d). Fig. 6(e) present the SEM of a released third-order 3CC-λ/4-bridged filter. Here, a spectrum analyzer measured the output power response to two-tone inputs with frequency spacings like those pictured in Fig. 2. These filters show an IIP3 of 17dBm for a tone separation of 125kHz; and in (b), IIP3’s of 22.7dBm and 27dBm for tone separations of 80 and 160kHz, respectively.

B. Measurement Results

Fig. 7 presents the frequency response spectra of the terminated micromechanical filters like those of Fig. 1. The filters are biased at 21V, 20V and 22.5V and terminated by 12kΩ, 10kΩ and 20kΩ resistors, respectively. Tuning voltages were adjusted to yield flat passbands. Table I summarizes the filter characteristics.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Designed/Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3CC-λ/4</td>
</tr>
<tr>
<td>µRes. Beam Length, L&lt;sub&gt;r&lt;/sub&gt;</td>
<td>40.8µm</td>
</tr>
<tr>
<td>µRes. Beam Width, W&lt;sub&gt;W&lt;/sub&gt;</td>
<td>8.0µm</td>
</tr>
<tr>
<td>µRes. Beam Thickness, t&lt;sub&gt;r&lt;/sub&gt;</td>
<td>2.0µm</td>
</tr>
<tr>
<td>Electrode Width, W&lt;sub&gt;e&lt;/sub&gt;</td>
<td>20.0µm</td>
</tr>
<tr>
<td>Gap Spacing, d&lt;sub&gt;o&lt;/sub&gt;</td>
<td>137nm</td>
</tr>
<tr>
<td>Coupling Beam Length, L&lt;sub&gt;12&lt;/sub&gt;</td>
<td>22.3µm</td>
</tr>
<tr>
<td>Coupling Beam Length, L&lt;sub&gt;12-4&lt;/sub&gt;</td>
<td>22.3µm</td>
</tr>
<tr>
<td>Coupling Beam Width, W&lt;sub&gt;W&lt;/sub&gt;</td>
<td>750nm</td>
</tr>
<tr>
<td>Coupling Location, l&lt;sub&gt;12&lt;/sub&gt;</td>
<td>4.7µm</td>
</tr>
<tr>
<td>Coupling Location, l&lt;sub&gt;12-4&lt;/sub&gt;</td>
<td>2.8µm</td>
</tr>
<tr>
<td>Filter Biasing Voltage, V&lt;sub&gt;f&lt;/sub&gt;</td>
<td>21V</td>
</tr>
<tr>
<td>Center Frequency, f&lt;sub&gt;c&lt;/sub&gt;</td>
<td>8.08MHz</td>
</tr>
<tr>
<td>Electromech. Coupling, C&lt;sub&gt;H&lt;/sub&gt;/C&lt;sub&gt;0&lt;/sub&gt;</td>
<td>14.77%</td>
</tr>
<tr>
<td>Resonator Quality Factor, Q</td>
<td>12,000</td>
</tr>
<tr>
<td>Bandwidth, B</td>
<td>10.7kHz</td>
</tr>
<tr>
<td>Percent Bandwidth, B/f&lt;sub&gt;c&lt;/sub&gt;</td>
<td>0.13%</td>
</tr>
<tr>
<td>Passband Ripple, PR</td>
<td>&lt;0.2dB</td>
</tr>
<tr>
<td>Insertion Loss, IL</td>
<td>1.2dB</td>
</tr>
<tr>
<td>20dB Shape Factor</td>
<td>2.08</td>
</tr>
<tr>
<td>Stopband Rejection, SR</td>
<td>38dB</td>
</tr>
<tr>
<td>Predicted third-order input intercept point, I&lt;sub&gt;P&lt;/sub&gt;</td>
<td>(f&lt;sub&gt;R&lt;/sub&gt; = 40kHz)</td>
</tr>
<tr>
<td>Measured third-order input intercept point, I&lt;sub&gt;P&lt;/sub&gt;</td>
<td>36.0dBm</td>
</tr>
<tr>
<td>Sp. Free Dyn. Range, SFDR</td>
<td>(SNR&lt;sub&gt;BW&lt;/sub&gt; = 10dB)</td>
</tr>
<tr>
<td></td>
<td>102.3dB</td>
</tr>
</tbody>
</table>

Fig. 8 presents IIP3 measurements on (a) a 3CC-3λ/4 bridged filter and (b) a 4CC-3λ/4 bridged filter. Here, a spectrum analyzer measured the output power response to two-tone inputs with frequency spacings like those pictured in Fig. 2. These filters show (a) an IIP3 of 17dBm for a tone separation of 125kHz; and in (b), IIP3’s of 22.7dBm and 27dBm for tone separations of 80 and 160kHz, respectively.
The expression for out-of-band spurious-free dynamic range takes the form [10]:

$$\text{SFDR} = \frac{2}{3}(IIP_3 + 174\text{dBm} - 10 \log(B)) - \text{SNR}_{\text{min}}$$  \( (20) \)

where $\text{SNR}_{\text{min}}$ is the required minimum signal-to-noise ratio. Assuming $\text{SNR}_{\text{min}}=10\text{dB}$, these $IIP_3$’s correspond to $\text{SFDR}$’s of 89.94dB for the 3CC-3λ/4 bridged filter with 125kHz tone separation; and 91dB and 93.9dB for the 4CC-3λ/4 bridged filter with 80 and 160kHz tone separations, respectively.

Fig. 9 presents $IIP_3$ measurements on a 3CC-λ/4 bridged micromechanical filter. Specifically, (a) shows two-tone measurements for different tone spacings, while (b) plots measured $IIP_3$ as a function of percent tone spacing ($\Delta\omega/\omega_0$). As expected, as out-of-band interferers move further away from the center frequency, induced displacements $X_1$ and $X_2$ decrease and $IIP_3$ increases. This filter achieves $IIP_3$ of 11dBm, 22dBm and 36dBm for tone separations of 80, 160 and 400kHz, respectively. These $IIP_3$’s are more than 30dBm higher than previous marks for single resonators [4, 5].

V. CONCLUSION

The analytical expression for multi-resonator capacitive-gap transduced channel-select CC-beam bridged micromechanical filters not only matches measured $IIP_3$’s as high as 36dBm for a tone separation of 400kHz, but also provides insight into just how good the linearity of such filters can be. In particular, the formulation shows that capacitive-gap transducer nonlinearity depends on not only material properties, but also on structure and electrode geometry, all of which serve as knobs through which one might design for a specific linearity requirement. Channel-select filters like those of this work, capable of rejecting all interferer signals and passing only the desired signal, and doing so with the high $IIP_3$’s demonstrated, stand to greatly reduce dynamic range requirements on subsequent stages, in turn enabling much longer battery lifetimes for ultra-low power wireless front-ends.

ACKNOWLEDGMENT

The work was supported by DARPA.

REFERENCES

Nonlinear Acceleration Sensitivity of Quartz Resonators

Jianfeng Chen and Yook-Kong Yong
Department of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
Piscataway, NJ, USA

Randall Kubena, Deborah Kirby and David Chang
HRL Laboratories, LLC
Malibu, CA, USA

Abstract—The nonlinear effects of initial stress/strain of the quartz plate resonator on its acceleration sensitivity was studied. Finite element models were developed using a theory of small deformations superposed on finite initial deformations in a Lagrangian formulation. AT- and SC-cut quartz circular plate resonators were studied. The plates were respectively subjected to diametrical compression force and bending force. The initial strains due to the application of diametrical force represented initial strains due to in-plane acceleration, while the initial strains due to bending force represented initial strains due to out-of-plane acceleration. Our model results using nonlinear initial strains showed good agreement with measured data by Ballato, Mingins, and Fletcher and Douglas. The model results using linear initial strains compared well only with the measured data for plates subjected to diametrical force but not for plates subjected to bending forces. Hence our model results showed that for accurate prediction of out-of-plane acceleration sensitivity the nonlinear initial strains must be used. The linear initial stress/strain cannot fully capture rotation and bending effects. The acceleration sensitivity model using linear initial strains could only be employed for in-plane acceleration, or for very low g out-of-plane acceleration. The SC-cut crystals showed better linearity of frequency change with respect to applied bending forces than the AT-cut crystals. The principle of superposition for out-of-plane acceleration sensitivity in AT-cut crystals is in general not valid, especially in cases of high g accelerations.

Keywords—Geometric nonlinearity, linear initial strains, nonlinear initial strains, In-plane acceleration sensitivity; Out-of-plane acceleration sensitivity, AT- and SC-cut crystals.

I. INTRODUCTION

When a piezoelectric quartz crystal is subjected to external force or vibration, the resulting nonlinear quasi-static stresses and strains on the crystal cause the natural resonant frequency to shift [1]. The main factor that causes frequency shift is deformation of the resonator. This frequency shift gives rise to unwanted sidebands in the frequency spectrums and in the degradation of output signals leading to reduced stability and accuracy of the system performance under dynamic environments.

Frequency shifts in quartz crystal resonators due to applied forces or acceleration have been studied extensively since the early 1970’s. The acceleration sensitivity of quartz resonators was thought to be well understood. More recently the acceleration sensitivity problems were solved using three-dimensional finite element analysis because of the easy availability of computing resources and the complex geometries of resonators. The initial strains in these acceleration sensitivity problems were computed using linear three-dimensional equations of elasticity.

In the present paper, we have found that the initial strains computed using linear equations of elasticity are accurate only for acceleration sensitivity in the case of in-plane forces or vibrations. For acceleration sensitivity in the case of out-of-plane forces or vibrations the linear equations of elasticity are not accurate. We have found that for accurate acceleration sensitivity predictions, the nonlinear terms in the initial stress/strain must be retained in order to fully capture the out-of-plane deformations in resonators. For cases involving out-of-plane accelerations, the linear equations of elasticity were accurate only for low accelerations of less than 10g.

II. GEOMETRIC NONLINEARITY AND EQUATIONS OF INCREMENTAL MOTION SUPERPOSED ON FINITE DEFORMATIONS

A. Geometric nonlinearity

Nonlinearity occurs in many practical applications of engineering. The underlying principle of nonlinear behavior is that cause and effect relationships are not proportional unlike the linear system. Nonlinear behavior can be grouped into two general behaviors: material nonlinearity and geometric nonlinearity. Material nonlinearity arises when the material exhibits nonlinear stress-strain relationship. Geometric nonlinearity arises when a system undergoes large deformation in which there are finite changes in the geometry of deforming body.

Normally it is small strains which define whether a problem is geometrically linear. However, there are geometric nonlinear problems that are small strains but with finite rotations and bending deflections. Geometric nonlinearity should be used whenever there is finite bending, where the linear strain tensor is replaced by the Green-Lagrange strain tensor, and the Cauchy stress tensor is replaced by the second Piola-Kirchhoff stress tensor. The strain-displacement equations for the linear strain tensor and Green-Lagrange strain tensor are:

\[
e_{ij} = \frac{1}{2}(U_{ij} + \mathbf{U}_{ij})
\]  

\[
E_{ij} = \frac{1}{2}(U_{ij} + \mathbf{U}_{ij} + \mathbf{U}_{k,j}U_{k,i})
\]

Linear strain tensor:

Green-Lagrange strain tensor:
The main difference between the two strains is the product term $U_{k,l}U_{k,j}$ which is responsible for finite bending deflections and rotations. We show in the present paper that for the 10 MHz, AT-cut circular (12 mm diameter) plate resonators the nonlinear initial strains are needed for accurate prediction of out-of-plane acceleration sensitivity. When the plate was cantilever mounted and subjected to 2 grams-force in the $X-Y$ plane, the dominant displacement gradients were $U_{1,2}$ and $U_{2,1}$, which had orders of magnitude $10^{-3}$ at the center of the disk, while other components of displacement gradient had orders of magnitude ranging from $10^{-6}$ to $10^{-8}$. The out-plane acceleration sensitivity was sensitive to these other components of displacement gradient and the product term $U_{k,l}U_{k,j}$ must be retained in eqn.(2).

In the case of the plate subjected to 2 grams force in-plane compression, the dominant displacement gradient was $U_{1,1}$ which had orders of magnitude $10^{-7}$ at the center of disk, while other components of displacement gradient had orders of magnitude ranging from $10^{-7}$ to $10^{-10}$. The product term $U_{k,l}U_{k,j}$ with orders of magnitude $10^{-14}$ can be neglected in eqn.(2), and linear strain tensor of eqn.(1) can be used in the case of in-plane acceleration sensitivity.

$\text{B. Equations of incremental motion superposed on finite deformations}$

For our study of acceleration sensitivity of quartz resonators we employ the Lagrangian three-dimensional equations of motion for incremental vibrations superposed on finite deformations[2]. The process of superposing small vibrations on initial deformations due to applied external forces or body forces in a crystal can be described by three consecutive states shown in Fig. 1. For simplicity here we write only the mechanical equations of incremental motion superposed on finite deformations since the electromechanical coupling of quartz is small. Our COMSOL models and results have included the piezoelectric effects.

\[ s_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i} + U_{k,l}u_{k,j} + U_{k,j}u_{k,l}) \]  

\[ (t_{ij} + t_{j,k}U_{i,k})_j + \rho b_i = \rho \ddot{u}_i \text{ in } V \]  

The terms $U_{k,l}$ and $S_{mn}$ are the initial deformation gradient and strain respectively.

$\text{III. COMPARISON OF MEASURED DATA IN CIRCULAR CRYSTAL DISK SUBJECTED TO DIAMETRICAL FORCES}$

We compared the experimental results by Mingins[3] and Ballato[4] of acceleration sensitivity due to in-plane diametrical forces with our COMSOL model results. An AT cut circular disk of diameter $d$ and thickness $2b$, was subjected to a pair of diametrical forces $F$ at an angle $\psi$ (azimuth angle) with the $X_1$-axis as shown in Fig. 2a. The frequency change is a measure of the acceleration sensitivity of the quartz plate for the given $\psi$ angle.

\[ X_1 \quad X_2 \quad X_3 \]

\[ \psi \]

\[ F \]

\[ \text{Fig. 2a. Experimental setup of circular disk subjected to diametrical forces.} \]

We have developed a COMSOL finite element model of eqns. (3)-(5) using the initial deformation gradient and strain from either eqn.(1) or eqn.(2). The Fletcher and Douglas[5] crystal plate of Fig.2b was modeled. The resonator plate dimensions were diameter $d=12$ mm, gold electrodes diameter 4 mm at the center and azimuth angle $\psi$ varying from $0^\circ$ to $90^\circ$. The resonant mode used was the fundamental thickness shear mode. The plate was fixed at one edge, compressive force $F$ applied at the diametrically opposite edge and prescribed displacements $u_y=u_z=0$ as shown in Fig. 2b. These boundary conditions yielded the same state of stress/strain as the pair of diametrical compressive forces of Fig. 2a.

\[ t_{ij} = (C_{ijkl} + C_{ijklmn}S_{mn})s_{kl} \]  

\[ \text{Incremental stress equations of motion:} \]

\[ (t_{ij} + t_{j,k}U_{i,k})_j + \rho b_i = \rho \ddot{u}_i \text{ in } V \]  

The terms $U_{k,l}$ and $S_{mn}$ are the initial deformation gradient and strain respectively.
The results of frequency change which represented the accelerations sensitivity were used to calculate a force sensitivity coefficient, \( K_f \), as defined by Ratajksi’s [6]:

\[
K_f = \left( \frac{\Delta f}{f_0} \times \frac{1}{f} \right) \frac{d}{f_0}
\]  

(6)

This force sensitivity coefficient \( K_f \) as a function of the azimuth angle \( \psi \) for our COMSOL model results and measured results are shown in Fig. 2c for AT-cut crystals. The solid green and solid red curves are our FEA model results using, respectively, the linear initial strains of eqn.(1) and nonlinear initial strains of eqn.(2). Note that only the solid red curve could be seen in the graph because it overlapped the solid green curve. We observed that the force sensitivity coefficient \( K_f \) using either the linear initial strains, or the nonlinear initial strains compared equally well with the measured results. Therefore, the use of linear initial strains was accurate for the prediction of in-plane acceleration sensitivity in this case of the plate resonator under in-plane stresses and deformations.

IV. COMPARISON OF MEASURED DATA IN CIRCULAR CRYSTAL DISK SUBJECTED TO CANTILEVER BENDING AND SYMMETRICAL BENDING

We now investigate the case of out-of-plane acceleration sensitivity in AT- and SC-cut crystals. The effects of bending moments on the frequency of the 10 MHz AT- and SC-cut circular plate resonators were studied. The crystal plate had the same dimensions as the plate in Fig. 2b. The bending was applied on either the cantilever mounted plate or the symmetrically mounted plate. The experimental setups of two mounting configurations are shown in Figs. 3. Fig. 3a shows the apparatus for applying the bending forces, Fig. 3b shows the cantilever bending and Fig. 3c shows the symmetrical bending.

1) AT cut crystal

In order to study the effects of bending force on the AT-cut plate mounted in the cantilever configuration, the crystal was clamp at one end and 2 grams force was applied at the diametrically opposite end to create an upward bending of the crystal shown in Fig. 3b. Fig. 4a shows our model of the cantilever plate of Fig. 3b. The angle \( \psi \) was the angle between the X-axis of the clamped edge to the crystal X-axis. The resulting frequency change (\( \Delta f \)) was a measured of the out-of-plane acceleration sensitivity of the quartz plate.
The model results of frequency change due initial strains from either the linear strain eqn.(1) or the nonlinear strain eqn.(2) were compared to Fletcher and Douglas[5] experimental results in Fig. 4b. The solid green curve is the frequency change due to linear initial strains while the solid red curve is the frequency change due to nonlinear initial strains. We observed that the solid red curve compared well with the experimental data while solid green curve did not compare well. Therefore for out-of-plane accelerations sensitivity the nonlinear initial strains must be used for accurate modeling. The frequency change due to linear initial strains was accurate only at two angles $\psi = 90^\circ$ or $270^\circ$.

The linearity of frequency change with the applied cantilever force is of fundamental interest because it determines whether the principle of superposition is applicable for acceleration sensitivity of quartz plates. Fig.4c shows the frequency change as a function of the applied cantilever force from 0 to 10 grams-force. Measured results from Fletcher and Douglas[5] showed that the frequency change with applied cantilever force was linear when the azimuth angle $\psi$ was $270^\circ$ and nonlinear when $\psi$ was $220^\circ$. Our COMSOL model results using nonlinear initial strains matched well the experimental results for the azimuth angle $\psi = 270^\circ$. For the azimuth angle $\psi = 220^\circ$ the model results was good only for applied cantilever force less than 2 grams-force. The principle of superposition of accelerations sensitivity is applicable for $\psi = 270^\circ$ but not for $\psi = 220^\circ$.

**Linear initial strains versus nonlinear initial strains:**

The theory of elasticity states that when strains are sufficiently small the nonlinear strains of eqn.(2) will converge towards the linear strains of eqn.(1). This was demonstrated in Fig.4c where the applied cantilever force was reduced to 0.1 gram-force. The frequency change using linear initial strains (green curve) converged towards the frequency change using nonlinear initial strains (red curve). Therefore, the frequency change using linear initial strains in these plate resonators were sufficiently accurate only for small bending force of 0.1 gram-force.

2) *SC cut crystal*

The 10 MHz *SC* cut plate was also studied in cantilever configuration, and our model results were compared with the measured results. The results of both the linear and the nonlinear initial strains models were compared with Fletcher and Douglas[5] measured data shown in Fig. 5a. We have found that our nonlinear model results (solid red curve) matched well with the measured data. We had to shift our model’s azimuth angle $\psi$ by $+90^\circ$ to compare with the measured data. Since the plate was circular, Fletcher and Douglas[5] might have erred in starting their measured data at $\psi = 90^\circ$ instead of $0^\circ$. The frequency change using linear initial strains (solid green curve) was accurate only at $\psi$ angles $0^\circ$ and $180^\circ$, respectively.
The linearity of frequency change with applied cantilever force was studied for the SC-cut plate resonators. The results shown in Fig.5b for azimuth angle $\psi = 40^\circ$ are linear and the principle of superposition for acceleration sensitivity is therefore applicable. Our model results (solid red line, using nonlinear initial strains) were also linear although the slope of the line was less than the measured results of [5].

![Fig.5a. Frequency change as a function of azimuth angle $\psi$ for cantilever bending of SC-cut crystal.](image1)

![Fig.5b. Frequency change as a function of applied cantilever force for the SC-cut circular plate resonator for azimuth angle $\psi = 40^\circ$.](image2)

**B. Symmetrical bending of circular plate resonators.**

Fletcher and Douglas[5] also measured the effects of bending force in a symmetrical configuration. Fig.3c shows the symmetric bending configuration: The crystal was clamped at one end, a knife edge placed over the top at the midway point to the opposite edge, and a line load of 5 grams-force was applied at the opposite edge. Fig.6a shows our model of the symmetric bending configuration. Fig.6b shows our model results using nonlinear initial strains in comparison with the measured data by Fletcher and Douglas[5]. Our model results (red curve) showed a trend similar to the measured data (blue curve) although the magnitude of frequency changes were off at azimuth angles $\psi$ such as 105°, 150°, and 225°.

![Fig.6a. Symmetrical bending of circular plate.](image3)

![Fig.6b. Frequency change as a function of azimuth angle $\psi$ for symmetrical bending of SC-cut circular plate resonator.](image4)

![Fig.6c. Frequency change as a function of applied force for symmetrical bending of SC-cut circular plate resonator for azimuth angle $\psi = 90^\circ$.](image5)

The linearity of frequency change with applied bending force for symmetrical bending at $\psi = 90^\circ$ was also studied. Fig.6c shows the measured data in comparison with our model results using nonlinear initial strains. Both measured data and model results have similar trends with slightly nonlinear curves. The slope of our model results (red line) was less than that of the measured data.

**V. RECTIFIED ACCELERATION SENSITIVITY IN QUARTZ RESONATORS**

Although the phenomenon of rectified acceleration sensitivity had been observed in experiments by Mingins et al. [7], it had not received much attention because it was not well understood. We define rectified acceleration sensitivity[8] as one in which the sign of frequency change $\Delta f$ is independent of the sign of applied force or acceleration. For our study of the cantilever plate bending in
**AT-** and **SC-cut** circular plate resonators, the bending force was applied in both positive and negative Y-axis direction, and the results were shown in Figs. 7a and 7b respectively. We found that for the **AT** cut crystal, the frequency change was rectified at $\psi = 0^\circ$ and $180^\circ$, while for the **SC** cut crystal, the acceleration sensitivity was rectified at $\psi \sim 140^\circ$, and $\sim 320^\circ$. (Please note that although the red curve (-2 grams force) and blue curve (+2 grams force) appear to intersect at frequency change $\Delta f = 0$, they in fact intersect at small values of $\Delta f \neq 0$.)

**SUMMARY**

Incremental equations for small vibrations superposed on initial deformations were implemented in COMSOL finite element models to study the acceleration sensitivity of quartz circular plate resonators. The plates were respectively subjected to diametrical compression force and bending force. The model results using **nonlinear** initial strains showed good agreement with measured data by Ballato[4], Mingins[3, 7], and Fletcher[5]. The model results using **linear** initial strains compared well only with the measured data for plates subjected to diametrical force but not for plates subjected to bending forces.

The initial strains due to the application of diametrical force represented initial strains due to in-plane acceleration, while the initial strains due to bending force represented initial strains due to out-of-plane acceleration. Hence our model results showed that for accurate predictions of out-of-plane acceleration sensitivity the nonlinear initial strains must be used.

A summary of our findings:

1) In practical applications, the quartz resonator is subjected to both in-plane and out-of-plane accelerations and/or vibrations. Since the model using linear initial strains is accurate only for in-plane acceleration sensitivity, it would be prudent to employ models using nonlinear initial strains in practical applications.

2) The acceleration sensitivity model using linear initial strains could only be employed for in-plane acceleration, or for very low $g$ out-of-plane acceleration.

3) The nonlinear initial stress/strain must be used in out-of-plane acceleration sensitivity models in order to fully capture the bending and rotation effects.

4) The **SC**-cut crystals showed better linearity of frequency change with respect to applied bending forces than the **AT**-cut crystals. The principle of superposition of out-of-plane acceleration sensitivity in **AT**-cut crystals is in general not valid, especially in cases of high $g$ accelerations.

5) Rectified acceleration sensitivity for the **AT**-cut crystal was found at $\psi = 0^\circ$ and $180^\circ$; while for the **SC** cut crystal, the rectified acceleration sensitivity was found at $\psi \sim 140^\circ$, and $\sim 320^\circ$.

**ACKNOWLEDGMENT**

Sponsored by DARPA Microsystems Technology Office (MTO), Program: Dynamics-Enabled Frequency Sources (DEFYS). Issued by DARPA/CMO under Contract No. HR0011-10-C-0109. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Advanced Research Projects Agency.

**REFERENCES**


Multiple SAW Resonance Sensing Through One Communication Channel with Multiple Phase Detectors

TAKIZAWA, Yoshinori†*; SHIBATA, Takayuki**; KASHIWADA, Shinji**; YAMAMOTO, Yasuo**; ESASHI, Masayoshi*; TANAKA, Shuji†

* WPI-AIMR, Tohoku University, Sendai City, Miyagi Prefecture, Japan
** Research Laboratories, DENSO CORPORATION, Nisshin City, Aichi Prefecture, Japan
† Graduate School of Engineering, Tohoku University, Sendai City, Miyagi Prefecture, Japan
E-mail: takizaway@mems.mech.tohoku.ac.jp

Abstract—SAW device is widely used in the various passive sensing application such as temperature, mechanical force and other various measurements. The SAW devices have relatively high the temperature dependency, thus in some applications, the multiple SAW sensors are used for the temperature compensation. In such a case, the communication between the sensor and controller is highly important. This paper describes the one channel communication with multiple SAW resonators for the measurement applications. The final target of this research is to investigate the near magnetic field wireless passive sensing capability using the continuous wave for higher response speed.

Keywords—SAW, Resonator, Phase Detection, Passive Measurement, TCF

I. INTRODUCTION

With using the SAW devices, many measurement applications especially for the temperature and mechanical force are proposed for the passive remote sensing.[1] The two types of SAW device are mainly used in such cases, “Resonator” and “Delay Line”. In signal processing wise, there also two main methods are commonly utilized, time division (TD) and continuous wave (CW) systems.

The most popular combination of SAW device and signal processing is “SAW delay line and TD”.[3-7] A transmitter sends the radio wave in a specific period and waits for the returned signal. After receiving the returned signal, the signal is analyzed and the next radio wave will be sent. The very precise measurement is feasible but the processing time might be slow.

The second popular one is “SAW Resonator and CW”. A transmitter always sends the signal and the returning signal is immediately processed. The higher speed is the most advantage of this type.[8,9] The other combinations “Resonator and TD” and “Delay Line and CW” [2] are rare cases.

One difficulty of SAW device is the relatively high Temperature Coefficient of Frequency (TCF). Therefore, if the measurement environments are in the environment of the temperature change, and the multiple SAW device systems are commonly proposed.[10]

In this paper, the feasibility of multiple SAW resonators along with the PLL frequency tracking is studied aiming at the higher speed at the order of $10^4$ Hz through one communication line. The experiment with near magnetic field sensing is challenged and mentioned.

II. BASIC THEOREM

The blockdiagram of the proposing system is shown in Fig. 1. The same numbers of PLL circuits as resonators are provided. The VCOs signals are summed and fed into the parallel connected multiple SAW resonator circuit through one communication line. Each VCO oscillating frequency is corresponding to each resonator frequency. The selective phase detection can be proved in the time-frequency domain analysis.[11].

Assuming the multiple sinusoidal signals are the linear summation of the different amplitude $f_n(t) = \sum_{n=1}^{N} a_n e^{-j \omega_n t + \theta_n}$, and the summed signal applies to a SAW circuit $h(t)$. The output signal $g(t)$ is expressed as

![Figure 1 Blockdiagram for the multiple SAW resonators measurement system with multiple PLLs](image-url)
\[ g(t) = h(t) \ast \sum_{n=1}^{N} f_n(t) = \sum_{n=1}^{N} h(t) \ast f_n(t) \]

\[ = \sum_{n=1}^{N} h(t) \ast a_n e^{-j\omega_n t + \theta_n}, \quad \cdots \cdots (1) \]

where * denotes convolution. Taking Fourier Transform of eq (1)

\[ \mathcal{F}(g(t)) = G(\omega) = 2\pi \mathcal{F}(h) \sum_{n=1}^{N} \mathcal{F}(f_n)e^{j\omega n} \]

\[ = 2\pi H(\omega) \sum_{n=1}^{N} a_n \delta(\omega - \omega_n)e^{j\theta_n} \quad \cdots \cdots (2) \]

The phase detection can be expressed as the multiplication of eq (1) and one of the original signals of \( k \)th.

\[ g_k(t) = a_k e^{j\omega_k t + \theta_k} \sum_{n=1}^{N} h(t) \ast a_n e^{j\omega_n t + \theta_n}, \quad \cdots \cdots (3) \]

Taking Fourier Transform of eq (3), the \( k \)th generator’s phase detection can be expressed as eq (3) in time domain and eq (4) in the frequency domain.

\[ \mathcal{F}(g_k(t)) = G_k(\omega) \]

\[ = a_k \delta(\omega_k)e^{j\theta_k} + 2\pi \mathcal{F}(H) \sum_{n=1}^{N} a_n \delta(\omega - \omega_n)e^{j\theta_n} \]

\[ = 2\pi a_k \sum_{n \neq k} \delta(\omega_k) + a_n H(\omega) \delta(\omega - \omega_n)e^{j\theta_n} \]

\[ = \pi a_k e^{j\theta_k} H(\omega) \sum_{n=1}^{N} a_n e^{j\theta_n} (\delta(\omega - (\omega_n - \omega_k)) + \delta(\omega - (\omega_n + \omega_k))) \quad \cdots \cdots (4) \]

This shows that the \( k \)th phase detector output is the series of the delta function with the magnitude of the SAW circuit frequency response \( H(\omega) \).

If the \( k \)th Loop Filter has cut off frequency of \( \omega_{zk} \), only the signal lower than \( \omega_{zk} \) can be passed through the \( k \)th loop filter, i.e.

- if \( |\omega_n - \omega_k| < \omega_{zk} \), \( \delta(\omega - (\omega_n - \omega_k)) \) is passed, and
- if \( |\omega_n + \omega_k| > \omega_{zk} \), \( \delta(\omega - (\omega_n + \omega_k)) \) is filtered out.

All the other \( \delta(\omega) \) terms of \( n \neq k \) are filtered out under the condition of \( (\omega_n - \omega_k) \ll \omega_{zk} \) for any \( n \). Eventually only the following term remain which indicate the single PLL.

\[ G_k(\omega) = \pi a_k^2 e^{j2\theta_k} H(\omega) \delta(\omega) \quad \cdots \cdots (5) \]

Therefore, when the VCO gain is appropriately designed to meet with eq (5), Fig.1 circuit behaves as a single phase lock loop against the \( k \)th resonator.

Fig. 2 shows the result of the two SAW resonators separation by the phase detection and loop filter at 433.42MHz and 433.92MHz. Fig. 2 (a) shows SAW resonators output or the input to the phase detectors. Figs (b) and (c) shows the phase detector outputs for 433.42MHz and 433.92MHz. The loop filter’s cut of frequency is 20.7 kHz of the 3rd order R, C filter. The two SAW resonator frequency difference is 0.5MHz and the phase detector and loop filter gain is approximately -60dB in voltage. Hence no interferences between two phase detector outputs were not observed in Fig. 2. This results were measured with Zurich Instruments Lock-In Amplifier, UHFLI.

III. EXPERIMENT WITH FOUR RESONATORS

The theory was confirmed with the parallel connected four different frequency SAW resonators shown in Table I. These were the on-the-shelf TCF stabilized SAW resonators to confirm the proposed theory.

Fig. 3 shows (a) the four parallel connected SAW resonator impedance response, and (b) the phase detector followed by the loop filter outputs. Two lines in (b) are the response of one of the phase detectors to the frequency sweep input from 400 to 440MHz at the temperature 20 and 200 °C. This result predicts the
assumption in this paper is supposed to be correct shown in eq (5).

Fig. 4 shows the frequency response of the four resonators. Four sinusoidal signals corresponding to the four resonance frequencies were added and applied to the SAW resonator circuit as shown in Fig. 1. Changing the temperature of SAW resonators from 20 to 200 °C, and the resonator behaviors were tested.

In order to investigate the PLL tracing and TCF, the four resonators were mounted on a PCB and the PCB was placed on the hot plate control the temperature. The resonant frequency change was measured as shown in Fig. 4. TCF was as good as approximately 5.2 ppm/K. The measurement point was every 20 °C and all the frequency differences (Δ Frequency Dependence) plots were very well fitted with the 2nd order linear approximation.

Even though the frequency dependencies on temperatures are estimated to the very good fitting curves, the response for the 403.55MHz resonator showed the convex response from 20 to 100°C. This means that the output voltage of phase detector followed by the loop filter cannot show the two possibility of the SAW resonant frequency. The PLL had been traced the frequency change properly, so that the SAW resonant frequency could be known by monitoring the VCO frequency. However, the frequency counters are required in this case and the system will be costly.

Therefore the linear frequency response is easier to measure the physical stress of temperature and mechanical force. Reasonably wide TCF cannot be denied for the measurement purpose and the good response fitting is extremely important.

IV. TWO RESONATORS AND WIRELESS

As already discussed in the previous chapter, while Fig. 4 shows the excellent TCF performance but this may not be always good especially in the case that the SAW device is used for the measurement purpose. Therefore, new linear LiNbO₃ SAW resonators with the resonance frequencies of 203MHz and 204MHz were developed.

Using the two SAW resonators, the temperature dependency was measured and the result is shown in Fig.6. The figure apparently shows the excellent TCF of 8.5 ppm/K. In many cases of mechanical stress measurement such as pressure, torque, tensile, and so on, the resonant frequency change of the temperature is much smaller in the order of 100° or less. Therefore, the TCF linearity is important, and the devices shown Fig. 6 will be the excellent candidate of the mechanical force measurement with the temperature compensation.

One of the most important application of the SAW device is the passive wireless sensing. The sensor might be used in the harsh environment and active components cannot be installed in the sensor unit. In such a case, the passive wireless communication is useful and the one channel communication is suitable for the purpose. A couple of 40nH coils, 10mm diameter 2 turn coil, were fabricated and test the possibility of the communication.

Fig. 7 shows one of the wireless application with the near magnetic field. The coil communication is fairly easy to fabricated and inexpensive solution for the near distance. In this type of analysis, S₁₁ parameter is sometimes is helpful. However, the SAW resonator’s IDT is driven by the electric field. Therefore the voltage analysis may be more appropriate analytical way.
Fig. 7 (b) shows the results by changing the coil coupling factor \( k \) which is relating to the distance of the coils. \( k=0.5 \) is very close each other. At \( k=0.3 \), there is no response from resonator2. This happens the resonator1’s anti-resonance and resonator2’s resonance are very close position and the results of both are interferences. When \( k \) goes smaller to 0.1, there was no more response from the resonators were readout. The load was 50 \( \Omega \) resister.

Generally to obtain a good communication efficiency, the resonance of the coil and SAW resonator’s resonance frequency should be matched. However, as the experiments in the paper already showed, the SAW resonance frequencies have relatively large range of frequency changes.

The Q factor of the couple of coil is sharp. Furthermore, if the \( k \) of the coil is changed, then the coil resonance frequency is also changed. Here is the difficult of the SAW resonator and the magnetic coil communication difficulty. This may have the chance to improve by optimizing the coil resonance frequency and the load against the SAW resonant frequencies, which has to be more studied. A phase lock loop was designed and fabricated with Analog Devices AD8302 phase detector, Maxim MAX2607 VCO with the gain of 11.9MHz/V, and 3rd order 37 kHz loop filter, more than 50 kHz convergence speed was confirmed.

V. CONCLUSION

The SAW resonators and PLL circuits combination will be the good solution for the high speed SAW passive sensing. One channel communication to read out the frequency change of multiple SAW resonators were proposed and confirmed its frequency tracking with PLLs. This method has the capability of wireless communication for passive sensing system was explained. As for the wireless sensing, the communication efficiency needs to be improved. Especially if the near magnetic field coil is used for the single line communication, the wide band stability is the issue to be improved.

ACKNOWLEDGMENT

This research was the joint work of DENSO Corporation, and Advanced Institute for Materials Research, Tohoku University supported by World Premier International Research Center Initiative(WPI), MEXT, Japan. The authors would like to express the appreciation for giving the opportunity of this research. The authors also would like to appreciate Mr. Hiroyuki Wado of DENSO who arranged this joint project, and Mr. Harutsugu Fukumoto of DENSO for his discussion and suggestions.

REFERENCES

Abstract—We use a Ramsey-type interaction scheme to measure spatially-resolved images of the static magnetic field (C-field) amplitude $B_{dc}$ applied across the Rb cell in the physics package of a high-performance vapor-cell atomic clock. Low field variations of $\leq 0.5\%$ are found across the recorded images, and Fourier analysis of the data indicates low variations of $B_{dc}$ also along the direction of laser propagation. Images of the $T_2$ relaxation time are obtained in a similar way, and show a distribution that correlates with the $B_{dc}$ distribution. This indicates inhomogeneous dephasing due to C-field gradients, which also results in spatial variation of the $T_2$ time for the clock transition.

Keywords—Atomic clocks, Magnetic field measurement, Microwave resonators, Microwave spectroscopy, Optical pumping.

I. INTRODUCTION

Precise knowledge of the microwave and static magnetic field (“C-field”) distribution across an alkali vapor cell is crucial for the development and optimization of high-performance vapor-cell atomic clocks, both based on continuous-wave (cw) [1] or pulsed optically-pumped (POP) [2] interaction schemes. We are working towards the realization of a POP clock based on a highly compact magnetron-type microwave cavity [3] that has the potential to reduce significantly the size and mass of the clock Physics Package (PP) [4]. In order to increase the understanding of the microwave field geometry across the Rb vapor cell used inside this clock PP, we recently reported on experimental imaging of the microwave magnetic field distribution in this magnetron-type microwave resonator cavity [5]. This microwave field imaging technique is based on time-domain pulsed laser and microwave interaction in the Rabi scheme, originally demonstrated for studying microwave structures in cold atom chips [6], and later applied to a thin, sealed Rb vapor cells inside a microwave resonator [7].

In this present communication we report on an extension of this imaging technique to also obtain images of the static magnetic field (“C-field”) applied to the Rb vapor cell in our clock PP, with high spatial resolution. Such spatially-resolved experimental measurements of the C-field distribution in the cell are generally difficult to achieve in a fully assembled clock PP, due to the presence of the vapor cell and other optical components that hinder a field sensor in the region of interest inside the cavity. In our approach, we employ the Ramsey interaction scheme for sensing the Rb atoms inside the vapor cell that act as magnetic field sensor here. Like in the microwave field imaging study of [5] a relatively thick cell is studied here (cell length $L=25\text{mm}$), which is in contrast to the studies reported in [6, 7] where thin ($L \leq 2\text{ mm}$) cells were studied. In the following we present this Ramsey-type imaging of the C-field distribution and the experimental results obtained on our high-performance clock PP using a thick Rb vapor cell. Fourier analysis of the time-domain signals is proposed to extract additional information on the field distribution along the laser propagation direction. Application possibilities of the method are discussed.

II. EXPERIMENTAL SETUP AND METHOD

The experimental setup (see Fig.1) uses a laser and microwave setup similar to the ones described in [7], but is applied to the high-performance clock PP described in [1, 4]. This clock PP is based on a Rb vapor cell of $25 \text{ mm}$ diameter and length, held in a compact magnetron-type microwave cavity of $\leq 45 \text{ cm}^3$ overall volume [3]. The cell contains enriched $^{87}\text{Rb}$ and a buffer-gas mixture, and is held at $55^\circ\text{C}$. A solenoid is used to apply a constant and uniform magnetic field oriented along the $z$-direction (laser propagation direction), and two layers of magnetic shields suppress external field fluctuations. For the pulsed Rabi and Ramsey schemes, the laser light is switched on and off using an acousto-optical modulator (AOM), and the microwave radiation is switched by microwave switches triggered by a synchronization control.

Spatially-resolved detection is achieved by imaging the laser light level transmitted through the cell onto a CCD camera, whose integration time is also triggered by the synchronization control. After data treatment, the lateral resolution of the CCD images (in the x and y directions transverse to the laser propagation direction) corresponds to

---

This work was supported by the Swiss National Science Foundation (SNSF grants 149901, 140712 and 140681) and the European Metrology Research Programme (EMRP project IND55-Mclocks). The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.
\[ \Delta OD = A + B \exp\left(-\frac{\delta t_R}{T_1}\right) + C \exp\left(-\frac{\delta t_R}{T_2}\right) \sin(\delta dt_R + \phi) \]
$B_{dc}$ are $<0.13 \ \mu T$ (0.4%) across the entire measured image. Images obtained on the other Zeeman transitions show the same shape of the distribution and absolute $B_{dc}$ values. For a more detailed analysis of the field variation, $B_{dc}$ is plotted against the distance $r$ from the central pixel of symmetry ($x=0\text{mm}, y=-0.72\text{mm}$), see Fig. 4a. The data is well-fitted by $B_{dc} = B_0 + B_2 r^2 + B_4 r^4$, with a small correction only due to the coefficient $B_0$. We attribute these field variations to residual inhomogeneities originating from the C-field coil geometry (edge effects for a solenoid of finite length) and the impact of the magnetic shields.

B. Relaxation time imaging

From the single pixel time series $\Delta OD(dt_k)$ (see Fig. 2) also the time constant $T_2$ of the Ramsey oscillation decay can be extracted, i.e. the $T_2$ relaxation time of the sampled Rb ground-state microwave transition. The images obtained closely resemble those for $B_{dc}$, with $T_2$ having a maximum value of 400 $\mu$s at the central pixel of symmetry ($x=0\text{mm}, y=-0.72\text{mm}$), and decreasing roughly linearly to $\approx 250 \ \mu$s at $r \approx 4\text{mm}$, see Fig. 4b. This decrease of $T_2$ with $r$ hints on inhomogeneous dephasing due to the Rb atoms travelling by diffusion between regions of slightly different $B_{dc}$ values. This inhomogeneous dephasing is bigger in regions of more pronounced field gradient $dB_{dc}/dr$, so the general trend $B_{dc} \sim r^2$ results in $T_2 \sim dB/dr \sim r$. Note that furthermore $T_2$ values are slightly above (below) the general linear trend in regions where the slope of $B_{dc}$ is slightly lower (steeper) than that of the average fitted polynomial, indicated by vertical solid (dashed) arrows, respectively, in Fig. 4. The same imaging and data analysis applied to the $m_i = 0 \leftrightarrow m_z = 0$ clock transition reveals that also here $T_2$ decreases with increasing $r$, see Fig. 4c, but with a slope significantly smaller than measured on the $m_i = 1 \leftrightarrow m_z = 1$ transition. Note that these spatial variations of $T_2$ do not correlate with the distributions of the microwave magnetic field in the same clock PP reported in [5]. We thus conclude that in our clock PP the $T_2$ of the clock transitions shows a spatial variation due to the small ($<0.15 \ \mu T$) residual inhomogeneities of the applied C-field.

C. 3D information from Fourier analysis

In images like that shown in Fig. 3, the data within each pixel is integrated along the path of the light propagation ($z$ direction). In case of sufficiently strong variation of $B_{dc}$ along $z$, the pixel’s Ramsey oscillations might thus contain more than one oscillation frequency, not taken into account by Equation (1). Using Fourier transform analysis inspired from NMR techniques [8], one can access the data’s Ramsey frequency spectrum and thus obtain statistical information on the variation of $B_{dc}$ along $z$, by mapping the Fourier frequency spectrum into dc magnetic field strength. Figure 5 illustrates this approach for five selected pixels as indicated in Fig. 3: pixel $(x=0\text{mm}, y=0\text{mm})$ at the center of the image in Fig. 3, and pixels $(x=\pm 2.42\text{mm}, y=0\text{mm})$ and $(x=\pm 4.84\text{mm}, y=0\text{mm})$ towards the edges of the image. The obtained single-pixel Fourier spectra show one main peak each, while the other regions of the $B_{dc}$ spectrum generally show much smaller relative contributions of $<10\% - 20\%$ of the main peak amplitude. The bump observed for some of the pixels around 39.8 $\mu T$ might be an artefact due to residual low-frequency fluctuations in the experimental setup. All five pixels’ peaks show clear shifts of their center $B_{dc}$ field values, in agreement with the values expected from Fig. 3. The FWHM of the main peaks is $<0.2 \ \mu T$ (0.5%), originating in approximately equal parts from the limited length of the timeseries data and the

Fig. 4. Field amplitude $B_{dc}$ (a) and relaxation time $T_2$ (b) as function of distance $r$ from the center of symmetry, for the $m_i = 1 \leftrightarrow m_z = 1$ transition. Dashed and solid vertical arrows mark regions where the gradient $dB_{dc}/dr$ locally is larger and smaller than the general trend, respectively. Sub-panel (c) shows the $T_2$ variation measured on the $m_i = 0 \leftrightarrow m_z = 0$ clock transition.

Fig. 5. Distribution of static magnetic field amplitudes, obtained by Fourier analysis of the Ramsey oscillation time series as shown in Fig. 2. Data is shown for the five pixels $(x, y) = (0 \text{mm}, 0 \text{mm}), (\pm 2.42 \text{mm}, 0 \text{mm})$ and $(\pm 4.84 \text{mm}, 0 \text{mm})$ here. Note the change in $x$-axis scaling at $B_{dc}=41.5 \ \mu T$. 

23
dominating $T_2$ time constant [8]. From the absence of stronger broadening of the main peaks we thus conclude that the variation of $B_{dc}$ along the z-direction is not above this value over significant parts of the z-extension of the cell.

IV. CONCLUSIONS

We have applied a pulsed Ramsey-type interrogation scheme to experimentally measure the static magnetic field (e-field) distribution inside a fully assembled physics package of a high-performance vapor-cell atomic clock. By using a CCD camera as light detector, images of the static magnetic field variation across the vapor cell’s cross section are obtained, with a spatial resolution of $\leq 100 \, \mu m$ in the x and y directions. The static magnetic field is found to vary by $< 0.4\%$ across the cell’s cross section. Using Fourier analysis of the same data we can also obtain information on the field variation along the laser propagation (z) direction that indicates that also variations along z direction are on or below this same level. Images of the $T_2$ relaxation time distribution across the cell are also obtained with the same technique, and indicate a variation of $T_2$ due to inhomogeneous dephasing due to the C-field gradients. Spatial variations of the clock transition’s $T_2$ time due to this effect are evidenced. The presented technique is useful for assessing the evolution of static magnetic field distribution in a fully assembled vapor-cell atomic clock physics package, even after extended clock operation times when e.g. the magnetization state of the magnetic shields and thus the static magnetic field over the cell volume may have changed.

ACKNOWLEDGMENT

We thank M. Pellaton and P. Scherler (both UniNe-LTF) for support in realizing the cavity and clock physics package, and A. Skrivervik and A. Ivanov (both EPFL-LEMA) for fruitful collaboration on the microwave cavity.

REFERENCES

87Rb Isoclinic Point Thermometry

N. Wells, T. Driskell, and J. Camparo
Electronics and Photonics Laboratory
The Aerospace Corporation
El Segundo, CA 90245
Email:nathan.p.wells@aero.org

Abstract: The concept of atomic thermometry is demonstrated in an experimental system consisting of two lasers frequency stabilized using FM spectroscopy. Utilizing the temperature insensitivity of the 87Rb Di isoclinic point, a reference laser is beat against a second laser locked to an atomic transition whose frequency is strongly dependent on the vapor temperature. The data presented here indicates that thermal fluctuations down to the mK level could potentially be sensed. This concept may be useful to obtain improved long term performance in Rb vapor cell atomic clocks.

Introduction

The long-term temperature instability of vapor-phase atomic systems, at the tens of milli-Kelvin level, is known to have a significant effect on the frequency stability of Rb atomic clocks and atom-stabilized lasers. Building on previous work examining the isoclinic point in (nuclear spin) I = 3/2 alkalies [1], we demonstrate in this paper an isoclinic-point thermometer.

We have previously shown that the minimum in the Doppler broadened D1 87Rb absorption transition midway between the F_g=2 \rightarrow F_e=1 and F_g=2 \rightarrow F_e=2 lines, the “isoclinic point,” is a suitable frequency reference for laser stabilization with a near zero temperature sensitivity [1,2]. Additionally, the isoclinic point can be employed for the measurement of the energy dependence of optical transition collisional shifts [3]. Here we combine our previous work to demonstrate isoclinic point thermometry, where we sense the temperature of an atomic vapor in an all optical fashion.

In brief, two 795 nm DBR lasers are frequency stabilized to optical transitions in Doppler or pressure broadened 87Rb vapor cells. Locking a single laser to the 87Rb isoclinic point in a vacuum cell provides a temperature insensitive reference point. The second laser is then locked to a second vapor cell with a temperature sensitive 87Rb optical transition. The beat frequency between these two lasers should then provide a measure of the temperature fluctuations of the second vapor cell.

The sensitivity and accuracy of this optical thermometer are inherently linked to the lasers’ long term frequency stabilization. We will discuss the influence of systematic FM spectroscopy effects on laser stabilization [4,5], and disentangle those effects from other instabilities of a purely atomic nature.

Experimental

Figure 1 shows the schematic of our isoclinic point thermometry test bed. The error signals used to stabilize the laser frequency to the Rb absorption resonances are derived from FM spectroscopy using electro-optic modulators to impart phase modulation on the laser field. The two EO phase modulators are driven at 125 MHz and 100 MHz respectively.
The electro-optic (EO) phase modulators are bulk MgO:LiNbO with LC resonant tank circuits for operation with a low power RF drive. The laser irradiance in the Rb cell is controlled with a $\lambda/2$ plate/polarizer pair. The polarizer is aligned along the electric field of the EO modulator. As will be discussed below, careful alignment of the polarizer/EO axis is needed to minimize the impact of residual amplitude modulation (RAM) [6].

filtered to remove the 2f component, and sent to a PID controller that actuates the laser injection current adjustment for locking the laser frequency.

The two frequency stabilized lasers are split before the EO phase modulator, and the unmodulated beams are combined and detected with a beat note detector (Vescent Photonics D2-160 [250 MHz-9.3 GHz]). The beat tone is analyzed with a microwave frequency counter or a spectrum analyzer.

One laser is locked to the isoclinic point of the $^{87}$Rb spectrum in a vacuum cell. The second laser is then locked to a peak in a vacuum cell (either $F_g=2\rightarrow F_e=1$ (a) or $F_g=2\rightarrow F_e=2$ (b), see Fig. 1). This produces a beat note at half of the excited state hyperfine splitting (ESHFS), approximately 406 MHz. Alternatively, the second laser is locked to a peak in a pressure broadened cell. The beat frequency will then depend on which peak is chosen, the buffer gas species and the pressure of the cell. For a Kr cell, the pressure shift coefficient is approximately -5 MHz/Torr and the beat note will be observed at $(5 \times \text{Cell Pressure}) \pm \text{ESHFS}$. For a 25 Torr Kr cell, the beat notes should be observed at 281 and 531 MHz corresponding to transitions b and a respectively, as shown in Fig 1.

The Rb cells are temperature controlled via an oven consisting of two Al blocks that sandwich the cell, two heaters on solid state relays, and a temperature controller that holds the oven temperature to better than 0.1 K. To measure the temperature sensitivity of any Rb resonance, one cell is stabilized to a fixed temperature while the other undergoes a stepwise ramp. The beat frequency is then recorded and analyzed when the ramp temperature stabilizes.

Fig. 1: a) $^{87}$Rb D$_1$ electronic manifold showing resonances for laser frequency stabilization. b) Two similar DBR lasers are frequency stabilized to Rb absorption resonances using FM spectroscopy. The beat frequency from these stabilized lasers is used to sense the temperature of the Rb vapor. See text for discussion.

Each phase modulated laser is passed through a Rb absorption cell and detected with a high speed photodiode. The photodiode signal is then sent to the RF port of a double balanced mixer (DBM) and is homodyned with a phase shifted copy of the EO drive signal. The RF+LO port of the DBM is then low pass
Results and Discussion

Figure 2 shows typical results of frequency locking one laser to the isoclinic point in a vacuum cell and one laser to peak a (see Fig. 1) in a vacuum cell. Both cells are held at a constant temperature (40 °C). As can be readily observed, a rapid decrease of approximately 1 MHz is initially observed in the beat note frequency, followed by a short period (I) of exceptional stability, and a slightly worse long term performance (II). Around a time of 5½ hours one laser becomes unlocked. After re-locking the laser, the beat note retraces within 100 kHz.

The cause of the initial rapid drop in the beat note frequency shown in Fig. 2a is not readily known, but as we discuss below, this behavior is likely caused by changes in the RAM imparted by the EO phase modulators due to temperature fluctuations in the room. This experiment was started early in the morning prior to the heating/cooling system of the laboratory engaging. After this initial period, the laser stabilization shows a nearly 5 hour period where the beat frequency moves by only 400 kHz (1×10^{-9}). A short period of time (I) shows exceptional stability where the Allan Deviation (Fig 2b) shows stability at 5×10^{-11} at 1000 s. The longer period of stability (II) shows an Allan Deviation of 1.5×10^{-10} at 1000 s.

Impact of RAM

It is well understood that EO phase modulators are far from perfect and will impart residual amplitude modulation (RAM) on the laser field [5]. In FM spectroscopy, RAM shows up, partly, as an unwanted DC offset in the demodulated atomic absorption signal [4,7]. Additionally, the RAM is dependent on the relative alignment between the optical polarization and the active EO crystal axes, as well as the EO crystal temperature [6].

In the beginning of our study, the sensitivity of RAM to the alignment of the optical polarization was underappreciated. Coupled with the temperature sensitivity of the RAM, this can lead to large changes in the FM error signal baseline. Figure 3a shows an example of the demodulated atomic absorption signal taken at different times of day when the laser polarization entering the EO phase modulator was not fully optimized. As the figure clearly
shows, we observed a large change in the DC level from the double balanced mixer (DBM) over time. Given this magnitude of baseline shift, we would routinely observe 40 MHz ($1 \times 10^{-7}$) fluctuations of the beat note when the frequency lock loops were engaged. Consequently, the systematic effects imposed by RAM on the beat note can lead to frequency shifts of the locked laser that are far greater than those arising from atomic resonance fluctuations. This behavior is suppressed significantly, but not entirely eliminated, when the input polarization is optimized: the results shown in Fig. 2a were obtained for optimal laser polarization alignment.

![Graph showing FM error signal with drifting baseline from RAM](image)

*Fig. 3: Example of FM error signal with drifting baseline from RAM. Data collected periodically throughout a day. $t = 0$ (Black), $t = 3$ hours (Blue), $t = 5$ hours (Red).*

It is not possible to completely eliminate the impact of the RAM via polarization control alone. The temperature sensitivity of the RAM is also a significant systematic effect, which arises from the temperature dependence of the EO’s index of refraction. This naturally leads one to try to temperature stabilize the EO phase modulator. Unfortunately, our initial attempts to stabilize the EO phase modulator temperature, and therefore stabilize the RAM, have so far been unsuccessful. In our present system, the ovens for the EO modulators use the same solid state relay pulsed heaters as the Rb vapor cells. While the EO oven thermocouple shows ~ 0.1 K stability on average, the pulsed nature of the temperature control dynamics impart (RAM induced) frequency noise onto the stabilized lasers. We have found that the RAM exhibits a phase cycle (*i.e.*, change of sign) with less than a 0.5 K change in the temperature of the EO modulator case.

![Graph showing laser beat frequency oscillations](image)

*Fig. 4: a) Laser beat frequency showing sinusoidal oscillations as a single EO phase modulator oven cools from 28 °C to room temperature. b) By 10,000 s the EO oven has re-equilibrated to room temperature and the beat frequency is stable about 403.5 MHz for at least 1000s.*

This finding is in agreement with other researchers using LiNbO$_3$ modulators [6]. In fact, simply changing the set point on the EO phase modulator oven imparts clear sinusoids
in the measured laser beat frequency during the heating or cooling phase of the oven (Fig 4a).

The amplitudes of the beat note fluctuations shown in Fig. 4a are approximately 1.5 MHz as the RAM phase cycles due to the EO crystal temperature change. This data supports the notion that small room temperature fluctuations can lead to shifts in the stabilized laser frequencies on the order of 1 MHz due to the dependence of RAM on EO modulator temperature. In all likelihood, the behavior of the laser locking shown in Fig. 2 is driven by thermal fluctuations of the EO crystal and the associated variations in RAM. Due to the strong dependence of RAM on the EO phase modulator temperature, we have currently chosen to let the EO modulators run at room temperature (Fig. 4b) while we re-evaluate our EO oven design.

**Isoclinic Point Thermal Sensitivity**

In this section we demonstrate an isoclinic-pint stabilized laser’s frequency to changes in locking cell temperature. As stated above, the premise of this atomic thermometry relies on the fact that one laser is locked to an atomic resonance that has a near-zero temperature sensitivity, $\partial \omega_0/\partial T \approx 0$. The isoclinic point in the D1 spectrum of $^{87}$Rb midway between $F_g=2 \rightarrow F_c=1$ and $F_g=2 \rightarrow F_c=2$ has previously been shown to have a sensitivity less than $1 \times 10^{-11}/K$ and is theoretically better than $1 \times 10^{-12}/K$ [1].

We reconfirm these results here by locking a laser to the isoclinic point in a vacuum cell and ramping the temperature between 35 °C and 50 °C in 2.5 °C steps. The other laser is locked to the peak of the $F_g=2 \rightarrow F_c=1$ transition in a vacuum cell with the temperature stabilized at 40 °C. We record the beat frequency between these lasers during the temperature ramp. These results are shown in Fig. 5.

During the course of this 18 hour experiment the laser beat note undergoes a maximum frequency excursion of approximately 1 MHz. However, these excursions are not correlated with the temperature change of the cell used to lock the laser to the isoclinic point. These frequency excursions are most likely related to a shifting base line due to the un-stabilized (temperature-dependent) RAM.

![Fig. 5: Laser beat frequency during an 18 hour experiment where the vacuum cell used to lock one laser to the isoclinic point was continuous temperature modulated between 35 and 50 °C. The second laser was stabilized to $F_g=2 \rightarrow F_c=1$ in a vacuum cell with the temperature stabilized to 40 °C. The lack of correlation between the temperature ramp and the laser beat frequency is striking.](image)

Analysis of the data shown in Fig 5 indicates that the isoclinic point frequency shows a sensitivity to the Rb vapor temperature of $\pm 2$ kHz/K $\pm 7$ kHz/K. The uncertainty is taken at a 95% confidence level (2σ). Consequently, in terms of fractional frequency the isoclinic point is most likely temperature insensitive at a level less than $5 \times 10^{-12}/K$, which reconfirms our prior measurements [1]. We stress that this is the third attempt, using three different experimental techniques, to measure a temperature coefficient for the isoclinic point, and all of these attempts agree. In
particular, the most accurate of these attempts indicates that at a 95% confidence level the temperature coefficient of the isoclinic point must lie in the range $-4.8 \times 10^{-12}/K$ to $+0.8 \times 10^{-12}/K$ [1,3].

**Pressure Cell Thermal Sensitivity**

In this section, we turn our attention to the case where the atomic transition is expected to have a large temperature sensitivity. Specifically, it is well known that when a buffer gas is added to the Rb cell, the optical transitions undergo a temperature dependent pressure shift.

We have previously verified the theoretical prediction that the temperature dependence of the pressure shift scales as: $\Delta \omega (T) = \Delta \omega_0 \left( \frac{T}{T_0} \right) ^\kappa$ [3]. Here $\Delta \omega_0$ is the pressure shift at a reference temperature ($T_0$) and $\kappa$ is a constant that is 0.31 for nearly all noble gases except helium [8].

An additional temperature dependence occurs when locking to a peak in a closely spaced doublet, such as the $F_g=2 \rightarrow F_c=1$, $F_c=2$ manifold due to the overlap of the two peaks. We have used direct absorption spectroscopy to measure the combined sensitivity, and have found that locking to a peak in an overlapped doublet leads to a temperature sensitivity coefficient as high as $-0.6$ MHz/K ($-1.6 \times 10^{-9}$/K) for a Rb cell filled with 25 Torr of Kr [3].

For the purpose of demonstrating atomic thermometry, this feature is ideal as the temperature sensing noise floor will depend inversely on the atomic temperature sensitivity. We also stress that for the intended application of atomic thermometry (i.e., temperature stabilization of Rb vapor cell atomic clocks) the Rb resonance cells of the clock are filled with 10’s to 100’s of Torr of an inert buffer gas. Consequently, in such applications atomic thermometry can exploit the (relatively large) temperature sensitivity of Rb optical transitions to stabilize resonance cell temperature, and thereby combat the temperature shifts of the ground state hyperfine clock transition, which affect the device’s long-term frequency stability [9]. Here, we demonstrate the efficacy of atomic thermometry by assessing the correlation between laser beat frequency and cell temperature when a laser is locked to the $F_g=2 \rightarrow F_c=2$ transition in a Rb cell filled with 25 Torr of Kr. The other laser is locked to the isoclinic point in a vacuum cell stabilized to 40 °C, since this gives a temperature insensitive reference laser frequency (Fig. 5).

![Fig. 6: Laser beat frequency during a 12 hour experiment where the Rb cell that was filled with 25 of Kr was used to lock one laser to the $F_g=2 \rightarrow F_c=2$ absorption resonance peak, and was continuously temperature modulated between 30 and 55 °C. The second laser was stabilized to the isoclinic point in a vacuum cell with the temperature stabilized to 40 °C. The anti-correlation between the temperature ramp and the laser beat frequency is striking.](image)

The results of this experiment are shown in Fig 6. During the course of this 12 hour experiment we observed a strong anti-correlation between the 25 Torr Kr cell temperature and the lasers’ beat frequency. Curiously, around the middle of the data collection at 6 hours, the magnitude of the
temperature sensitivity appears to have decreased, and then spontaneously increased again at 8 hours. The root cause of this apparent non-repeatability is unclear. One likely explanation is the non-stationary (temperature-dependent) RAM and its effect on the laser locked to the 25 Torr Kr cell. The impact of RAM on this laser lock will be larger than that shown in Fig 2, since the pressure broadening of the optical transitions will lower the Q of the locking transition, thereby magnify the influence of RAM changes on the laser’s locked frequency.

the third temperature ramp cycle, the data fits to a line with a coefficient of $-118$ kHz/K. Finally during the last temperature ramp cycle, the initial behavior returns and the temperature coefficient returns to $-589$ kHz/K at 40 °C.

**Vapor Temperature Sensing Potential**

The data presented here show that a Rb cell filled with a buffer gas at several tens of torr, as would be used in a Rb vapor cell atomic clock, will likely show an optical transition frequency with a temperature coefficient of about 1 MHz/K (i.e., $3 \times 10^{-9}$/K). Therefore, if the lasers can maintain an inherent stability of 4×10^{-11} (best case shown in Fig. 2) over the course of an hour, then atomic thermometry should be able to assess hourly vapor temperature variations at the level of 10 mK. More optimistically, if laser stability can be made to average down to 8×10^{-12} at 100 s by improving the EO ovens, then the temperature stability capable with atomic thermometry could potentially reach 3 mK at two minute averaging times. We note that this laser frequency stability level is achievable for a laser stabilized to a Doppler broadened absorption line [10].

The current limitation to our implementation of atomic thermometry is the inherent frequency instability of the locked lasers. As shown in Fig. 2, 1 MHz shifts that are not associated with the atomic thermal sensitivities can occur due to systematics of the FM derived locking signal, and this limits our present temperature stability to no better than 1 K. However, since this limitation largely stems from RAM, we believe that a mitigation strategy for the elimination of RAM will allow our atomic thermometry to reach its potential.

Prior groups have used methods for detection and closed loop feedback on the EO

Fig. 7: Frequency shift thermal sensitivity for data shown in Figs. 5 and 6. The isoclinic point (blue squares) thermal frequency shift in a vacuum cell is $-5 \times 10^{-12}$/K ±1.9×10^{-11}/K at a 95% confidence interval. The thermal frequency shift sensitivity of the $F_y = 2 \rightarrow F_y = 2$ transition in an $^{87}$Rb cell with 25 Torr of Kr shows quadratic behavior (black diamonds and orange circles) with a sensitivity coefficient of -1.6×10^{-9}$/K at 40 °C. Due to some unknown processes, we have also measured the sensitivity for this transition (grey triangles) as linear with a thermal sensitivity coefficient of $-3 \times 10^{-9}$/K.

Nevertheless, we clearly observe anti-correlation during the entire experiment as expected. Figure 7 shows the analysis of this data over three different time periods. During the first two temperature ramp cycles, the data show a non-linear dependence of the beat frequency on the 25 Torr Kr cell temperature. We empirically fit this to a quadratic function and obtain a temperature sensitivity of $-664$ kHz/K at 40 °C. During
modulators to stabilize and eliminate RAM [6,11]. One method is to split the laser and set up an auxiliary FM detection line without the absorption cell [11]. This should be straightforward to implement. Other groups have noted that the RAM is correlated to polarization shifts of the light exiting the EO modulator, and have detected the RAM using polarization methods [6]. In both cases, RAM can be suppressed to acceptable levels. We plan on exploring these methods for RAM stabilization as we continue this work.

Summary
The concept of atomic thermometry utilizing the temperature insensitivity of the $^{87}$Rb D$_1$ isoclinic point has been demonstrated in a system consisting of two lasers frequency stabilized using FM spectroscopy. When a Rb vapor cell contains a buffer gas, as will be the case for vapor cell Rb clocks, the temperature sensitivity of the optical transitions can approach the 1 MHz/K level. Under these conditions, the thermal sensitivity can be exploited as an internal vapor thermometer by monitoring the beat frequency between a reference laser locked to the isoclinic point and a laser locked to the temperature sensitive optical transition. Our results suggest that accuracy down to the few mK level should be possible. The current limitation appears to be related to laser frequency instability imparted by RAM, a systematic imposed on the lasers’ stabilization by the EO modulators, and not an inherent atomic limitation. Prior methods used to stabilize and eliminate RAM in FM spectroscopy should prove useful for the work presented here.

Acknowledgements
This work was supported under The Aerospace Corporation’s Sustained Experimentation and Research for Program Applications (SERPA).

References
Spectroscopy and hyperfine clock frequency shift measurements in Cs vapor cells coated with octadecyltrichlorosilanes (OTS)

Moustafa Abdel Hafiz, Vincent Maurice, Ravinder Chutani, Nicolas Passilly, Christophe Gorecki, Stéphane Guérandel, Emeric De Clercq and Rodolphe Boudot

1FEMTO-ST, CNRS, Université de Franche-Comté, 26 chemin de l’épitaphe 25030 Besançon, France
2LNE-SYRTE, Observatoire de Paris, CNRS, UPMC, 61 avenue de l’observatoire 75014 Paris, France
Email: rodolphe.boudot@femto-st.fr

Abstract—We report the characterization using coherent population trapping (CPT) spectroscopy of an octadecyltrichlorosilane (OTS)-coated centimeter-scale Cs vapor cell. The linewidth of the narrow CPT resonance is compared to the linewidth of an evacuated Cs cell of similar size. The Cs-OTS adsorption energy is measured to be \((0.42 \pm 0.03)\) eV. A hyperfine population lifetime, \(T_1\), of 1.6 ms is reported, corresponding to about 37 useful bounces. Ramsey CPT fringes are detected using a pulsed CPT interrogation scheme.

I. INTRODUCTION

The performance of vapor cell atomic clocks or magnetometers is improved with increased relaxation time of the observed atomic state. In alkali vapor cells, two main techniques are used to slow down the depolarization of the atoms due to wall collisions. The first method is to confine the alkali atoms with a pressure of buffer gas in order to operate in the Dicke regime [1]. Alkali-buffer gas collisions result in a slow diffusive motion of atoms in the cell and contribute to increase the time for alkali atoms to collide against the cell walls. However, the use of buffer gas induces several drawbacks including a frequency shift and broadening of the optical resonance signal [2] but also a temperature-dependent frequency shift of the atomic clock transition frequency [3], [4], [5]. A second technique is to deposit an anti-relaxation wall coating onto the inner walls of the cell. The presence of this film material helps to make the atoms an increased number of bounces onto the cell walls before depolarization. Different materials were tested in the literature. Paraffin-coatings have demonstrated to support up to \(10^4\) atom-wall collisions [6]. Recently, Balabas et al. reported alkene-based coatings \((C_nH_{2n+1})\) demonstrating polarized alkali-metal vapor with minute-long transverse Zeeman population and coherence lifetimes in a 3 cm diameter cell. This corresponds to about \(10^6\) polarization preserving bounces [7].

A drawback of those coatings is their low-temperature melting point, preventing them to be used in chip scale atomic devices in which the microfabricated cell is produced using anodic bonding techniques that require typically high-temperature. Silane layers are found to be interesting candidates for this purpose. Seltzer et al. reported the measurement of up to 2100 collisions before the population relaxes \((T_1)\) in a K vapor cell with OTS multilayers [9]. 25 bounces were reported with OTS monolayers in Rb vapor [10]. Yi et al. demonstrated in a 8 mm side cubic Rb vapor cell up to 40 bounces with the cell walls before coherence relaxation and measured their adsorption energy to be 0.065 eV [11]. In 2014, Straessle reported a microfabricated vapor cell with OTS anti-relaxation coating and the demonstration of 11 surviving wall collisions [12].

This article aims to report basic characterization of a Cs vapor cell coated with OTS. We report the detection, spectroscopy and hyperfine clock frequency shift measurements of coherent population trapping (CPT) resonances in a centimeter-scale OTS-coated Cs vapor cell. Measurements of adsorption energy are given.

II. EXPERIMENTAL SETUP

Figure 1 shows the experimental set-up used to perform coherent population trapping spectroscopy and to characterize the cells.

![Experimental setup](image)

Fig. 1. Experimental setup.

The laser source is a distributed-feedback (DFB) diode laser tuned on the Cs D\(_1\) line at 894.6 nm [13]. The laser beam is sent into a pigtailed Mach-Zehnder intensity electro-optic modulator (MZ EOM - Photline NIR-MX800-LN-10). The latter is driven at 4.596 GHz by a low noise microwave frequency synthesizer. At the output of the EOM, light is sent through a Michelson delay-line and polarization orthogonalizer system. This allows to produce the so-called push-pull optical
pumping interaction scheme [14], [15] where atoms interact with an optical bichromatic field that alternates between right and left circular polarization. This technique allows to enhance the clock signal. The diameter of the output beam is expanded to 1.5 cm to cover the cell diameter. The cell is placed into a temperature-controlled oven. A homogeneous longitudinal static magnetic field is applied to isolate the clock transition. A magnetic field flux density value of 1.55 G for both the Cs cell and the Cs-OTS cell was applied. The assembly is surrounded by a double-layer mu-metal magnetic shield. A photodiode detects the light transmitted through the cell. The laser is frequency stabilized on the bottom of the Doppler-broadened optical line using a lock-in amplifier based (LA1) synchronous modulation-demodulation technique. CPT resonance spectroscopy is realized by scanning slowly the 4.596 GHz output frequency. The MZ EOM can be used as a light switch in order to make the atoms interact with a sequence of optical pulse trains. This technique is convenient to perform relaxation times measurements through the Franzen’s technique of relaxation in the dark [16] or to produce a Ramsey-like pulsed CPT interrogation.

We tested 2 different cylindrical vapor cells made of borosilicate glass. The first cell is a reference evacuated Cs cell with a diameter of 15 mm and a length of 15 mm. The second cell, with the same dimensions than the evacuated Cs cell, is equipped with OTS anti-relaxation coating and does not contain any buffer gas.

III. EXPERIMENTAL RESULTS

Figure 2 reports the CPT clock resonance in the Cs-OTS cell for a total incident laser power of 100 μW. The dual-structure of the dark resonance, signature of the anti-relaxation effect, is obvious. The pedestal of the resonance is characterized by a Doppler-broadened structure whereas the top of the resonance is narrowed thanks to the coating material. Figure 3 shows the linewidth of the CPT resonance in the Cs cell and the Cs-OTS cell versus the total laser power \( P \) for a cell temperature of 35°C. In a Doppler-broadened system, it can be shown that the CPT linewidth is proportional to the square root of intensity for low laser intensity of the driving field and is independent of the Doppler width [17]. At higher laser intensities, the CPT linewidth is increased linearly with the laser power. In our experiment, the CPT linewidth - laser power dependence curve shows two distinct regimes with a sudden linewidth decrease for laser powers lower than about 50 μW. The CPT linewidth extrapolated at null laser power in the evacuated Cs cell is measured to be 7 kHz, in excellent agreement with theoretical calculations yielding a total linewidth of 7.07 kHz (sum of spin-exchange and transit-time contributions).

For the Cs-OTS cell, the experimental width in Hz is well fitted by a linear function \( 634 + 8.6P \), with \( P \) in μW. The CPT linewidth extrapolated at null laser power is about 10 times narrower in the Cs-OTS cell compared to the pure Cs cell. In the Cs-OTS cell, the laser power broadening is about 13 times smaller than in the evacuated Cs cell.

The adsorption energy can be estimated from the measurement of the clock frequency shift \( \delta \nu \), defined as the difference between the actual clock frequency and the exact unperturbed Cs atom frequency (9.192631770GHz), versus the cell temper-

\[
|\delta \nu| \propto E_a \exp(E_a/kT)
\]

We measured the clock frequency versus the cell temperature (from 30 to 47°C) to extract the value of \( E_a \) (see figure 4).

For each cell temperature, the clock frequency is measured for different values of laser intensity and extrapolation to null laser intensity is performed. A negative frequency shift of a few hundreds of Hertz is measured as usually observed on wall-coated cells [11]. The shift rate is measured to be +24.6 Hz/K. According to Eq. (1), the plot of \( \ln |\delta \nu| \) against \( 1/T \) is a straight line of slope \( E_a/k \). This yields to \( E_a = 0.42 \) eV with a statistical uncertainty of 0.03 eV. This value is in good agreement with the one reported in [20], the only one we found in the literature for Cs-Pyrex-OTS.
Figure 4. Clock frequency shift (from the exact unperturbed Cs atom frequency (9.192631770GHz)) versus the cell temperature.

Figure 5 reports, for an incident laser power of 500 μW, the measurement of the relaxation time $T_1$ in the dark of the hyperfine level ($F = 4$) population. The Franzen’s technique is used. Experimental data are well fitted by an exponential decay function with a time constant $T_1 = (1.66 \pm 0.007)$ ms. Such a time constant corresponds to about 37 useful bounces.

Figure 6 shows a CPT-Ramsey fringe detected in the Cs-OTS cell. Atoms interact with light in a pulsed CPT sequence where the CPT pumping time $\tau_p$ is 2 ms and the free evolution time in the dark $T_R$ is 0.5 ms. The laser power is 650 μW. The beam diameter covers the whole cell diameter.

$E_a = (0.42 \pm 0.03)$ eV, in agreement with the value reported in the literature for Cs-OTS-Pyrex [20] by a completely different method. A clock frequency shift rate of 24.6 Hz/°C was measured. Measurements of hyperfine population lifetime ($T_1$) of about 1.6 ms were reported, corresponding to about 37 bounces. Ramsey CPT fringes were detected using a pulsed CPT interrogation scheme.

IV. CONCLUSION

This paper reported the characterization of a centimeter-scale Cs vapor cell wall-coated with octadecyltrichlorosilane (OTS) using CPT spectroscopy. The presence of a narrow peak Lorentzian structure on the top of a Doppler-broadened broad structure is a relevant signature of the OTS anti-relaxation coating. Using clock frequency shift measurements, the adsorption energy of Cs atoms on OTS surface was measured to be $\tau_p = 2$ ms, $T_R \approx 0.5$ ms. The laser power is 650 μW. The cell temperature is 35°C. The beam diameter covers the whole cell diameter.

ACKNOWLEDGMENT

The authors thank P. Bonnay (LNE-SYRTE) for the Cs filling of Cs vapor cells. M. Abdel Hafiz PhD thesis is co-funded by the Région Franche-Comté and the LabeX FIRST-TF (Facilities for Innovation, Research, Services, Training in Time & Frequency). V. Maurice PhD thesis is funded by the Délégation Générale de l’Armement (DGA) and the Région Franche-Comté. This work is supported in part by the Agence Nationale de la Recherche and the DGA (ISIMAC project ANR-11-ASTR-0004).

REFERENCES

Buffer Gas Consumption in Rubidium Discharge Lamps

Bernardo Jaduszliwer, Michael Huang, and James C. Camparo
Physical Sciences Laboratories
The Aerospace Corporation
El Segundo, CA, USA

Abstract—We present a physics-based empirical model of a newly discovered potential failure mode of rubidium atomic clocks: exhaustion of the noble gas buffer in the rubidium discharge lamp. We attribute the buffer gas loss to noble gas ion capture (NIC) by the glass walls of the lamp. The noble gas ions are produced by multistep ionization in collisions with discharge electrons. The model explains the observed pressure dependence of the buffer gas loss rate, and predicts an extremely high sensitivity of the loss rate to discharge electron temperature. That prediction is confirmed by comparison with experimental data. The model needs further work to be fully validated. We propose that longest lamp life can be achieved by minimizing noble gas light emission while keeping Rb light emission at the level required to achieve the desired atomic clock performance.

Keywords—Rubidium clocks; Discharge lamps; Buffer gas consumption

I. INTRODUCTION

Rubidium atomic clocks provide accurate timekeeping for many applications, including on board navigation satellites such as GPS, Galileo, QZSS and BeiDou. A typical rubidium clock physics package is shown schematically in Fig. 1. Light generated by $^{87}$Rb atoms in a discharge lamp is transmitted through filter and resonance cells containing $^{85}$Rb and $^{87}$Rb vapors, respectively, and then detected by a photodiode.

A typical discharge lamp is a quasi-cylindrical alkali-resistant glass bulb, approximately $10^{-2}$ m in diameter and $1.5 \times 10^{-2}$ m in length, surrounded by a RF exciter coil. It contains a few hundred micrograms of liquid $^{87}$Rb; it is filled to a pressure $p$ of a few torr of a buffer noble gas (typically Kr or Xe) and sits inside an oven held at $T \approx 380$ K, resulting in a fraction of a millitorr of Rb vapor. The inductively-coupled RF discharge inside the bulb is driven by an approximately 100 MHz current in the coil. Electrons oscillate in the rf field, and extract energy from the field through elastic collisions with the noble gas atoms. Rb atoms are ionized by successive electron collisions whose added kinetic energies eventually exceed 4.176 eV (the Rb ionization threshold), and $\text{Rb}^+$/electron recombination near the glass walls [1,2] ends up producing the Rb resonant light required to operate the clock. Unfortunately, the lamps slowly lose their buffer gas, giving rise to a life-limiting mechanism. This has been previously studied by Russian investigators [3,4,5], who collected a significant amount of experimental data. In particular, they measured loss rates for a set of widely spaced discharge powers $\mathcal{W}$ ranging between 1.7 W and 10 W, from which they postulated that the service life of the lamp ($T_{LAMP}$) scales with increasing discharge power $\mathcal{W}$ as $T_{LAMP} \propto 1/\mathcal{W}^{1.7}$ [3].

Typically, the discharge electron temperature $T_e \approx 3200$ K, so that the electron average kinetic energy $\langle E_e \rangle = kT_e \approx 0.3$ eV, where $k$ is Boltzmann’s constant. Since the ionization threshold of Xe is 12.127 eV, approximately 40 times larger than the average electron energy, one can expect that most of the ions in the plasma will be Rb ions. Since the buffer gas pressure is thousands of times higher than the Rb vapor pressure at 380 K, the plasma will be weakly ionized. Bulk plasma charge neutrality is preserved by the azimuthal symmetry of the electric field and ambipolar radial diffusion, except in a positively charged plasma sheath [6] that develops next to the glass surface, a few Debye lengths in thickness (i.e., a few tenths of a millimeter). It is worth remarking here that the lamp discharge is a highly nonlinear system, where changes in the exciter and oven control circuits drive changes in the discharge plasma, and changes in the discharge parameters couple back into the exciter circuit operation [7].

Data has recently become available in which lamp buffer gas pressure was measured as a function of operating time for lamps with different initial Xe fills. Our analysis of those data allowed us to develop a physics-based empirical model of the noble gas loss process which, if fully validated, should allow high confidence predictions of lamp life against buffer gas loss.

II. THE NOBLE GAS LOSS MECHANISM

Lamps with nominal Xe fill pressures $p_0 = 1$ torr, 2 torr and 3.5 torr were studied at nominal, as well as higher, RF discharge power $\mathcal{W}$. The buffer gas pressure was measured using a technique first described by Kazantsev et al. [8]: the lamp is...
inserted in lieu of the resonance cell in a rubidium atomic clock; since the measured clock frequency shift (relative to the Rb atom unperturbed clock transition frequency) is dominated by the noble gas density, it becomes a measure of the Xe pressure in the lamp. The data obtained from those studies showed that:

(i) As expected, the Xe loss rate increases with increasing discharge power.

(ii) As shown in Fig. 2, the consumption rate increases sharply with time (i.e., with decreasing Xe pressure).

After considering a number of possible buffer gas loss processes, including gas phase excited state Xe chemistry and Xe loss to the liquid Rb in equilibrium with the vapor, we concluded that capture by the glass lamp envelope was the only credible Xe loss mechanism. A variety of microanalytical techniques were used at The Aerospace Corporation to search for Xe trapped in lamp glass. Fig. 3 shows two examples of the results, which confirmed the capture by glass assumption.

A striking fact about the observed Xe loss to the lamp glass is that the loss of a noble gas atom to the glass wall is a very unlikely event. The weakly ionized discharge plasma can be considered, to a very good approximation, to behave like an ideal gas. The Xe number density will be \( n = p/kT \approx 6.5 \times 10^{22}/m^3 \). The mean thermal velocity of the Xe atoms will be \( \bar{v} = \sqrt{8kT/(\pi M)} \approx 250 \text{ m/s} \). If the probability of a Xe atom being lost in a collision with the wall is \( \gamma \), then, from ideal gas kinetic theory considerations, the number of Xe atoms lost per unit wall area and unit time is

\[
\frac{d^2N}{dAdt} = -\gamma \left( \frac{1}{4} n \bar{v} \right) \tag{1}
\]

Integrating over the area of the lamp wall yields

\[
\frac{dN}{dt} = -\frac{\gamma A \bar{v}}{4V} N \tag{2}
\]

where \( N \) is the total number of Xe atoms in the lamp, \( V \) is the internal lamp volume, \( V \approx 7 \times 10^{-7} \text{ m}^3 \), and \( A \) is the area of the lamp glass wall, \( A \approx 4.5 \times 10^{-4} \text{ m}^2 \). Since \( p \propto N \),

\[
\dot{p} = \frac{dp}{dt} = -\left( \frac{\gamma A \bar{v}}{4V} \right) p \tag{3}
\]

or \( \dot{p} = -(4V/A\bar{v})(p/p) \). For measured loss rates of the order of 2 millitorr/week at 2 torr, \( \gamma \approx 4 \times 10^{-14} \). Clearly, losing a Xe atom in a collision with the glass wall is an extremely low probability event. There are three plausible explanations:

(i) There are very few sites in the glass able to capture Xe.

(ii) Xe capture is a very high activation energy process.

(iii) Xe is not lost as ground state neutral Xe, but either as an excited Xe species or as Xe\(^+\), in either case present in very low numbers.

Fig. 2. Xe pressure vs. operating time for a nominal 1 torr fill Xe lamp

Fig. 3. Left: energy dispersive spectroscopy (EDS) map of Xe distribution on a lamp faceplate; lighter grey indicates higher Xe presence, looking down on the lamp’s faceplate; the curved edge extending from lower left to upper right marks the lamp’s wall. Right: X-ray photoelectron spectroscopy (XPS) Xe depth profile in a lamp barrel wall.
Another striking fact about the observed Xe loss to the lamp glass is its pressure dependence. The loss rate \( R = \text{d}p/\text{d}t \) for any process mediated by collisions of ground state neutral Xe atoms with the walls should decrease as the collision rate decreases, and thus \( \text{d}^2p/\text{d}t^2 > 0 \), since the atom-wall collision rate will be proportional to the decreasing Xe pressure. But the long term loss rate is seen to increase with decreasing pressure (i.e., \( \text{d}^2p/\text{d}t^2 < 0 \)), and so, the observed pressure dependence of the long term Xe loss is inconsistent with (i) and (ii) above, leaving (iii) as the only possible alternative.

We are proposing noble gas ion capture (NIC) by the glass surface as the most likely Xe loss mechanism. We will show that this mechanism explains the extremely low probability of Xe loss per collision with the wall, as well as the observed pressure dependence of the loss rate. The NIC mechanism also predicts a strong dependence of Xe loss on discharge electron temperature, which has been confirmed by our measurements.

In the NIC model, the Xe loss rate to the glass can be written in a most general way as

\[
\frac{\text{d}p}{\text{d}t} = \frac{kT}{V} \frac{\text{d}N}{\text{d}t} = -\frac{kT}{V} \int_{\vec{x}}^{\infty} \int_A \Phi(E, \vec{x}) \gamma(E, \vec{x}) \text{d}E \text{d}A \tag{4}
\]

where \( \Phi(E, \vec{x}) \) is the flux of Xe\(^+\) ions of kinetic energy \( E \) impinging on the glass surface at point \( \vec{x} \), and \( \gamma(E, \vec{x}) \) is the probability that a Xe\(^+\) ion of kinetic energy \( E \) impinging on the glass surface at point \( \vec{x} \) will be captured in the glass. \( \Phi(E, \vec{x}) \) will be determined by conditions in the discharge plasma, while \( \gamma(E, \vec{x}) \) will be determined by the properties of the glass surface (and possibly by the plasma conditions in its immediate vicinity). The area integral is carried out over the inner surface of the lamp.

A full calculation of \( \Phi(E, \vec{x}) \) would require complete modeling of a two-component (Rb and Xe) weakly ionized plasma in a non-homogeneous, non-isotropic electromagnetic field distribution, with ill-defined boundary conditions. Furthermore, it would require full knowledge of all the relevant collision cross sections: (e-,Xe), (e-,Rb), (Xe,Xe), (Rb,Rb) and (Rb,Xe). Such a calculation is well beyond the scope of a reasonable effort to understand and mitigate the Xe loss problem. Instead, we will use the results of empirical observations to guide us in the formulation of the simplest possible physics-based model capable of making reasonable lamp life predictions.

If we assume the electrons in the discharge to be fully thermalized at temperature \( T_e \), then their energies will fall within a Boltzmann distribution,

\[
f(E) = \frac{1}{\sqrt{2\pi kT_e}} \left[ \frac{E}{kT_e} \right]^{1/2} e^{-E/kT_e}. \tag{5}
\]

Fig. 4 shows the energy distribution of thermal electrons at a typical discharge lamp \( T_e \approx 3200 \text{ K} \). The fraction of electrons in the discharge having energies higher than some atomic energy threshold \( E_n \) can be obtained by integrating the Boltzmann distribution over energies \( E \geq E_n \):

\[
F(E \geq E_n) = 1 - \text{ERF}\left(\sqrt{E_n/kT_e}\right) + \frac{2}{\sqrt{\pi}} \sqrt{E_n/kT_e} e^{-E_n/kT_e}. \tag{6}
\]

Fig. 5 shows the atomic energy levels diagram for Xe. The Xe ionization threshold is 12.127 eV, and the first excitation threshold is 8.3204 eV. Thus, for \( T_e = 3200 \text{ K} \) the fraction of electrons capable of exciting ground state Xe by collision is given by the area under the curve.
We will show that \((m+1)\)-step ionization leads directly to a \(p^m\) dependence of the loss rate on Xe pressure. We also note that the first step up the ladder limits the ionization rate; it will happen very rarely for \(T_e \approx 3200\) K because of the very small value of \(F_{exc}\) explaining the very low probability of Xe loss to the wall.

The process we envision is illustrated in Fig. 6, indicating the competition between \((e^-,\text{Xe})\) collisional excitation and \((\text{Xe}^*,\text{Xe})\) collisional de-excitation. Specifically, the probability of ionization after \((m+1)\) collisions with electrons will be given by

\[
P_+ = P_{0,1} \prod_{k=1}^{m} P_{k,k+1}
\]

where \(P_{k,k+1}\) is the probability of an atom being promoted by a collision with an electron from state \(k\) to state \(k+1\). \(P_{0,1}\) will be independent of Xe pressure, since a ground state atom cannot be demoted to a lower energy state by a collision with another Xe atom. For \(k \neq 0\), \(P_{k,k+1}\) will depend explicitly on pressure, since the atom in state \(k\) could be demoted to a lower energy state by a collision with another Xe atom before colliding with an electron and being promoted to state \(k+1\).

The probability of an atom colliding with an electron of enough energy to excite it to a higher energy level between \(x\) and \(x+dx\) will be \(dP_+ = (1/\zeta)e^{-x/\zeta}dx\), where \(\zeta\) is the appropriate collision mean free path. The probability that the excited atom will then not collide with another Xe atom between \(x\) and some other point \(z\) where the next collision with an electron will promote it to the next step in the excitation ladder will be \(P_0(z-x) = e^{(z-x)/\lambda}\), where \(\lambda\) is the mean free path for collision with another atom. Thus, the probability \(P_{k,k+1}(z)\) that an atom in state \(k\) will have been promoted to the next step in the ladder at \(z\) will be

\[
P_{k,k+1}(z) = \int_0^z \frac{e^{x/\zeta}}{\zeta} e^{(z-x)/\lambda} dx = \frac{e^{x/\zeta}}{\zeta} \int_0^z e^{x/\zeta} dx
\]

where \(1/\xi = 1/\lambda - 1/\zeta\). After integration over \(x\),

\[
P_{k,k+1}(z) = \frac{\xi}{\zeta} \left[ e^{z/\xi} - e^{z/\lambda} \right].
\]

Since the plasma is weakly ionized, for the range of Xe pressures of interest \(\zeta >> \lambda\), and thus \(\xi/\zeta \approx \lambda/\zeta\), and \(e^{z/\xi} - e^{z/\lambda} \approx e^{z/\lambda}\). In general, each atomic state \(k\) will see different mean free paths \(\lambda_k\) and \(\xi_k\) for collisions with other Xe atoms and with electrons capable of exciting it to state \(k+1\), respectively. We can estimate \(P_+\) by letting the \(\{z_k\}\) take their expectation values, \(z_k = \langle z_k \rangle = \xi_k\). Thus,

\[
P_+ \approx P_{0,1} \prod_{k=1}^{m} \lambda_k e^{z_k/\xi_k} = P_{0,1} \frac{e^{m/\lambda}}{n_m} \prod_{k=1}^{m} \frac{1}{\sigma_k \lambda_k}
\]

where \(\lambda_k = 1/n \sigma_k\), and \(\sigma_k\) is the total de-excitation cross section for a Xe atom in state \(k\) by collision with a thermal ground state Xe atom. Since the Xe pressure \(p\) is proportional to the Xe number density \(n\), the probability of \((m+1)\)-step ionization (and thus the Xe loss rate) will be proportional to \(p^m\), explaining the observed increase in loss rate with decreasing pressure.

Thus, the NIC mechanism explains the main observed characteristics of the Xe loss process. As we will show, a simple empirical consumption model based on the NIC mechanism is consistent with all the relevant data on Xe consumption in lamps, although additional physical processes may need to be introduced when examining consumption in high power discharges.

---

1 Here we are assuming that the probability of radiative decay is much smaller than the probability of collisional de-excitation. This assumption is based on expected effective radiative lifetimes \(\geq 10^5\) s, and average times between \((\text{Xe},\text{Xe})\) collisions \(\leq 3\times 10^8\) s.
III. THE NIC MODEL AND LAMP XE PRESSURE

Fig. 7 shows \( p(t) \) data obtained for a lamp with an initial Xe fill of approximately 1 torr, operated at nominal power. The \( p(t) \) data in Fig. 7 displays the behavior characteristic of most lamps. There is a beginning-of-life (BOL) loss characterized by its positive curvature (i.e., the loss slows down with time)\(^2\), and a longer term behavior characterized by its negative curvature (i.e., the loss accelerates with time). The BOL process can be modeled by an exponential, \( \Delta p_{\text{BOL}} \exp(-t/\tau) \), added to the long term model. The time constant \( \tau \) is of the order of 20 days, and \( \Delta p_{\text{BOL}} \) is a few percent of the Xe fill pressure. After the BOL process losses become negligible, and before the data begin to display an appreciable negative curvature, there is a long period during which the \( p(t) \) data can be fit rather well by a straight line (i.e., a “constant” loss rate). While the data obtained in the quasi-linear regime could possibly be used to predict lamp life using a validated physical consumption model, it has no value as a discriminant between different consumption models. The higher the lamp fill pressure, the longer this quasi-linear regime will last. None of the lamps with initial 2 torr and 3.5 torr nominal fills were tested for a long enough period to begin displaying the negative curvature behavior. That is why we will focus our analysis of the pressure vs. time lamp behavior on lamps with 1 torr nominal Xe fills.

The NIC model predicts a \( R = \kappa / p^m \) dependence of the Xe loss rate \( R \) on Xe pressure \( p \), where \( m \) is the number of additional electron collisions\(^3\) required to ionize a Xe atom after it has been excited to one of the \( 5p^56s \) states at \( e_i \approx 8.3 \text{ eV} \). \( R \) will be proportional to the probability of multistep Xe ionization, \( P_+ \), and thus, by Eq. (10) \( \kappa \) will be proportional to \( P_{0,1} e^{-m} \prod_{k=1}^{m} 1/\zeta_k \sigma_k \). Of course, the \( R \propto 1/p^m \) behavior cannot continue indefinitely as \( p \to 0 \). As a matter of fact, \( R \) must stay smaller than \( \left( \frac{\Delta V}{\Delta V} \right) P_0 \), which would be the loss rate if every Xe atom that collides with the wall is lost. Thus, \( |d^2p/dt^2| \) must start decreasing as \( p \) becomes very low. As we will see, that behavior is actually observed for lamps operating at very low pressures. We will postulate the simplest possible pressure dependence for \( R \) that displays a high pressure asymptote \( \propto 1/p^m \) and a low pressure asymptote \( \propto p \), characteristic of a collisionless regime\(^4\):

\[
R = -\frac{dp}{dt} = \frac{\kappa p}{p^{m+1} + p^{m+1}} \tag{11}
\]

The critical pressure \( p_C \) is the pressure at which the \( \text{(Xe,Xe)} \) collision mean free path \( \lambda \) becomes larger than some critical distance \( D_c \). Fig. 8 shows a generic example of the \( R(p) \) given by Eq. (11). This pressure loss rate equation can be rewritten as \( p^m dp + p^{m+1} d[\ln(p)] = -\kappa dt \); integration yields an implicit expression for the time-dependence of the pressure in the operating lamp:

\[
t(p) = \frac{1}{\kappa} \left[ \frac{p_{0,1}^{m+1}}{m+1} - \frac{p^{m+1}}{m+1} - p_C^{m+1} \ln \left( \frac{p}{p_0} \right) \right] \tag{12}
\]

\(^2\) The BOL loss displays the \( d^2p/dt^2 > 0 \) expected for a process in which Xe is lost via collisions of neutral ground state Xe atoms with the wall. If this is true, the number of sites capable of trapping Xe atoms must be relatively small and must saturate, since that loss process seems to become negligible after a small fraction of the Xe in the lamp is lost. Alternately, what appears as Xe loss may not at all be a process involving Xe, but the loss of some low concentration gas phase contaminant that has the same sign pressure shift as Xe, being consumed either by gas phase chemistry in the discharge or by loss to the wall. Or it might represent evolution from the walls of a gas having the opposite pressure shift sign as Xe. Even if it indicates actual Xe loss, the observed BOL loss mechanism would be unimportant, since it results in a total loss of just a few percent of the initial Xe fill.

\(^3\) Since there are multiple paths taking an excited Xe atom in one of the \( 5p^56s \) states to \( Xe^+ \), each requiring a different number \( m \) of steps, the empirical exponent of \( p \) in the observed loss rate \( R \) will be a properly weighted average of the different \( m \)'s for all the possible paths to ionization.

\(^4\) Of course, this is a gross oversimplification. Among other issues, elastic \((e,\text{Xe}) \) collisions are necessary for the electrons to extract energy from the RF field. If the Xe pressure becomes too low, the discharge electrons will start cooling too much. Our purpose is not to develop an accurate description of Xe loss at very low pressures, which is irrelevant to real lamp operation, but to insert a reasonable low pressure cutoff to the \( 1/p^m \) loss behavior.
where \( p_0 \) is the lamp pressure at \( t = 0 \). As we will see, Eq. (12) fits the available pressure vs. time data very well.

The directly measurable BOL lamp parameters are \( p_0 \) and \( R_0 \). At \( t = 0 \), \( R_0 = \kappa p_0 / (p_0^{m+1} + p_C^{m+1}) \approx \kappa / p_0^m \), or \( \kappa \approx R_0 p_0^m \). Fig. 9 shows the expected pressure vs. time life histories for lamps filled at the same initial Xe pressure \( p_0 \) and displaying the same initial consumption rate \( R_0 \), where different discharge characteristics would lead to different exponents \( m \) and critical pressures \( p_c \). This figure illustrates the difficulties involved in predicting lamp life based on just lamp screening data: during the first 100 weeks of operation, the consumption of the four lamps shown in the figure are indistinguishable from each other, even though the lamps following an \( m = 5 \) behavior would end up lasting 50% longer than the lamps following the \( m = 8 \) behavior. One would have to screen lamps for about 200 weeks to establish the actual value of the exponent \( m \), and by that time most of the useful life of the lamp would have been used up. Only a reasonably accurate model of noble gas consumption, based on the understanding of the consumption physics, will enable the use of lamp screening data to predict discharge lamp lifetimes.

If we assume that the discharge electron temperature does not depend strongly on Xe pressure, then identical lamps with different initial fills should display similar \( \kappa \)'s and similar exponents \( m \). Fig. 10 shows the effect of initial Xe fill on otherwise identical lamps, illustrating the very large increases in predicted lamp life \( T_{\text{LAMP}} \) as the Xe initial fill pressure is increased from 1 torr to 3.5 torr.

Four 1-torr nominal fill lamps were tested for a long enough time to allow the data to have value as a discriminant between different loss models. Fig. 11 shows the pressure vs. time data and corresponding NIC model fits for those lamps.

Note that the NIC exponents \( m \) derived for lamps operated at nominal power range between 7 and 9, while for both lamps operated at five times nominal power we obtained \( m = 4 \). This observation is consistent with the basic physical assumptions behind the NIC model: as the discharge power increases, the electron temperature should increase, which then favors ionization paths involving fewer steps (i.e., smaller \( m \)).

Fig. 12 shows a log-log plot of the consumption rates derived from the \( p(t) \) data, \( R \approx \Delta p/\Delta t \), for lamp (a) in Fig. 11. In a log-log plot, a simple power law dependence of \( R \) on \( p \), \( R \propto p^{-m} \), would show as a straight line of slope \( -m \). While the higher
pressure \( \Delta p/\Delta t \) data falls on such a line (of slope -8), it is clear that as the pressure becomes lower, \( R \) departs from simple power law behavior. Also shown is the NIC model fit to \( R(p) \), given by Eq. (11).

IV. THE NIC MODEL AND DISCHARGE ELECTRON TEMPERATURE

As discussed in Section II, the probability of a Xe atom becoming ionized (and thus, in the context of the NIC model, the Xe loss rate \( R \)) will be proportional to \( P_{0,1} \), the probability of a Xe atom being excited by electron impact to one of the 5p6s states. That, in turn, will be proportional to \( F_{ex} \), the fraction of electrons in the discharge having \( E_e \geq 8.3 \) eV. Fig. 13 shows \( F_{ex} \), calculated using Eq. (6), as a function of \( T_e \). It is apparent that \( F_{ex} \) is a very sensitive function of \( T_e \), increasing \( 350 \) K, from 3200 K to 3550 K, will increase \( F_{ex} \) (and thus the Xe loss rate \( R \)) by a factor twenty.

---

Fig. 11. Four lamps operated until failure by Xe exhaustion. Black dots are measurements; red curves are NIC fits. The fit parameters \( m, \alpha \) and \( p_c \) are given for each case follow. (a) 1 torr fill, operated at nominal power; \( m = 8, \alpha = 2.4 \times 10^{-4} \) torr/hr, \( p_c = 1.2 \) torr. (b) 1 torr fill, operated at nominal power; \( m = 9, \alpha = 4.5 \times 10^{-4} \) torr/day, \( p_c = 0.8 \) torr. (c) 1 torr fill, operated at 5 times nominal power; \( m = 4, \alpha = 1.25 \times 10^{-4} \) torr/hr, \( p_c = 0.49 \) torr. (d) 1 torr fill, operated at 5 times nominal power; \( m = 4, \alpha = 1.8 \times 10^{-4} \) torr/hr, \( p_c = 0.37 \) torr.

Fig. 12. Xe loss rate vs. pressure for lamp (a) in Fig. 11; the black dots are derived from the \( p \) vs. \( t \) data, and the red curve is the corresponding NIC model.
We performed spectroscopic measurements on lamps operating at nominal power to determine their electron temperatures, following the approach described by Camparo and Fathi [11] and Encalada et al. [12]. Briefly, light emitted by the lamp is spectrally analyzed by an Ocean Optics fiber-coupled spectrometer that measures the relative intensities $S_{ij}$ of $|i\rangle \rightarrow |j\rangle$ Xe optical transitions. The transitions used in this case are listed in Table 1. In a simple model of multistep Xe excitation, we can write an expression relating the relative intensity $S_{ij}$ for the $|i\rangle \rightarrow |j\rangle$ optical transition to the electron temperature:

$$\ln \left[ \frac{S_{ij}}{g_i A_{ij}} \left( \frac{\epsilon_i}{\epsilon_i - \epsilon_1 + kT_e} \right) \right] \sim \frac{\epsilon_i}{kT_e}.$$  \hspace{1cm} (13)

Here (Table 1 and Eq. (13)), $\Lambda$ is the transition wavelength, $g_i$ is the degeneracy of the upper state of the transition, $A_{ij}$ is the Einstein-A coefficient, $\epsilon_i$ is the upper-state energy, $T_e$ is the electron temperature, and $T_o$ is a zeroth-order estimate of $T_e$, which we take as $\approx 4500$ K. The model assumes that the excitation of the upper state $|i\rangle$ of the transition occurs mostly via a two-step process, where the first step is the excitation of one of the $5p^6s$ states of Xe, at $\epsilon_1 \approx 8.3$ eV. Using (13), we perform a linear least-squares fit of the $S_{ij}$ vs. $\epsilon_i$ data for the five Xe spectral lines listed in Table 1, and obtain $T_e$ from the slope. Fig. 14 shows the results obtained for sixteen lamps operating at nominal power.

Since all the lamp ovens and exciter circuits are identical, one would expect that the discharge electron temperatures in all these lamps would be similar. However, for these sixteen lamps, the electron temperatures are dispersed over a 400 K range. If we choose to compare only electron temperatures of lamps having the same pressure, we still see significant dispersion (200 K for the 2.5 torr lamps and 150 K for the 3.5 torr lamps). Lamp #1, which showed the highest rate of Xe consumption, also displays the highest $T_e$.

Xe loss rates for the same sixteen lamps were also determined. Fig. 15 shows the measured loss rates vs. measured electron temperatures. It also shows $F_{ex}(T_e)$, arbitrarily positioned to match the set of measurements at each of the three nominal fill pressures. It can be seen that, for each nominal fill pressure, $F_{ex}(T_e)$ represents fairly well the increase in Xe loss rate $R$ with increasing discharge electron temperature $T_e$.

---

The actual value of $T_o$ used in Eq. (15) has a very small effect on the derived values of $T_e$, and it is not critical.
Since an increase in the discharge RF power will increase $T_e$, the NIC model predicts a high sensitivity of Xe consumption to discharge power. Some evidence supporting this prediction was provided by the consumption data. All the lamps were operated at $W = 1.3\times$nominal power (i.e., 1.3X) for 1000 hours, and then operated at $W =$ nominal power (i.e., 1.0X) for an additional 1500 hours. Those data allow us to estimate loss rates for each lamp at 1.3X and 1.0X. Fig. 16 shows the 1.3X-to-1.0X ratios of loss rates, averaged over all lamps filled at the same nominal fill pressure. The data in Fig. 16 show a very high sensitivity of the Xe consumption to discharge power. A 30% increase in power results in a factor 3 to 4 increase in Xe loss rate. This increase is much larger than the 1.56 factor obtained using the $R \propto W^{1.7}$ power law reported by Gevorkyan et al. [3], derived from data obtained at widely spaced discharge RF powers.\(^6\)

Finally, we observe that $T_e$ depends not just on RF discharge power, but also on RF frequency. The time-averaged kinetic energy of a free electron placed initially at rest in an RF electric field, $E\cos(\omega t)$, will be $\langle E \rangle = e^2 E_0^2/4M\omega^2$, where $e$ and $M$ are the electron charge and mass, respectively, so we would expect $T_e$ to be proportional to $1/\omega^2$. In reality, energy transfer from the RF field to the electron is mediated by elastic collisions with Xe atoms, and so the collision frequency will play a very important role in determining electron average kinetic energy (and thus electron temperature). However, all other things being equal, we could expect to see electron temperatures falling with increasing RF field frequency; N. Encalada et al. verified this behavior [12], and their data are shown in Fig. 17. The data are quite telling: within the explored frequency range (65 MHz to 85 MHz), $T_e \propto 1/\omega^2$, and increasing the RF frequency by 20 MHz lowers $T_e$ by 600 K.

V. CONCLUSIONS

A reasonably accurate model of noble gas consumption, based on the understanding of the consumption physics, could be valuable to assure that rubidium atomic clock discharge lamps will meet their anticipated lifetimes. This would be particularly true for rubidium atomic clocks used on space missions. We believe the physics-based, empirical NIC model presented in this report is reasonably accurate, and could be used to assess (and possibly extend) expected lamp operating lives. The model is consistent with the available data, and it is the only proposed model that explains the observed pressure and electron temperature dependences of the Xe loss rate. The Russian investigators that first studied noble gas loss in Rb

\(^6\) Large increases in RF discharge power might trigger other physical processes that could actually reduce to some extent the net noble gas loss. For instance, heating of the glass wall could result in the release of previously trapped Xe.
discharge lamps made a very important observation that validates the basic hypothesis behind the NIC model (i.e., that noble gas is lost to the glass via stepwise ionization): “Changes in buffer gas pressure also were not observed in the lamps in which the pure rubidium discharge was excited without the excitation of krypton” [13]. If noble gas light is not emitted by the discharge, that means that the first steps in multistep ionization are not taken; noble gas ions are not produced, and thus they cannot be lost.

An important result of the NIC model is that the Xe loss rate is not a property of the bare lamp-bulb; the loss rate is extremely sensitive to the electron temperature, and thus can only be properly assessed for a complete lamp-bulb/RF-exciter/lamp-oven assembly operating at the design RF power, RF frequency and oven temperature.

As discussed in Section III, in the NIC model,

$$ R \sim \int_{\mathbb{A}} dA \int_{\mathbb{E}} dE \Phi(E, \bar{x}) \gamma(E, \bar{x}) $$

and thus there are two ways of reducing the Xe loss rate:

(i) Minimizing \( \gamma \), the probability of loss in an ion collision with the glass wall, or

(ii) Minimizing \( \Phi \), the flux of Xe ions reaching the glass wall.

Choosing option (i) requires controlling the glass surface properties to minimize \( \gamma \). That would require extensive research on the physics of Xe+ capture in the glass. The required experimental work could be quite difficult to perform. Additionally, even if we knew what to do to the glass surface to minimize Xe+ capture, it would probably be difficult to implement the required level of process control, given the artisanal character of typical discharge lamp manufacturing. We already know what is required to implement option (ii): operating the lamp at the lowest electron temperature, consistent with emission of the required intensity of Rb light, will minimize \( P_{0,1} \) and thus minimize Xe loss. In practice, this can be achieved by minimizing noble gas light emission while keeping Rb light emission at the level required to achieve the desired atomic clock performance.

REFERENCES


mSTAR: Testing Special Relativity in Space Using High Performance Optical Frequency References

Thilo Schultd∗, Shailendhar Sarat†, Alberto Stochino†, Klaus Döringshoff‡, Sasha Buchman†, Grant D. Cutler†, John Lipa†, Si Tan†, John Hanson§, Belgacem Jaroux§, Claus Braxmaier*, Norman Gürlebeck∥, Sven Herrmann∥, Claus Lämmerzahl∥, Achim Peters†, Abdul Alfaouwaz¶, Abdulaziz Alhussien§, Badr Alsuwaidan¶, Turki Al Saud¶, Hansjörg Dittus*, Ulrich Johann**, Simon P. Worden§ and Robert Byer†

∗Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Raumfahrtsysteme, Bremen, Germany
†Stanford University, Hansen Experimental Physics Laboratory, Stanford, California/USA
‡Humboldt-Universität zu Berlin, Institut für Physik, Berlin, Germany
§NASA Ames Research Center (ARC), Mountain View, California/USA
¶King Abdulaziz City for Science and Technology (KACST), Riyadh, Saudi-Arabia
∥Zentrum für Angewandte Raumfahrttechnologie und Mikrogravitation (ZARM), Universität Bremen, Germany
**Airbus Defence and Space GmbH, Friedrichshafen, Germany

Email: thilo.schultd@dlr.de

Abstract—The proposed space mission mini Space-Time Asymmetry Research (mSTAR) aims at a test of special relativity by performing a clock-clock comparison experiment in a low-Earth orbit. Using clocks with instabilities at or below the 1·10⁻¹⁵ level at orbit time, the Kennedy-Thorndike coefficient will be measured with an up to two orders of magnitude higher accuracy than the current limit set by ground-based experiments. In the current baseline design, mSTAR utilizes an optical absolute frequency reference based on molecular iodine and a length-reference based on a high-finesse optical cavity. Current efforts aim at a space compatible design of the two clocks and improving the long-term stability of the cavity reference. In an ongoing Phase A study, the feasibility of accommodating the experiment on a SaudiSat 4 bus is investigated.

I. INTRODUCTION

Special Relativity is classically tested by performing three types of experiments, investigating the orientation-dependence of the speed of light (Michelson-Morley experiment), the boost-dependence of the speed of light (Kennedy-Thorndike experiment) and the effect of time dilation (Ives-Stilwell experiment). To date, the test with the lowest demonstrated accuracy is the Kennedy-Thorndike (KT) experiment. The proposed mini Space-Time Asymmetry Research (mSTAR) space mission will perform a KT experiment in space by comparing an absolute (iodine-based) frequency reference to a length-based frequency reference (i.e. a laser frequency stabilized to a cavity) – both with frequency instabilities at or below the 1·10⁻¹⁵ level at orbit time. This allows to determine the Kennedy-Thorndike (KT) coefficient with an up to two orders of magnitude higher accuracy than current ground-based experiments [1].

Performing the experiment in space offers mainly two advantages. The velocity modulation is a factor of ten higher, compared to a ground based experiment and the (putative) science signal is shifted to Fourier frequencies where the stability of oscillators is better compared to sidereal frequencies. Further, space offers a vibration free environment and elimination of large DC gravity forces. The velocities relevant for a KT experiment are shown in figure 1. For the measurements on ground, the relevant laboratory velocity has contributions from the motion of the sun (relative to the cosmic microwave background, 370 km/h), the Earth’s orbital motion around the sun (30 km/s) and the Earth’s daily rotation (approx. 300 m/s). For a space-based experiment in low-Earth orbit, the orbital velocity is about 7.4 km/s with an orbit time of 90 min.

In the baseline design, mSTAR utilizes an absolute frequency reference based on a hyperfine transition in molecular iodine near 532 nm. A frequency-doubled Nd:YAG laser is foreseen as laser that is stabilized to the iodine reference. Part of the fundamental (1064 nm) stabilized laser light is split off and sideband locked to the resonance frequency of a high finesse optical cavity made of ultra-low expansion (ULE) glass using an electro-optic modulator (EOM). This way, the frequency difference between the absolute and the length reference can be extracted from the EOM sideband frequency, which is then analyzed with respect to variations at the orbit frequency for obtaining the KT coefficient.

The mSTAR iodine clock is based on a DLR-funded setup on Engineering Model (EM) level, realized using specific
assembly-integration technologies. A frequency stability below $5 \cdot 10^{-15}$ at integration times between 10 s and 5000 s was demonstrated. The length reference is based on the space-qualified cavity setup under development at JPL within the GRACE follow-on mission [2]. A design with adapted thermal shielding required for improved long-term stability and fiber coupling to the cavity is currently realized at Stanford University.

The mSTAR mission is investigated in an international collaboration including the King Abdulaziz City for Science and Technology (KACST, Riyadh, Saudi-Arabia), Stanford University (USA), NASA Ames (USA) and a German Team consisting of the German Aerospace Center (DLR Institute of Space Systems, Bremen), the Center of Applied Space Technology and Microgravity (ZARM, University Bremen) and the Humboldt-University Berlin. In an ongoing Phase A study, the feasibility of the payload accommodation within the SaudiSat 4 satellite bus is evaluated.

We will give an overview over the mission outline and the scientific measurement. Special emphasis will be placed on the optical frequency references that are foreseen for the mission.

II. MISSION AND SPACECRAFT BUS OVERVIEW

The proposed baseline orbit for mSTAR is a circular 6 AM dawn-dusk sun-synchronous low-Earth orbit with an altitude of 650 km. A low altitude is beneficial for the KT experiment as it provides larger boost velocity vector variation in a short period of time. The 650 km altitude minimizes the need for radiation hardening and hardened components and also provides a natural de-orbiting within 25 years. The sun-synchronous orbit provides good thermal stability of the payload, required by the mission, by maintaining an approximately fixed orientation of the satellite relative to the sun.

A ground station at KACST will be used for primary communications with the mSTAR satellite. The Science Operations Center (SOC) will be located at Stanford University with a backup SOC located at KACST and an auxiliary Science Center at ZARM, receiving flight data in near real time.

The spacecraft bus will be contributed by KACST, based on their SaudiSat 4 spacecraft with an envelope size of $672 \times 606 \times 1227$ mm$^3$. The spacecraft provides $\sim 100$ W of electrical power for a scientific payload with a mass up to 30 kg and a volume up to 140 liters. The attitude control system will use sun-vector information from sun-sensors and magnetic field information from a three-axis magnetometer. Reaction wheels are not foreseen as they introduce undesired vibrations possibly affecting the cavity performance.

III. PAYLOAD OVERVIEW

An overview over the mSTAR payload is given in the functional diagram shown in figure 2. One solid-state Nd:YAG laser with a wavelength of 1064 nm is used as light source for both frequency references. The iodine reference consists of a beam preparation unit, a spectroscopy unit and corresponding locking electronics. The laser output is sent to the beam preparation unit generating pump and probe beam for the iodine spectroscopy. A secondary laser output is delivered to a modulation bench preparing the laser light for the cavity setup.

As standard Pound-Drever-Hall method for cavity frequency stabilization can not be applied due to the frequency offset between the iodine and cavity resonance frequencies, an electro-optic modulator is used employing sideband locking. The feedback signal to the EOM sideband frequency is analyzed with respect to a possible KT signal at orbit frequency.

A. Iodine based frequency reference

The development of a space compatible iodine frequency reference setup is currently under way in a cooperation of ZARM Bremen, DLR Bremen, University of Applied Sciences Konstanz, Airbus D&S Friedrichshafen and the Humboldt-University Berlin [3]. A compact and ruggedized setup of the spectroscopy unit on engineering model (EM) level was realized, see figure 3. The optical components are joined to a fused silica baseplate using adhesive bonding technology in combination with a space-qualified two-component epoxy. This technique allows for higher long-term stability of the iodine frequency reference due to reduced pointing instability, which is a limiting effect in standard setups. The setup takes into account space mission related criteria such as compactness, MAIVT (manufacturing, assembly, integration, verification and test) and robustness with respect to shock, vibration and thermal stress. It utilizes a specifically designed multi-pass iodine cell in nine-pass configuration. The cell has dimensions of $100 \times 100 \times 30$ mm$^3$ resulting in an interrogation length of approximately 90 cm and utilizes a specifically designed robust cold finger design. With this setup, a frequency stability of $7 \cdot 10^{-15}$ at an integration time of 1 s and below $5 \cdot 10^{-15}$ at integration times between 10 s and 5000 s, was demonstrated in a beat measurement with a second laboratory setup of an iodine frequency reference. This frequency stability is similar to the one of the best current state-of-the-art laboratory setup of an iodine-based frequency reference [4].

The proposed baseline laser source for mSTAR is a Nd:YAG non-planar ring oscillator (NPRO) laser with a wavelength of 1064 nm and 65 mW output power provided by Tesat.
Spacecom. This type of laser has high technology readiness level; they are baseline for other space missions like Aladin, ATLID, GRACE follow-on, LISA pathfinder and LISA. It is operated in space aboard the TerraSAR-X laser communication terminal. Scientific laboratory versions are available from JDSU Corp. and Coherent, Inc., and are used in our EM setup.

Within the beam preparation unit, the laser light is split into pump and probe beam. Each beam is passing an acousto-optic modulator (AOM) for frequency modulation, RAM control, intensity stabilization and for generating a frequency offset between pump and probe beam. Afterwards both beams are frequency-doubled using second-harmonic generation (SHG) waveguide modules. The current setup uses modules provided by NTT electronics.

B. Cavity based frequency reference

The cavity development for mSTAR is based on the development of a space-qualified setup for the GRACE follow-on laser ranging instrument (LRI) carried out by JPL together with Ball Aerospace (cf. figure 4). For the scientific goals, a cavity setup with a frequency instability at or below $1 \times 10^{-15}$ at orbit time is aimed for. This long-term stability requirement especially demands for a very sophisticated thermal design.

The mSTAR baseline design foresees a similar designed mid-plane mounted cavity with a finesse $> 160.000$ made of ultra-low expansion (ULE) glass with a coefficient of thermal expansion (CTE) of $\sim 10^{-9}/K$ within an operating temperature range of $10 – 30^{\circ}C$ and a CTE zero crossing near $15^{\circ}C$. By operating the cavity close to the CTE null, the effective CTE can be further decreased. Mirror substrates are made of fused silica in order to reduce thermal noise and ULE compensation rings are foreseen in order to maintain the CTE zero crossing temperature [5]. The thermal enclosure consists of 4 gold coated aluminum cans with titanium alloy supports, as detailed in figure 5. Thermal simulations yield to an attenuation factor $> 10^{10}$, so that a 1 K temperature swing at the outer shield will have negligible stress effect on the cavity.

For frequency locking, standard Pound-Drever-Hall method cannot be applied as the cavity resonance can occur at a variable offset to the input light. This offset can be as large as half the FSR (full spectral range) of the cavity, in our case 750 MHz. Therefore sideband locking using an electro-optic modulator is foreseen [6].

IV. CONCLUSION

We presented the mission concept of a small satellite mission dedicated to perform a test of the boost dependence of the speed of light. By using state-of-the-art optical frequency references aboard a satellite flown in low-Earth orbit, the Kennedy-Thorndike coefficient can be determined with an up to two orders of magnitude higher accuracy than the currently best ground-based experiment. The payload consists of two frequency references, one based on Doppler-free spectroscopy of molecular iodine, the other based on a high-finesse optical cavity, with an intended frequency instability at or below the $10^{-15}$ level.

The experiment will be accommodated in a SaudiSt 4 satellite bus flying in a sun-synchronous orbit for enhanced science signal and thermal stability of the payload. In a currently performed study, the feasibility of the payload accommodation is investigated.
ACKNOWLEDGMENT

Financial support by the German Space Agency DLR with funds provided by the Federal Ministry of Economics and Technology (BMWi) under grant numbers 50 QT 1102, 50 QT 1201 and 50 QT 1401 is highly appreciated. The authors thank Jan Hrabina and Josef Lazar from the Institute of Scientific Instruments of the Academy of Sciences of the Czech Republic, Brno, for fruitful discussions on the design of the multipass iodine cell.

REFERENCES


Geometrical scale-factor stabilization of square cavity ring laser gyroscopes

J. Belfi and A. Di Virgilio
INFN Section of Pisa,
Largo Bruno Pontecorvo 3, Pisa, Italy
Email: belfi@pi.infn.it

N. Beverini, G. Carelli, E. Maccioni, A. Simonelli
Physics Dept., Univ. of Pisa,
Largo Bruno Pontecorvo 3, Pisa, Italy

R. Santagata
Physics Dept., Univ. of Siena,
Via Roma 56, Siena, Italy

Abstract—Large frame ring laser gyro performances are ultimately limited by the instabilities of their geometrical parameters. We present the experimental activity on the GP2 ring laser gyro. GP2 is a ring laser gyro devoted to develop advanced stabilization techniques of the ring cavity geometrical scale-factor. A method based on optical interferometry has been developed for canceling the deformations of the resonator. The method is based on the measurement and stabilization of the absolute length of the cavity perimeter and of the resonators formed by the opposite cavity mirrors. The optical frequency reference in the experiment is an iodine-stabilized He-Ne laser, with a relative frequency stability of $10^{-11}$. The measurement of the absolute length of the two resonators has been demonstrated up to now on a test bench. We discuss the experimental results on GP2: the present performances as a ring laser gyro and the stabilization scheme to be implemented in the near future.

I. INTRODUCTION

Ring lasers gyroscopes (RL) are inertial sensors able to measure absolute rotations [1]. If they are placed at rest in a ground-located laboratory, the measured rotation is that of our planet $\Omega_{\oplus}$. Ultra sensitive, large frame RLs are promising detectors for ground based tests of General Relativity (frame-dragging). Earth’s rotation frame-dragging or Lense-Thirring effect consists in a dragging of the local inertial frame of reference caused by the perturbation of the local metrics in the proximity of a spinning massive body like Earth [2]. On the Earth surface, relativistic effects are of the order of 1 part in $10^9$ of $\Omega_{\oplus}$. Frame dragging effects have been experimentally observed up to now only in space-experiments using orbiting satellites [3] [4].

The GINGER (Gyroscopes IN GEneral Relativity) experiment [5] [6] aims at measuring the Lense-Thirring effect for the first time in a ground-based laboratory, by using an array of large RLs made solid with the ground. The requirements can be clearly understood by examining the sensor response. The basic setup of a RL is made up of a stable ring optical cavity along which an active medium, typically a He-Ne mixture, is present (Fig. 1). Two laser beams are generated and propagate in opposite directions along the loop. By mixing on a photodiode the beams exiting the cavity in the opposite directions, a Sagnac beat frequency is measured [7]:

$$f_S = \frac{4A \cdot \Omega}{\lambda P},$$

where $\Omega = \Omega_{\oplus} + \Omega'$ is the rotation relative to the local Lorentz inertial frame (being $\Omega'$ any correction term), $A$ is the area vector enclosed by the ring optical path $P$ and $\lambda$ is the wavelength of the laser. The sensitivity limit of a RL is given by the shot-noise:

$$\Omega_{sn} = \frac{vP}{4AQ} \sqrt{\frac{hf}{P \cdot T}},$$

where $v$ is the velocity of the laser beam along the cavity, $Q$ is the quality factor of the resonator, $h$ the Planck constant, $P$ the detected optical power and $T$ the measuring time. From equation 1 two important features follow: the dependence of Sagnac frequency $f_S$ on the laser path geometry via the scale factor $k_S = 4A/\lambda P$, and the scalar nature of the sensor, being measured only the projection of the velocity vector $\vec{\Omega}$ along the area vector $\vec{A}$. The development of a high sensitivity RL requires:

- a large frame structure. To increase the size of the ring cavity, in fact, implies to increase the signal to noise ratio (SNR), being the signal proportional to the ratio $A/P$ via the scale factor, and the shot noise of the sensor proportional to $P$ via the quality factor $Q$.
- a high-Q resonator ($> 10^{11}$ and higher). Low-loss ‘five-9s quality’ super/mirrors must be used.
- a multi-axial system of RL, in order to reconstruct the modulus of the Earth rotational vector and compare the Earth rotation rate measured locally with the one provided by IERS (International Earth Rotation and Reference Systems Service).
- to reduce the instrumental drift in the measurement of rotation rate to less than $\Omega / \Omega_{\oplus}$. This needs a long-term strict control on the fluctuation of laser active medium, cavity geometry and, in a RLs array, of relative dihedral angles.
- to reduce all the sources of Earth-surface and environmental noise, installing the detector in a very stable geological environment, well coupled to the solid rock, in a low environmental noise laboratory, possibly located underground.

---

A. Simonelli is presently a PhD student of the Scuola di Dottorato Regionale in Scienze della Terra (XXX ciclo), Università di Pisa.
To this day, the best RL is the Grössring G, located at the Geodätisches Observatorium in Wettzell, Bavaria [8] [9]. It has achieved a resolution better than $5 \times 10^{-13}$ rad/s with an integration time of few hours, becoming of geodetic interest for measuring short-term fluctuation, with periods of hours to days, in Earth rotation. This stability record is mainly due to the strong passive stabilization of the optical cavity. G, in fact, is a semi-monolithic device made in Zerodur, a glass ceramic with an especially small thermal expansion, high mechanical stability and consistency of shape and length. Four bars are rigidly connected to a base plate forming the edges of a square stability and consistency of shape and length. Four bars are with an especially small thermal expansion, high mechanical is a semi-monolithic device made in Zerodur, a glass ceramic

The measure of Lense-Thirring effect in a ground-based laboratory requires a stabilization of the scale factor $k_S$ better than $10^{-10}$. This implies an accuracy on the mirror positioning better than 1 nm. It can be shown [13] that for a single square RL, the accuracy request on the mirror positions can be reduced if the absolute length of the diagonal cavities is stabilized in addition to the perimeter one. We showed that if the lengths of the two diagonals are locked to the same value, the perturbations to the mirror positions affect only quadratically the ring laser scale factor. This constraint reduces the mirror position fluctuation at a level of 1 part in $10^{10}$, even if the two lengths are stabilized to values that differs at a micrometric scale. These results motivated the design of GP2 [6], an intermediate prototype of GINGER specifically devoted to test the active control strategies and, in particular, to implement the length stabilization of the diagonal resonators by means of optical interferometry.

III. GP2 RING LASER

GP2 is the seed device for the next generation heterolithic active-stabilized RLs. The optical setup consists in four supermirrors each one contained in a steel holder placed at the corner of a granite support fixed on a concrete base. Steel pipes connected by the mirror holders define the vacuum chamber that encloses the optical path of the circulating beams along a square loop. High finesse of the two diagonal resonators is guaranteed by a special mirror coating ensuring a reflectivity of about 99.9% at normal incidence, in addition to a reflectivity $> 99.999\%$ at 45 deg angle of incidence.

The vacuum chamber, that includes also the crossed diagonals optical path, and is filled with He-Ne. The active medium for the ring laser operation is excited by means of a capacitive radiofrequency discharge, applied through a pyrex capillary located in the middle of one side of the ring.

Fig. 2 shows a drawing of GP2 (above) and its installation in the clean room at INFN Pisa laboratories. The granite slab supporting the laser cavity is oriented with its normal axis parallel to the Earth’s rotation axis in order minimize the orientation errors. The four mirrors holders (Fig. 5) are placed at the corner of a square granite slab and the vacuum chamber encloses the beam optical path along a square loop 1.60 m length in side. Fig. 3 shows a preliminary Sagnac spectrum; a Sagnac frequency of 184 Hz has been observed, as expected.
To check the quality factor $Q = 2\pi f \tau$ of the laser cavity we made a ring-down time $\tau$ measurement of the laser by short-cutting the discharge capacitor. In Fig. 4 we report the laser intensity decay trace acquired by the oscilloscope. Fitting to data the exponential function $I = I_0 e^{-t/\tau} + C$, where $I_0$ is the initial intensity and $C$ is a constant, we have obtained a measure for the ring-down time $\tau = 154.4 \pm 0.5 \mu s$. This corresponds to a quality cavity factor $Q = 4.6 \times 10^{11}$. The expected shot noise level is $\Omega_{\text{sn}} \sim 1(\text{mrad/s})/\sqrt{\text{Hz}}$.

The slab whereon the holders are mounted is made of precise black granite, a rock well suited for metrology application for his long term thermal and dimensional stability, high flatness accuracy, high bending strength and insensitivity to mechanical overloading. It has been machined with a precision better than 10 $\mu$m to guarantee a preliminary well positioning of the corner mirrors. The mirrors are accessible through big optical transparent windows installed parallel to them on the holders. The window allows the circulating beams to exit the cavity and to be monitored; an external optical setup detects the beat frequency. As showed in the lower part of Fig.5, it is possible to inject an external laser probe into the diagonal resonators.

In [13] we defined the eigenvectors basis of the cavity deformations, identified the rigid body motion of the cavity and then classified the residual 6 cavity deformation modes. These correspond to the minimum number of degrees of freedom to control for implementing a rigid constraint. GP2 has 6 piezoelectric transducers moving the mirrors holders: one 3-axial, and three 1-axial along the diagonals. Each piezo stage has a dynamic range of 80 $\mu$m with a driving voltage of about 180 V. In Fig. 6 is shown the measured frequency response of a 1-axial actuator. The response is flat up to about 80 Hz.
where a resonance appears. This sets the limit of the control bandwidth to about few tens of Hz.

A. Measurement principle

To stabilize the absolute length of a square RL diagonal resonators with respect to an interrogating high-stability laser we worked out an interferometric metrology technique and we tested it on two Fabry-Perot cavities simulating the ring diagonals on an optical bench. The technique we used is based on an accurate frequency measurement of the resonant longitudinal mode and an univocal determination of the interference order [14].

The laser source is a 10 mW diode laser emitting at 633 nm. A high spectral purity is gained referring it to an optical reference frequency provided by a 100 µW He-Ne laser frequency stabilized on the saturated absorption line R-127 11-5 of Iodine. The relative Allan deviation of the laser frequency has a minimum at 100 s at the level of $10^{-11}$.

This is achieved implementing a light amplifier based on injection-locking. Fig. 7 shows the optical scheme that will be used on GP2. The frequency stabilized laser is split in two parts and modulated by two electro-optic modulators driven by a modulation signal $M_{i=1,2}(t)$

$$M_i(t) = \alpha \sin \omega_A t + \beta \sin (\omega_B t + \Delta \sin \omega_C t)$$

where $\omega_A$ is the modulation providing the Pound-Drever-Hall signal for carrier lock; $\omega_B$, is the modulation allowing us the detection of the $i-th$ dynamical cavity resonance; $\omega_C$ is a frequency dithering applied for shifting the FSR detection in the kHz frequency range.

The electronic apparatus for the GP2 diagonals length stabilization is shown in Fig. 8. It is based on the multi-frequency digital lock-in amplifier (Zürich Instruments UHF2). It allows us to digitally synthesize the phase modulations and to process the two beams intensities $I_{1r}$ and $I_{2r}$, reflected from the two cavities. For each resonator, two phase error signals $\phi_A$ and $\phi_C$ at the modulation frequencies $\omega_A$ and $\omega_C$ are calculated and two digital PIDs are generated. Four PIDs are implemented. Two PIDs drive the cavities PZTs, and the other two PIDs drive the VCOs synthesizing the modulations at $\omega_B^1$ and $\omega_B^2$ (see details in the figure caption). Once the four loops are closed, cavities optical frequency resonances are locked to the reference laser. The cavity dynamic resonances (in the range of few hundreds MHz) are measured with two RF frequency counters.

IV. Conclusion

We described the heterolithic, active-stabilized GP2 gyroscope. The expedients adopted to achieve a stabilization of the scale factor better than $10^{-10}$ are reported. An optical interferometry method already verified in [14] is going to be implemented by means of digital controls. The expected outcome provided by the implementation of the control scheme described in this paper is to make GP2 run as a gyroscope with a geometrical stability only limited by the stability of the laser frequency reference. The achievement of these results will be
of key importance toward an heterolithic large frame detector for fundamental physics research.

REFERENCES


Möbius Metamaterial Topology: Applications in
Resonators and Tunable Oscillator Circuits

1,2 Ajay K. Poddar
1 Synergy Microwave Corp., NJ, USA
2 University of Oradea, Romania

Abstract—Manipulating and tailoring the electromagnetic-wave coupling properties, numerous interesting properties of topology driven Möbius Metamaterials Strip is emerging that goes beyond the conventional characteristics and structure functionality relationship. The goal of this paper is 3-folds: (i) to address the key concept of topology and topological notions in a wide class of materials with potential technological applications, (ii) discuss the possibility of application of topology driven negative refractive index composite Metamaterial Möbius strips, and (iii) validation examples (4GHz-12GHz/12GHz-18GHz voltage controlled oscillator circuits for RF/MW applications).

Keywords— Metamaterial, Möbius Strips, Oscillator, Topology

I. INTRODUCTION

Topology is a field of science that studies invariance of certain properties under continuous deformation, such as stretching, bending, or twisting, of the underlying geometry. Topological symmetry defined as a property, conserved when the system undergoes an alteration (deformation, twisting and stretching of objects), for example, Möbius strips deform in a way that its metrical properties barely change, and violate the Hückel rules [1]. It is interesting to note that the Möbius strips allow launch pad to complex geometries and topological exploration for realization of -ve index material and building components for electronics and biomedical applications.

Recently groundbreaking experimental results reported [2]-[10] based on topology symmetry, which has received widespread attentions in the field of DNA (Deoxyribonucleic acid) optical polarization, metamolecules, and high frequency components for applications in communication systems.

A. Möbius DNA Strip: Sensors Appications

Figure (1) shows the nano-architectures of Möbius DNA strip in which each colored band represents a different DNA double helix; such nano-architectures used in high sensitivity biological and chemical sensing devices [2].

Fig.1 The typical nano-architectures of Möbius DNA strip: (a) colored Möbius strip, (b) A fraction of the Möbius strip in (a) is illustrated in the DNA helical mode, and (c) Image of Möbius DNA strip [2].

B. Möbius Polarized Light: Optical Fabrication Application

Figure (2) illustrates the artificial Möbius Strip formed by EM (electromagnetic wave) waves, which demonstrates “a light-beam can be controlled so that its polarization twists follow a contour of Möbius-Strips” [3]. As shown in Figure (2), the typical Möbius strip surface made of polarization states of light beam, which is a non-vanishing curvature, exhibits special symmetry. The creation of EM (electromagnetic wave) around Möbius contour is remarkable for understanding of the optical polarization and the complex light beam engineering extends the capability of developing optical micro-and nano-fabrication structures for sub-wavelength imaging.

C. Benzene Metamolecules: Metamaterial Möbius Symmetry

The EM symmetry discovered in the metamaterial is equivalent to the structural symmetry of a Möbius-Strip, with the number of twists controlled by the sign change of the electromagnetic coupling between the meta-atoms. Figure (3) illustrates the Metamaterial Möbius symmetry made from metals and dielectrics [4].

Fig.2. Optical polarization twist around Möbius-Strip: (a) single knot, (b) multi knots, and (c) observed Möbius Strip polarization twist [3]

Fig.3: Möbius-Strips symmetry in metamolecular trimers [4]
As depicted in Figure (3), the hyperspace Möbius mechanism transforms ordinary benzene molecules into metamolecules with Möbius symmetry: “the topological phenomenon that yields a half-twisted strip with two surfaces but only one side”. The prototype of metamolecular trimer shown in Figure (3) is a 3-body system like a trimmer, metallic resonant meta-atoms configured as coupled split-ring resonators, symbolized as metamolecules exhibits topological Möbius-Cyclic-Symmetry (C$_3$-symmetry) through 3-rotations of 120-degrees, and Möbius-twists result from a change in the signs of the electromagnetic coupling constants between the constituent meta-atoms [4]. The interesting phenomena is “with different coupling signs exhibit resonance frequencies that depend only on the number of turns but not the locations of the twists,” thus confirming its Möbius Symmetry.

**D. Graphene Möbius Strip: Microwave Components**

Recently, Graphene has received extensive attention due to its remarkable structural (single layer of graphite, exhibits mechanical properties like thin paper or plastic with large bulk modulus along the plane of Graphene, easily bent or curved) and superior electronic and optical properties [5]. This unique characteristic allows Graphene wrap into carbon nanotubes without deformation, qualifies to use Graphene to build Möbius-Strips for microwave and optical components.

The bandgap and cohesive energy of Mobius Graphene as depicted in Figure (4) depends on the width of strip, and augment by increasing the width, described by [6]

$$E_{CO}^{(M)} = \frac{E_{SCF}^{(M)} - N_{c}E_{SCF}^{c} - N_{H}E_{SCF}^{H}}{2(N_{c} + N_{H})}$$

where the superscript $r(M)$ corresponds to the Graphene ribbon (Mobius strip), $N_{c}$ is the number of carbon atoms, $N_{H}$ is the number of hydrogen atoms, $E_{CO}^{(M)}$ is the self-consistent field energy of the Graphene ribbon (Mobius strip), $E_{SCF}^{c}$ and $E_{SCF}^{H}$ are the self-consistent field energy of alone carbon and hydrogen atoms, respectively.

Secondly, the magnetic moment and spin-properties of Möbius Graphene is very interesting. It is interesting to note that Graphene Möbius strip keeps its metallic surface states in the presence of an external electric field. For sufficiently higher applied electric field, the spin flipping can take place in the Mobius strip. In contrast with the Graphene nanoribbons, Graphene Mobius strips show half-semiconducting properties when an external electric field is applied.

Figure (5) shows the typical orbital of Graphene Möbius strip and spin dependent DOS (density of states). As shown in Figure (5), ferromagnetism and spin-flipping properties of Möbius Graphene makes attractive applications for spintronic devices and quantum oscillator applications [6].

![Fig.4: Graphene Möbius structure: Möbius strips with length L=18 and width N=3 (Möbius axis and edge C atom index are shown) [6]](image)

![Fig.5: (a) depicts spin-dependent DOS (density of states) of Graphene Möbius strip as a function of the electron energy, and (b) molecular orbital of Graphene Möbius strip at E=0eV [6]](image)

![Fig.6: (a) Möbius molecular rings made of carbon atoms, (b) Energy spectra $E_{\sigma}$ of the molecular ring, (c) Relative permittivity as a function of $\Delta \omega$, and (d) Relative permeability as a function of $\Delta \omega$ [7]](image)

Figure (6) shows the characteristics of Möbius molecular rings that support negative refraction (-ve permittivity $\varepsilon$ and -ve permeability $\mu$), known as metamaterial properties. As shown in Figure (6), two energy bands are denoted by their different pseudo spin labels $\sigma = \uparrow$ and $\downarrow$; detuning $\Delta \omega = \omega - \omega_0$ [7]. The difference between the Möbius molecular ring and the common annulenes lies in the boundary condition. The negative index properties offer remarkable properties such as image cloaking, sub-wavelength imaging, enhancement of evanescent fields resulting improved Q-factor. Figure (7) depicts the application of Graphene Möbius strips for the realization of microelectronic components.

![Fig.7: The structures of the Möbius Metamaterial strips formed by Graphene for the realization of microelectronic components [5]](image)
As shown in Figure (7), cylindrical closed ring strips structures formed by Graphene-nanoribbons possess 2-edges, exhibit anti-ferromagnetic (Zero-magnetic moment, whereas Mobius closed ring strips structure formed by Graphene-nanoribbons possess single-edge, exhibit ferromagnetism (nonzero magnetic moment). Graphene nanoribbons with ‘zigzag’ edge structure exhibits magnetism at their edges, most stable configuration of these ribbons is anti-ferromagnetic, so that magnetic moments at opposite edges point in opposite directions, reducing the total magnetic moment of the ribbons to zero. However, Graphene Möbius strip is only one continuous edge, hence no magnetic cancellation between the opposite edges, resulting nonzero magnetic moment. Applications of these topological materials (Graphene Metamaterial Möbius strips) illustrated in Figure (7) ranges from sensing, energy storing elements- resonators, catalysis to DNA structure, and nanomedicine.

As discussed above, there are numerous topological applications envisage by virtue of Möbius transformation and Möbius metamaterial coupling mechanism, this paper is focused on design of high performance oscillator circuits using Möbius metamaterial strips resonator.

II. MöBIUS TRANSFORMATION AND MOBIUS STRIPS

A. MöBIUS Transformation

Figure (1) shows the nano-architectures of Möbius DNA

The topological Möbius metamaterial symmetry is due to Möbius transformation characterizes $f(z)$, is given by [8]

$$f(z) = \frac{az+b}{cz+d}, \ \ (a, b, c, d \in \mathbb{C} \text{ and } ad - bc \neq 0) \ (1)$$

$$z \mapsto \rho \mapsto r [\cos(\theta) + is \sin(\theta)] \ (2)$$

where $a, b, c, d$ are complex numbers, and the numerator of (1) is not a multiple of denominator (i.e. $ad - bc \neq 0$). From (1), the properties of Möbius transformation $f(z)$ are

(i) $f(z)$ can be expressed as a composition of affine transformations (scaling: $z \rightarrow \rho z$, translation:

$$z \mapsto \rho \mapsto r$$

rotation: $z \rightarrow e^{i\theta}z$, complex conjugation:

$$z \rightarrow \bar{z}$$

where $t, p \in \mathbb{C}$

(ii) $f(z)$ maps $\mathbb{C}$ one-to-one onto itself, and is continuous

(iii) $f(z)$ maps circles and lines to circles and lines

(iv) $f(z)$ is conformal

From (i)-(iv), the family of functions is the composition of functions; the identity element is the identity map, and the inverse is given by inverse function. The Möbius group consists of those fractional linear transformations, which maps the open unit disk $\mathbb{D}$ = \{ $z \in \mathbb{C}$: $|z| < 1$ \} to itself in a one-to-one way. These transformations and their inverses are analytic on $\mathbb{D}$ and map its boundary (the unit circle $S^1$ =\{ $z \in \mathbb{C}$: $|z| = 1$ \}) to itself. The automorphisms of the disk is in the form:

$$f(z) = e^{i\phi} \frac{z-a}{1-az} \ (3)$$

$$z = M(w) = e^{i\psi} \frac{w+a}{1+aw} \ \ (4)$$

$$w = M^{-1}(z) = e^{-i\psi} \frac{z-a}{1-az} + a \ \ (5)$$

where $\phi \in \mathbb{R}$, $\psi \in \mathbb{R}$, and $a \in \mathbb{D}$.

From (1)-(5), Möbius transformation used for developing different topological symmetry Möbius strips resonators for the host of frequency generating circuits [9].

B. Möbius Strips

The concept of the Mobius strips based on the fact: “a signal coupled to a strip shall not encounter any obstruction when travelling around the loop”, enables large group-delay and high quality factor. The particular interest in this paper is planar closed ring resonator for monolithic integration, equivalently represented by the simple lumped L-C network, shown in Figure (8a). In solving, the electric currents on the resonator can be formulated by a periodic boundary condition of the form described by

$$I_{j+k} = I_j \ (6)$$

where $I_j$ represents the electric current around the $k^{th}$ closed loop on the periodic ladder structure of $k$-elements. The boundary condition of the general form shown in (6) governs that $I_k$ are a conserved quantity that gives invariance of solutions under a $2\pi$ rotation with a definite handedness. The non-dissipative wave equation of LC network shown in Figure (8a) for the $k^{th}$ element:

$$\frac{\omega^2 - \frac{1}{LC}}{I_k} - \gamma \frac{\omega^2 - \frac{1}{2LC}}{(I_{k+1} + I_{k-1})} = 0 \ (8)$$

$$\omega^2 = \frac{\gamma}{LC} \left( \frac{2\sin^2 \frac{\pi k}{N}}{N^2 - 2\pi \cos \frac{2\pi k}{N}} \right) \ (9)$$

where $p$ is an integer specifying the normal mode, $\gamma$ is mutual coupling coefficient (Mutual inductance ‘M’=2$\gamma$L), $k$ is number of element structure. From (7)-(9), for even value of $k$, there are $k-1$ eigenvalues, including $(k-2)/2$ degenerate doublet and one singlet. A typical ring resonator, whose Eigen functions satisfy (6), defines a distinct inner and outer surface of the ring, shown in Fig. (8b). Figure (8c) shows a topological transformation resulting in a Mobius strip resonator, whose current dynamics formulated by applying twisted boundary condition as

$$I_{j+k} = -I_j \ (10)$$

From (10), a simple topological transformation on the resonator ring (Fig. 8b) results in a sign reversal of current ($I_j$) upon a $2\pi$ rotation of the solutions, and a $4\pi$ rotation is now required for invariance of the Eigen functions. Note that the eigenfunctions satisfying the condition for twisted boundary are of the same form as (6) provided that the mode indices are given half-integer values ($p = 1/2, 3/2, 5/2, \ldots (k-1)/2$) relative to a ring consisting of identical components.

![Fig. 8](image-url)
The dispersion relation for Mobius ring is same as (9), however, the wave-vectors are shifted by
\[ \Delta \lambda = -\frac{a}{2} \] (11)

The two distinct topologies shown in Figures 8(b) and, 8(c) can be considered as a complementary pair related by a single transformation. The eigenfunctions of the Mobius resonator form an orthogonal basis set; presents an interesting possibility for the design of metamaterial for the application in tunable oscillators, antenna, and filter circuits. The inherent multi-phase, multi-mode nature of the Möbius strip provides additional degrees of freedom relative to traditional designs for generating multi-phase and quadrature oscillators, which are essential modules of many electronic systems, such as image rejection demodulator in wireless transreceivers, half-rate clock-data-recovery (CDR) circuitry in high-speed optical receivers, phased-arrays, direct-conversion transmitter, and fractional-N frequency synthesizers.

C. Multi-Phase Oscillators

Figure (9) shows the typical ring oscillator, exhibits multi-phase but it suffers from high phase noise [10]. LC-boosted ring oscillator as shown in Fig. 9(c) minimizes the noise but still inferior to single-stage LC oscillator, largely due to the trade-off between phase noise and phase error. Strong coupling between the gain stages (implies large-coupling MOSFETs) can minimize the phase errors; on the other hand, large coupling MOSFETs will increase noise and power consumption [10]. In all possible multi-phase LC oscillator configurations, MOSFETs play a dual role: compensate the energy loss in LC tanks, and couple LC tanks together and maintain a desired phase relation [11]. However, the second goal deteriorates phase noise, especially 1/f (flicker) noise up-conversion [12]. These problems can be partially or fully overcome by replacing MOSFETs with metamaterial Möbius connected high order LC ring resonator in distributed topology, shown in Figure (10). The resonant frequency of metamaterial (-ve index material) LC-ring resonator increases with the number of stages \( N \) as compared to right-handed (+ve index material) LC-ring resonator. Therefore, metamaterial LC-ring resonator is well suited for high frequency generation.

As shown in Figure (10), MOSFETs acts as negative resistance generating devices, needed only for compensating the loss of the LC-ring resonator, implies that current from MOSFETs are always injected into the resonator in phase with voltage and do not resonate with LC resonator components. For N-stage distributed LC-ring oscillator, the phase noise scales down as 1/N, while the power consumption increases linearly with N. This allows synthesizing multi-phases, while maintaining the identical figure of merit (FoM) as a 1-stage LC oscillator [12]. The main challenges of these topologies are mode selections and desired phase relations. In addition to this, assumption for number of noise sources increases linearly with \( N \), the total phase noise should be proportional to 1/N no longer hold good under large signal drive level condition.

The alternative approach is wave-based Möbius oscillator. The energy recycling nature of Möbius strips allows very high frequencies oscillations with minimum power consumption for a given class of frequency generating circuits.

D. Wave-Based Möbius Oscillator

Figures 11(a) and 11(b) show the differential open ring and closed cross-connected ring. When the voltage source connected in Figure 11(a) replaced by a cross-connection of the inner and outer conductors, leads to a signal inversion. If there were no losses, a wave could travel on this differential ring indefinitely, providing a full clock cycle every other rotation of ring, defined as Möbius effect [13]. However, in real applications, multiple anti-parallel inverter pairs added to the line shown in Figure 12(a) to overcome losses and provide rotation lock of Möbius connected differential ring.

Figure (12) shows the lumped equivalent of coupled stripline used for Möbius ring described in Figures (11)[14].

Fig. 11: shows the oscillation based on the Möbius effect: (a) open loop differential ring, (b) Möbius connected differential ring, and (c) practical tunable Möbius connected differential oscillator circuit [15]

As shown in Figure (12), \( Z_e \) are even and odd mode impedance, \( \gamma_e, \gamma_o \) are even and odd mode propagation constant, \( L_e, L_o \) are even and odd mode length [14].
As shown in Figure (12), coupled line would support even and odd mode propagation. The even mode (fast wave mode) when the two lines are excited with the same signal, as a result, the coupling capacitance $C_r$ between the two lines (Fig.12), will have no effect on the propagation constant '$\gamma$'. The odd mode (slow wave mode), is obtained as the line pair is excited differentially, so the coupling capacitance between the two lines has a doubled effect on the propagation constant of this mode. For resonator application, slow wave propagation is of interest, care must be taken to suppress the even mode propagation. Figure (13) shows the method of stable oscillation, inverters across the lines are providing gain and theoretically forcing the line to operate differentially, varactor diodes and centered ring provides even mode cancellation [13]-[15]. As shown in Figure (13) Möbius effect allows energy harvesting effects, energy is not lost at each stage but re-circulates in closed path).

While the benefits of the Möbius connected wave-based oscillators are evident, this topology still limits the tuning range, limited to 10-20%. In this paper, a novel approach of improving the tuning range is suggested and validated with the practical example of 4-12 GHz/12-18 GHz oscillator based on Metamaterial-Möbius-Strips resonator. From (26), locking range restricted by the phase shift:

$$\Delta \phi = \frac{\phi_0}{2Q} \tan \left( \sin^{-1} \left( \frac{\eta}{2} \right) \right)$$

where $\phi_0$, $Q$, and $\eta$ are the resonant frequency, quality factor of the LC tank, and the injection ratio $(i_{inj}/I)$ respectively. Using (12)-(16), 18-25 GHz Metamaterial Möbius-Strips resonator based oscillator designed and validated.

III. METAMATERIAL MöBIUS STRIPS RESONATOR VCO

Recent publications [10]-[15] describe the oscillator circuits using wave-based oscillators but they have limited tuning range, prone to mode-jumping, poor quality factor due to dispersion phenomena at microwave frequencies, therefore inferior phase noise performances. These drawbacks can be overcome by using Metamaterial based Möbius strip resonator in conjunction with phase-injection-mode-locking mechanism. The unique characteristic of Möbius strip is self-phase-injection properties along the strips that enables higher $Q$ factor for a given size of the printed resonator.

The phase-injections (perturbations) at different nodal points on strips improve the tuning-range and phase noise performance of Möbius strip based frequency-generating circuits. Figure (14) illustrates the typical phase-injection mechanism using phase-perturbation $\psi_i$ at $i^{th}$ node on Möbius-Strip resonator based frequency generating electronic circuits. The $Q$-factor of resonator is

$$Q_L = \frac{\phi_0}{2} \frac{d\phi(\omega)}{d\omega} = \frac{\phi_0}{2} r_d \; ; \; r_d = \left| \frac{d\phi(\omega)}{d\omega} \right|_{\omega = \omega_0}$$

$$r_d = \frac{d\phi(\omega)}{d\omega} = \frac{\phi_0}{2} \sin 2\tan 2$$

where $\phi(\omega)$ is the phase of the resonator impedance and $\tau_d$ is the group delay. The effective inductance of the Möbius strips varies with the current/voltage passes through strips, building evanescent field, therefore higher quality factor:

$$Q_m(\omega, \omega_0) = \left[ \frac{\omega}{2(\max - \min)} \int_{\min}^{\max} Q_m(\omega, i, i) \; di \right]$$

From (26), locking condition given by [9]

$$2(f - f_0) = \frac{\eta \sin(\phi)}{2 + \cos(\phi)}$$

where $f_0$, $Q$, and $\eta$ are the resonant frequency, quality factor of the LC tank, and the injection ratio $(i_{inj}/I)$ respectively. Using (12)-(16), 18-25 GHz Metamaterial Möbius-Strips resonator based oscillator designed and validated. Figure (15) shows the PCB layout of oscillator circuit, fabricated on 8-mil thick Roger substrate, with dielectric constant value of 2.2. As shown in Figure (15), the Metamaterial Möbius-Strips resonator acts as an evanescent mode cavity, restoring the recycled energy due to Möbius effect under a $2\pi$ rotation with a definite handedness. The resonator coupling coefficient ‘$\beta_j$’ depends upon the geometry of the perturbation:

$$\beta_j = \left[ \left( \frac{\mu H_e E_m dv}{\mu H_e E_m dv} \right)_e + \left( \frac{\mu H_m E_e dv}{\mu H_m E_e dv} \right)_m \right]$$

where $E_m$ and $H_e$ are the electric and magnetic fields produced by the Möbius strip, and $E_m$, $H_m$ are corresponding fields due to perturbation or nearby adjacent resonator, subscript ‘$e$’ and ‘$m$’ are the electrical and magnetic coupling.
Fig. 15: Layout of 12-18 GHz oscillator (0.75x0.75inches), using evanescent mode Metamaterial Möbius strips resonator network realized using 8 mils RT/Duroid 5880 with $\varepsilon_r = 2.2$

The oscillator circuit shown in Figure (15) works at DC bias of 5V and 35 mA, measured output power is typically 0dBm over the operating frequency band.

The oscillator circuit shown in Figure (15) works at DC bias of 5V and 35 mA, measured output power is typically 0dBm over the operating frequency band. The concept and basis of Metamaterial Möbius strips based oscillator reported in Figure (15), can be utilized to make tunable multi-band oscillator circuits in compact size. Figure (17) shows the typical circuit schematic of multi-band tunable oscillator 4-12 GHz, uses Infineon BFP 740 SiGeHBT, works at DC bias of 3.5Volt and 25 mA, measured output power exceeds -3 dBm over the desired tuning range.

IV. CONCLUSION

With the development of MMIC fabrication techniques, Metamaterial Möbius strips resonator based oscillator is a promising alternative to high frequency planar VCO solutions.

REFERENCES

Frequency Signal Source’s PN (Phase Noise) Measurements: Challenges and Uncertainty

Ulrich L. Rohde  
BTU Cottbus 03046 Germany

Ajay K. Poddar  
Synergy Microwave NJ USA

Enrico Rubiola  
FEMTO-ST Inst. Besancon France

Marius A. Silaghi  
University of Oradea, Romania

Abstract—This paper describes oscillator noise measurement techniques, challenges and associated measurement uncertainty. The cross-correlation method used in modern PN measurement equipments, can present erroneous result, depending upon phase-inversion, harmonics, o/p load mismatch, and cable length. This discussion is imperative for low phase noise signal sources, validated with 2.4 GHz SAW oscillator, and discussed steps for mitigating these issues by using filtering/phase-matching N/W.

Keywords—Cross-Correlation, Oscillator, Phase Noise

I. INTRODUCTION

Prediction and estimation of oscillator phase noise is highly desirable for accurate timing, and other measuring purpose. Several noise models developed for estimating the phase noise dynamics of oscillator systems [1]-[6].

A. Oscillator Phase Noise Model

The simplest noise model is linear time invariant (LTIV) system in which oscillator’s output considered as a resonant response to the input noise [1]-[2]. However, LTIV model fails to explain the uncertainty in the phase of noise upon each addition of noise to the autonomous oscillatory field [7]. Consequently, relative phase between the signal and the added noise results in partition of the noise perturbation into quadrature of amplitude and phase, forming the basis for linear time variant (LTV) model [2]-[3]. The LTV model addresses the issues of LTIV approach by linearizing the oscillator phase, however introduces nonphysical artifact of infinite spectral power at frequencies approaching the carrier signal. The scientific alternative is nonlinear time variant (NLTV) model. NLTV techniques uses perturbation method based on numerical techniques, involves rigorous stochastic differential equations [4]-[6]. The NLTV approach applies Floquet theory to demonstrate the oscillator spectrum to be a Lorentzian for white-noise perturbation. The Lorentzian distribution [4] institutes the power at the oscillation frequency to be finite, mitigate the problems of infinite power [1, 3]. Although, the NLTV approach is precise, consequential solution is non-intuitive and not easy to apply since it requires comprehensive analysis of the oscillator dynamics, e.g., the determination of the time consuming Floquet eigenvectors [7].

Table 1 shows the comparative analysis of 3-most cited phase noise models (LTIV, LTV, and NLTV) [1]-[6]. One can argue the superiority of any of the three models based on easy to use, accuracy, reliability, simulation time, and convergence for a given oscillator topology. The ultimate goal of unified noise model is still a subject of discussion, what is needed to have an intuitive functional description of oscillator phase noise, comparable to the empirical model proposed by Leeson [1], but incorporates physics based technique [8].

Similarly, the reliable phase noise (PN) measurement is a difficult and time-consuming task.

B. Phase Noise Measurement Challenges And Uncertainty

The PN measurement is ratiometric measurement, which requires high dynamic range to perform reliable measurement. The most effective methods therefore rely on removing carrier (by filtering or phase/frequency detection), so removing the reference value for any measurements. This may lead to “stitching” errors in the phase noise plots, which can be noticed when the PN measurement equipment changes settings to measure noise at different offset frequencies.

It is common to witness a test system produce a phase noise plot where the phase noise suddenly changes in level by a few dB’s. It is typically a sign that the calibration of the plot is incorrect and the measured data cannot be relied. The problem frequently arises because the equipment changes bandwidth and other critical setting (amplifier or PLL setting) that has an impact on the correction values applied to the measurement. Sometimes it can indicate that part of the test system is overloaded during one set of measurements. These errors make it tricky to assign a traceability figure to phase noise measurements.

The software algorithm in the PN measurement equipment may attempt to categorize noise and spurious signals, and correct the reading automatically in real time. Unfortunately, software do not succeed all time, particularly when the signal is slightly above the surrounding noise levels. The software algorithm can be made efficient by narrowing measurement bandwidth (at the expense of measurement time), because this lowers the level of the measured noise compared to the spurious signal. This makes it easier for the software to distinguish between noise and a spurious signal for a given spurious signal level; this only moves the problem to a lower level but never eradicate the problems.

Table 1: Describes the strengths and weaknesses of the noise models
Besides the concern of identifying spurious signals and noise, instruments use subtle method to expand the displayed dynamic range or to make noise signals appear to be practical and reasonable. These artifacts can be problematic, may lead to error in interpreting the real signal levels measured. The manufacturer of commercially PN measurement instruments continuously putting effort to improve the measurement method, and consequently may be able to take measures that partially overcome but in most cases disguise the limitations.

### C. Influence of Harmonics on Phase Noise Measurement

For low phase noise measurement, we noticed that third order harmonics and equipment dynamic ranges influence the measurement with greater degree. Device under test (DUT) harmonics can influence significantly on PN measurement accuracy. For validation, we postulate that the DUT’s harmonics are combined by the mixers and contribute to phase noise measurement variations. The mixers convert to baseband the fundamental and the harmonics of the signal generated by the DUT. Because the noise components of the harmonics and the fundamental are correlated their combination depends on the phase delay between the harmonics.

Figure 1(a) shows the spectral components of the DUT signal. H1 and H3 are the fundamental and the 3rd harmonic, n1,fo and n3,fo are the phase noise (PN) associated with H1 and H3 at an offset frequency (fo). The 3rd harmonic noise is correlated to the noise of the fundamental and its level is 9.542dBc higher. For ease, amplitudes of the fundamental and of the 3rd harmonic, H1 and H3, are not represented to scale.

In Figure 1(b), c1,fo and c3,fo represent PN associated with H1 and H3 that were down-converted. The PN associated with the 3rd harmonic and fundamental will be added as two rotating vectors. The magnitude of this component depends on the relative phase between the 3rd harmonic and fundamental. The resulting PN measurement uncertainty (peak-to-peak) at any given offset frequency is:

\[
\frac{c_{fo}^{\text{max}}}{c_{fo}^{\text{min}}} [dB] = 20 \log \left( \frac{1+3N1,N3}{1-3N1,N3} \right) \quad (1)
\]

The phase of the 3rd harmonic is rotated over 360 degrees producing a maximum and a minimum of the c3,0 signal.

Figure (2) illustrates the set-up, Signal Source Analyzer (SSA) used to measure phase noise of a signal, rich in harmonics. We use Rhode & Schwarz SPREF Reference Synthesizer, which can generate a 1GHz signal with 13dBm power, -35dBc harmonic content and a phase noise of -149 dBc at 10 kHz offset. The SPREF harmonically produces 2, 3, 4, 5, and 6GHz and these frequencies are available simultaneously on different outputs, all internally locked to a common crystal.

For the first part of the experiment, we selected the 1GHz and 3GHz outputs. The 3GHz output is passed through a variable attenuator and through a coaxial phase shifter (Narda’s model 3752), then is added to the 1GHz reference by means of a 6dBc resistive divider. The 3GHz signal is produced harmonically from the 1GHz signal and thus their phases are coherent. The cables, variable attenuator and power combiner determine the relative position of the 3GHz with respect to 1GHz. With the help of phase shifter, the phase of 3GHz signal’s can be rotated 360 degrees to perform the measurements. In Fig. 3 we present the results captured while using 6dB, 12dB and 19dB attenuator and rotating the phase of the 3GHz signal in 30 degrees increments from 0 to 360 degrees. We take the phase noise plot at 10 kHz offset and show the variation trend based on the phase shift delay. We plot the phase shift in radians while emphasizing that the origin is arbitrarily set by the cable length. The PN measured with the 3GHz signal not connected was -149.1dBc. With the 6dB attenuator, the amplitude of the signal at the phase noise analyzer is +6.2dBm for the 1GHz and -0.7dBm for the 3GHz. The corresponding 3rd harmonic levels, considering 1dB loss through the coaxial phase shifter, are -7, -13 and -20dBc. As we can see in Fig. 3, a 3rd harmonic with a level of -7dBc is showing a variation between -8 and +4dB around the level measured without harmonics. We can also observe that the values at 6.28 radians are approximately equal with the values at 0 radians as expected from a shift over 360 degrees.

**Fig. 2** Instrument setup for generating signals with harmonics and measuring the phase noise

**Fig. 3** PN depending on phase shift delay and harmonic level
With a harmonic level of -20dBc exhibits a variation of +1.1/-1.5dB. This suggests that using a filter that reduces the level of harmonics to below -20dBc would reduce the uncertainty to +/-1dB. We verified the 2^{nd}, 4^{th} and 5^{th} harmonic conversion using the 2, 4 and 5GHz outputs from the SPREF. In summary, a -7dBc 2^{nd} harmonic has shown a variation about +/-2dB, while the -13dBc results were in the range of +/-1dB. The 4^{th} and 5^{th} harmonics were showing no discernible variation of uncertainty of measurements over all phases. Figure (4) shows the relationship between phase noise and the conversion loss difference between the fundamental and third harmonic, for various levels of 3^{rd} harmonic content in the signal presented at the input of the test equipment. The phase noise variations shown in Figure (4) are observed when fundamental is presented with high harmonic content at the input to the test equipment. These mixers (double balanced, image reject or triple balanced mixers) suppress the even order and as such show no significant variations in the case of even harmonics, while the 5^{th} harmonic conversion loss is high enough and does not create measurement issues.

Analyzing Figure (4) if the mixer has a conversion loss difference between the fundamental and the 3^{rd} harmonic of 7.5dB and the 3^{rd} harmonic level is -7dBc (Red trace) we can expect a phase noise measurement variation of about 12dB (Red dot in Fig. 4, corresponding to the red trace in Fig. 3). For the same signal, if the conversion loss difference would be greater than 23dB, we could achieve a measurement variation smaller than +/-1dB. However, if the signal has a 3^{rd} harmonic with a level of -20dBc (Green trace), and the conversion loss difference (between fundamental and 3^{rd}) is higher than 10dB then the phase noise uncertainty will be less than +/- 1 dB.

All phase noise measurement methods that use mixers to down-convert the signal to baseband are subject to the effects described in Figure (4). The Phase Detector method and the Residual Phase Noise method will also demonstrate similar behavior. If the mixed signals have harmonics, the mechanism that converts the harmonics to baseband will degrade the measurement accuracy.

D. Estimation of Uncertainty

For reliable measurement, estimation of uncertainty is key requirement. As per guidelines [9], measurement uncertainty grouped into two categories: type-A (statistical), and type-B (various components-instruments and temperature control).

![Phase noise uncertainty estimate peak-to-peak (1) for different levels of 3^{rd} harmonic with respect to fundamental](image)

The type “A” characterized by estimated variances, accounts for statistical methods such as reproducibility, repeatability, special consideration about Fast Fourier Transform analysis, and the experimental standard deviation. The type “B”, accounts for uncertainty due to instruments and temperature control variable; characterized by quantities that may be considered as approximations to the corresponding variances, the existence of which is assumed [10].

There are several methods to estimate the uncertainty and measure the PN of signal sources. However, it is essential to know the weaknesses-strengths of different measurement techniques, since none of these methods are correct for every situation. The modern cross-correlation engine takes into the account of most part of the uncertainty associated with the PN measurement instruments. Many test and measurement manufacturers implement cross-correlation techniques in one form or another to produce best phase noise measurement instruments available in the market today. However, cross-correlation engines is not flawless, simultaneous presence of correlated and anti-correlated signals can lead to gross underestimation of the total signal in cross-spectral analysis at localized offset or over a wide range of frequencies [11]-[12].

II. OSCILLATOR PHASE NOISE MEASUREMENT

The oscillator noise is normally described in terms of the power spectrum density \( S(f) \) of the amplitude and phase noise, thus \( S_a(f) \) and \( S_p(f) \), as a function of the Fourier frequency \( f \). Designers should be aware that the effect of amplitude noise may not be negligible, and that the resonant frequency of some resonators may be affected by the amplitude. The first definition of phase noise \( L(f) \) was given as \( L(f) = \frac{\text{SSB power}}{\text{carrier power}} \). The problem with this definition is that it does not divide AM noise from PM noise, which yields to ambiguous results [11]. The IEEE Std 1139-1999 [13] redefines \( L(f) \) as \( L(f) = (1/2) \times S_p(f) \).

The physical unit of \( S_p(f) \) is \( \text{rad}^2/\text{Hz} \). Phase noise spectra are plotted on a log-log scale. \( L(f) \) is always given in dBc/Hz, which stands for dB below the carrier in 1-Hz bandwidth. In decibels, \( L(f) = S_p(f) - 3\text{dB} \). The fact that phase noise expressed in terms of the relative power (to the total carrier power) in a 1 Hz BW does not mean that the signal measured in a 1 Hz bandwidth. The phase noise data displayed by the equipment normalized to 1Hz measurement bandwidth. The modern phase noise measurement equipment measure the phase noise in measurement bandwidths, which increases with increasing carrier offset frequency. This is done for two reasons: (1) it results in shorter overall measurement time, and (2) at high carrier offset frequency (> 100 kHz), many measurement systems employ analog spectrum analyzers that are not capable of 1Hz resolution.

Noise measured in a 1 kHz bandwidth, for example, is 30dB higher than that displayed in a 1Hz bandwidth. That means that low-level discrete spurious signals (and narrowband noise peaks typically encountered under vibration due to high Q mechanical resonances) will go undetected. The second problem is that the commercial available software employed in the PN measurement equipment discriminates between random noise and discrete spurious signals but when
a reasonably sharp increase in noise level is detected; the system software assumes that the increase marks the presence of a “zero bandwidth” discrete signal. It is therefore when displaying the phase noise on a 1 Hz bandwidth basis, applies a bandwidth correction factor to the random noise, but does not make a correction to what is interpreted as a discrete signal. This results in an erroneous plot if when the detected “discrete” is really a narrowband noise peak.

A. Phase Noise Measurement Techniques

Following are the primary phase noise measurement techniques, listed in the order of increasing precision: (i) Direct Spectrum Technique, (ii) Frequency discriminator method—Heterodyne digital discriminator method, (iii) Phase detector techniques—Reference source/PLL method, (iv) Residual Method, and (v) 2-channel cross-correlation technique. The Direct Spectrum Method, PLL method, delay line discriminator method, and cross-correlation methods are frequently used to measure the oscillator phase noise. The first one is the simplest and has the biggest limitation. The last one requires the most complex measurement system but is the most versatile, can measure PN performance better than that of its reference oscillator.

Figure (5) shows the typical block diagram of the 2-channel cross-correlation technique. The oscillator under test is measured simultaneously by two separate phase-to-voltage transducers (in the dashed boxes), and by a dual-channel FFT analyzer. The O/P of each channel as shown in Figure (5) is

\[
x(t) = a(t) + c(t) \leftrightarrow X(f) = A(f) + C(f) \quad (2)
\]

\[
y(t) = b(t) + c(t) \leftrightarrow Y(f) = B(f) + C(f) \quad (3)
\]

where \(a(t), b(t)\) and \(c(t)\) are random signals; \(a(t)\) and \(b(t)\) are the noise of the transducers, and \(c(t)\) is the DUT noise; the upper case is used for the Fourier transform, and the ‘\(\leftrightarrow\)’ stands for the transform / inverse-transform pair. All signals are sampled at a suitable rate, and each acquisition takes the measurement time \(T\). The PSD (Power Spectral Density) is a complex concept of mathematical probability, defined as the Fourier transform of the covariance function. The cross PSD is

\[
S_{xy}^2(f) = \frac{2}{T} Y(f) X^*(f) \quad (4)
\]

where the superscript ‘1’ means single-sided (no negative frequencies) and will be omitted hereinafter; the symbol ‘\(^*\)’ means complex conjugate; and the factor ‘2’ is necessary for power consistency, after removing the negative frequencies.

Equation (4) implies the mathematical expectation, which in experiments is replaced with the average over a suitable number ‘\(m\)’ of samples, denoted with \(S_{xy}\)\(_m\). The average measurement takes a time \(mT\), plus computing time. Using (2)-(3), and omitting ‘\(f\)’ we get

\[
(S_{xy})_m = \frac{1}{m} \sum_{n=1}^{m} (B_n A_n^* + B_n C_n^*) + (C_n A_n^* + C_n C_n^*) \quad (5)
\]

Notice that \(B_n A_n^*\), \(B_n C_n^*\), and \(C_n A_n^*\) are in general complex, while \(C_n C_n^*\) is always real. The random signals \(a(t), b(t)\) and \(c(t)\) are statistically independent because they originate from fully separate circuits.

Also, and for the same reason, their Fourier transforms are statistically independent. As a consequence, the background noise \((B_n A_n^*, B_n C_n^*,\) and \(C_n A_n^*)\) terms is rejected proportionally to \(\frac{1}{\sqrt{m}}\), and for large \(m\) the average \((S_{xy})_m\) converges to the DUT noise \(S_{DC}^\). From (5), the DUT noise through each channel is coherent and is therefore not affected by the cross-correlation, whereas, the internal noises generated by each channel are incoherent and diminish through the cross-correlation operation at the rate of \(\sqrt{M}\) (\(M = \text{number of correlations}\)), given by

\[
[\text{Noise}]_{\text{meas}} = [\text{Noise}]_{\text{DUT}} + [\text{Noise}]_{\text{channel1}} + [\text{Noise}]_{\text{channel2}} \quad (6)
\]

where \([\text{Noise}]_{\text{meas}}\) is the total measured noise at the display; \([\text{Noise}]_{\text{DUT}}\) the DUT noise; \([\text{Noise}]_{\text{channel1}}\) and \([\text{Noise}]_{\text{channel2}}\) are the internal noise from channels 1 and 2, respectively; and ‘\(M\)’ is the number of correlations.

From (6), the 2-channel cross-correlation technique offers superior noise measurement capability but the measurement speed suffers when increasing the number of correlations. Figure (6) shows the measured phase noise plots (with and without cross-correlation) of 10 MHz crystal oscillator, cross-correlation offers 10-20 dB improvement in PN performance.

B. Cross-correlation: Uncertainty in PN Measurement

Cross-spectral analysis is a mathematical tool for extracting the power spectral density of a correlated signal from two time series in the presence of uncorrelated interfering signals. A major crux of the system described is the inherent impossibility to divide the DUT (device under test) noise from any other correlated effect. It is self-evident that vibrations or EMI (electromagnetic interference) hitting simultaneously on the two channels as depicted in Figure (5) cannot be rejected. However disturbing, experience is often useful to identify these perturbations as artifacts. Other effects are more subtle, and difficult to identify.

The cross-spectrum of two signals \(x(t)\) and \(y(t)\) is defined as the Fourier transform of the cross-covariance function of \(x\) and \(y\). Introducing a disturbing signal \(d(t)\) impacting on the two channels, from (2)-(3) rewrite as

\[
x = a + c + c x d \leftrightarrow X = A + C + c x D \quad (7)
\]

\[
y = b + c + c y d \leftrightarrow Y = B + C + c y D \quad (8)
\]
The high quality factor resonator provides the frequency-determining module; the electronic feedback system injects the energy needed to sustain the oscillation without much disturbing the resonator frequency. The sustained oscillation state forms a limit cycle in the phase space of dynamical variables of the system. A limit cycle in a deterministic system is periodic with frequency precision; divergence from this state caused due to inherent noise from thermal, electronic, vibrational and other sources [15]. In general, the oscillator phase noise will result from the combination of many different noise sources, perhaps with different spectra, leading to complicated frequency dependences for the total noise.

By varying the drive-level and optimizing the resonator nonlinearity using tuning to special operating points can improve the noise performance. Two types of noise acting directly on the resonator are expected. Firstly, theromechanical noise for a mechanical resonator and Johnson noise for an electronic resonator. The spectrum is usually white and the noise intensity is proportional to the temperature and the dissipation coefficient. There may also be additive noise associated with the nonlinear dissipation. The second type of noise is a parameter fluctuation. The spectra of these noises may be white, white filtered by a response of the device, e.g. thermal fluctuations will be quenched above a time scale determined by thermal contact to the environment.

### A. Resonator Operation in Nonlinearity Regime

The conventional approach is to drive resonator within its linear regime, which results in oscillator phase noise being inversely proportional to the oscillator carrier power. However, there is a demand for reduction in the oscillator size and power without compromising phase noise performance. When resonator is driven into its nonlinear regime, oscillation frequency becomes dependent on the amplitude of oscillation. In addition to this, with the reduction in size, their dynamic range also diminishes because nonlinear effects manifest at lower amplitudes [14].

The solution to above problems rely on the local elimination of frequency to energy dependence, evasion of amplifier noise, use of either parametric feedback, non-degenerate parametric drive, or coupling to internal resonances; however these approach are not cost-effective and offers unified solution [14]. The novel approach is to use the nonlinear behavior of the resonator to reduce the phase noise, exploiting the phase space geometry of the noise forces and the phase sensitivity of the resonator. Recent publication shows the possibility of improved phase noise performance of high q-factor resonator under nonlinear regime, where the phase noise performance is improved beyond the limitations of the linear regime. There is a possibility of existence of a special region in the parameter space, above the nonlinear threshold where the dominant contributions to the phase noise are suppressed [14]. By varying phase and feedback signals of the feedback oscillator, full exploration of the input parameter space can be obtained. The conventional phenomenological understanding is questionable, which assumes that resonator-operating regime beyond the threshold of nonlinearity necessarily degrades phase noise. The recent research reported in [14], using piezoelectric NEMS doubly clamped beam...
resonator made from an aluminum nitride (AIN) and molybdenum (Mo) multilayer, operating in nonlinear region where the signal level can be increased to large values without the conventionally expected performance phase noise degradation; the improvement in phase noise performance is experimentally verified. It is therefore possible to overcome fundamental limitations of oscillator performance due to thermodynamic noise. As is known for nonlinear resonators, when the driving force is sufficiently large, the system of resonator networks can bifurcate into three possible solutions at a given drive frequency; two of these are stable, and one is unstable [16]. We report to results for the oscillator phase noise, focusing in particular on special operating points of the oscillator where the detrimental effects of the resonator noise are reduced by incorporating dynamically tuned conduction angle feedback circuitry. Figure (7) shows the typical schematic and layout of ultra low phase noise 2.4 GHz SAW oscillator circuit, where SAW resonator driven into nonlinear regime window. The reduction in phase noise is achieved by using the feedback phase to tune the oscillator to operating points where the sensitivity to particular noise sources are reduced and choosing a feedback level and phase so that the resonator is driven where the amplitude-frequency and phase-frequency curves of the driven resonator become non-monotonic. The phase noise plots shown in Figure (8) is best performance (-150 dBc/Hz @ 10 kHz offset) reported to date for this class technology.

B. Collapse of Cross-Correlation Engine: Remedy!

The cross-correlation PN measurement technique was applied to isolate any form of injection locking or inadvertent non-linearity but inherent flaw of phase inversion gave erroneous result. Figure (8) shows the misleading plot for 2.4 GHz SAW oscillator, an optimistic value of -202dBc/Hz at 1.2 MHz off the carrier and 15 dB inferior at 6 MHz offset depending upon +ve/-ve phase inversion respectively.

Fig. 8: Measured phase noise plots of 2.4GHz SAW oscillator circuit

We witness the variation in PN measurement due to harmonics, load and phase mismatch, and cable length (delay). These variations can be minimized by incorporating tuned phase-tuned filter and impedance level at output so that impact of +ve/-ve phase inversion can be reduced significantly.

IV. CONCLUSION

The reported paper presents low phase noise 2.4 GHz SAW oscillator circuit, demonstrates stable signal source for the applications in clock and high frequency sources.

REFERENCES

Model for Acoustic Locking of Spin Torque Oscillator

Tanay A. Gosavi
OxideMEMS Lab, Cornell University
Ithaca, NY, USA
tag75@cornell.edu

Sunil A. Bhave
Analog Devices Inc.
Woburn, MA, USA
sunil.bhave@analog.com

Abstract—This paper presents a model for locking Spin Torque Oscillator (STO) to an out-of-plane AC strain generated using a mechanical transducer like High-Overtone Bulk Acoustic Resonator (HBAR). We model the magnetization dynamics of the free layer magnet in the STO using the Landau-Lifshitz-Gilbert-Slonczewski (LLGS) equation modified to include magneto-elastic coupling term. Locking is clearly demonstrated from the simulated frequency spectrum of the acoustically-locked STO which has a narrower linewidth and higher signal power as compared to a free running STO. We have shown locking of different modes of the STO and compare the amplitude of out-of-plane AC strain needed to achieve lock. Acoustic locking illustrated here can be used for locking multiple STOs to a common strain transducer and is a potential platform for developing hybrid magneto-acoustic oscillator systems.

Keywords—Spin Torque Oscillators; Frequency Locking; HBARS, Macro-Spin dynamics

I. INTRODUCTION

STOs are nanoscale GHz frequency self-oscillating magnetic tunnel junctions (MTJ) or spin valve devices [1,2]. In STOs, spin transfer torque generated by spin-polarized electrons overcomes damping in the free layer magnet of the MTJ or spin valve causing steady state precession of the magnetization. Using Tunnel Magneto-Resistance effect (in MTJ) or Giant Magneto-Resistance effect (in Spin Valves) magnetization precession is read out as oscillations of effective impedance of the device. This oscillating impedance modulates the DC bias current to generate an AC current at the frequency of impedance oscillations. STOs have extremely small form factor and octave spanning tuning range but have low output power and poor phase noise limiting their practical use. Locking multiple STOs to a common external reference oscillator has been proposed as a solution and injection locking scheme improves the output power and reduces the oscillator factor and octave spanning tuning range but have low output power and poor phase noise limiting their practical use.

II. MODEL AND SIMULATION SETUP DESCRIPTION

A. LLGS: Macrospin Model of STO

Magnetization dynamics of the free layer of a spin torque oscillator are be modelled using the LLGS equation shown in (1). \( \ddot{\mathbf{m}} = \frac{\mathbf{M}}{M_s} \) is a unit vector which represents the magnetization \( \mathbf{M} \) of the free layer, where \( M_s \) is the saturation magnetization. The first term on the right side of (1) describes the precession of the magnetization \( \mathbf{m} \) around the effective magnetic field \( \mathbf{H}_{\text{eff}} \), where \( \gamma \) is the gyromagnetic ratio of the electron in free layer of the STO.

\[
\frac{d\mathbf{m}}{dt} = -\gamma \mathbf{m} \times \mathbf{H}_{\text{eff}} + \alpha \mathbf{m} \times \frac{d\mathbf{m}}{dt} + \frac{\gamma h}{2e \mathbf{M}_s} \mathbf{m} \times (\mathbf{m} \times \mathbf{d})
\]  

The second term on the right side is the damping term, where \( \alpha \) is the phenomenological Gilbert damping constant. The effect of the damping term is to relax the magnetization \( \mathbf{m} \) along \( \mathbf{H}_{\text{eff}} \). The value of Gilbert damping constant is material dependent and has origin in magnon-magnon and magnon-phonon scattering [2]. The third term on the right side of (1) is the spin torque term, where \( J \) is the current density, \( h \) is the reduced plank constant, \( \eta \) is the spin polarization efficiency, \( e \) is the charge of an electron and \( d \) is the thickness of the free layer. Here \( \mathbf{d} \) is the spin polarizer vector which is along the direction of the polarizer or the fixed layer magnet in a STO. \( \mathbf{d} \) in (1) is used to represent the direction of the spin polarized electrons that are applying a torque to the magnetization of the free layer. Amplitude and the sign of the current density \( J \) have direct control over the spin transfer torque amplitude and direction. Increasing the amplitude of bias current the spin torque can overcome the damping in the free layer and result in steady state precession of magnetization.

\[
\mathbf{H}_{\text{eff}} = \frac{dE_{\text{eff}}}{dm}
\]

\[
\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{ap}} + H_{an}(\tilde{\mathbf{t}}, \mathbf{m}) + H_{d}(\tilde{\mathbf{d}}, \mathbf{m}) + H_{II} + H_{sd}(\tilde{\mathbf{t}}, \mathbf{m})
\]

The total effective magnetic field \( \mathbf{H}_{\text{eff}} \) for the free layer magnet is the derivative of its total energy \( E_{\text{eff}} \) with respect to its magnetization \( \mathbf{m} \) as shown in (2). The components of \( \mathbf{H}_{\text{eff}} \) shown in (3) are derived from constituent energy terms of the magnet. \( \mathbf{H}_{\text{ap}} \) is the external applied magnetic field which is used as a bias field. By varying \( \mathbf{H}_{\text{ap}} \) the frequency of oscillation of the STO can be changed. \( H_{an} \) and \( H_d \) are the anisotropy and demagnetization fields whose amplitude and
directions (\(\vec{n}\) and \(\vec{d}\)) are dependent on the shape of the magnet [6]. The size and shape of the free layer magnet also determines the initial direction of the magnetization \(\vec{m}\). In simulation, direction of magnetization is given in polar coordinates with the azimuthal angle \(\varphi\) and polar angle \(\theta\).

B. Thermal Noise in LLGS

The effect of non-zero temperature \(T\) and the corresponding thermal noise in dynamics of magnetization is accounted via term \(H_T^\prime\) in (3). \(H_T^\prime\) is a Gaussian-distributed Langevin term with random direction whose RMS amplitude calculated by fluctuation dissipation argument is shown in (4). In (4) \(\Delta t\) is the simulation step size, \(V\) is the volume of the free layer magnet and \(k_B\) is the Boltzmann constant [7]. Thermal noise is also accounted for in the model by assuming that the starting azimuthal angle \((\varphi_0)\) and polar angle \((\theta_0)\) of the magnetization have random fluctuations around them. The RMS amplitudes of these thermal fluctuations are calculated using equipartition theorem and are as shown in (5) and (6) for \(\varphi_0\) and \(\theta_0\) respectively.

\[
H_T^{\text{RMS}} = \sqrt{2\alpha k_B T/\gamma M_s V \Delta t} \quad (4)
\]

\[
\varphi_0^{\text{RMS}} = \sqrt{k_B T/H_{an} M_s V} \quad (5)
\]

\[
\theta_0^{\text{RMS}} = \sqrt{k_B T/H_{ad} M_s V} \quad (6)
\]

\[
H_{st} = B_{eff} S/M_s \quad (7)
\]

\[
\Delta R = (1 - \vec{m} \cdot \vec{\theta})/2 \quad (8)
\]

C. Magneto-Elastic locking term

The effect of out-of-plane strain applied by the HBAR on the free layer magnet is incorporated in the magnetization dynamics as an effective out-of-plane uniaxial shape-anisotropy term \(H_{st}\), which is part of \(H_{\text{eff}}\) as shown in (3) [5]. Equation (7) gives the amplitude of \(H_{st}\), where \(B_{eff}\) is the material dependent magneto-elastic coupling coefficient. Amplitude of \(H_{st}\) is proportional to \(S\) which is the amplitude of the AC strain along direction \(\ell\) acting on the free layer of STO.

D. Simulation Setup

In our simulation of STO we use a MTJ with CoFeB for both free and fixed layer which are separated by MgO based tunnel barrier. The size of the free layer and its relevant magnetic properties taken from [8], are shown in Table I. Applied bias magnetic field and DC bias current amplitude used in simulation are also shown in Table I. For our simulation model with the chosen elliptical shape and size of the free layer the \(H_{an}\) field is along the long axis of the magnet in X direction and \(H_{ad}\) field which keeps the magnetization from pointing out-of-plane is along the Z direction. The initial direction of magnetization (\(\vec{n}\)) is along the +X axis. External applied bias field \(H_{an}\) of amplitude 300 Oersted is along the easy axis of the magnet in +X direction as shown in Fig. 1. HBAR as the transducer will generate strain in Z direction which is the out-of-plane direction for the magnet. Thus the magneto-acoustic field due to strain \(H_{st}\) will be along the \(Z\) direction as shown in Fig. 1. The direction of the polarizer \(\vec{\theta}\) is chosen along the +X direction. In the simulation we use step time \((\Delta t)\) of 10ps, while the total simulation time is 1\(\mu s\) giving a resolution of up to 2MHz in the extracted frequency spectrum. Simulations were first done without the AC strain \(S\) to extract the original spectra of the STO and then were repeated with strain field at the oscillation frequency. The strain field is turned on after the initial 30ns which is sufficient time for magnetization to settle into steady state precession resulting in oscillations of impedance \(\Delta R\) of STO.

III. EXPERIMENTAL RESULTS AND DISCUSSION

Frequency of the simulated STO changes with the applied bias current as shown in Fig. 2. There are two main regions of operation as seen in Fig. 2 which correspond to two different modes of oscillation for the STO. First mode is the in-plane mode shown in Fig. 3A, where the oscillation frequency decreases with the increasing bias current. Second is when the oscillation frequency increases with the bias current. This mode of oscillation is called the out-of-plane mode (shown in Fig. 3B). As the STO transitions from in-plane mode to out-of-plane mode it goes through a region where the mode has clam-like shape. In this mode the magnetization tilts out-of-plane in both the +Z and –Z directions. The clam-shaped mode (shown in Fig. 3C) is considered as a subset of the in-plane mode as it follows the same frequency-current relation.

![Fig. 1. Directions of M, H_{ap}, H_{st}, Polarizer (\vec{\theta}) and the coordinate system used in the simulations.](Image)

![Fig. 2. STO frequency and oscillation mode as a function of bias DC current.](Image)

### Table I. Simulation Parameters

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free layer size</td>
<td>170 nm × 100 nm × 1.5 nm</td>
</tr>
<tr>
<td>(M_s)</td>
<td>1100 erg/cm(^3)</td>
</tr>
<tr>
<td>(H_{an})</td>
<td>219 0e</td>
</tr>
<tr>
<td>(H_d)</td>
<td>13000 0e</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.01</td>
</tr>
<tr>
<td>(B_{eff})</td>
<td>(-7 \times 10^7)  emu/cm(^3)</td>
</tr>
<tr>
<td>(\eta)</td>
<td>0.3</td>
</tr>
<tr>
<td>(H_{ap})</td>
<td>300 0e</td>
</tr>
<tr>
<td>Current Range</td>
<td>1.8 – 8 mA</td>
</tr>
</tbody>
</table>

Fig. 2. STO frequency and oscillation mode as a function of bias DC current.
Effect of the applied AC strain and the resultant locking of the STO can be studied using FFT based frequency spectrum of its modulated impedance $\Delta R$. $\Delta R$ at any given time is dependent on the relative directions of magnetization ($\vec{m}$) and the polarizer ($\vec{e_p}$) as shown in (8). Amplitude of $\Delta R$ is also determined by the material properties and the device structure but is not relevant for the frequency spectrum analysis. For FFT we consider the data only after initial 60ns which is sufficient time for magnetization to settle into steady state precession and for it to respond to the locking AC strain $S$, which is turned on 30ns after starting the simulation. Fig. 4 shows the FFT spectrum of different oscillation modes without locking strain in blue and with AC locking strain in red. The relative amplitude of oscillation seen in the blue spectrum across different oscillation modes is reflective of the actual amplitude measured in different experiments [1,3,9]. Specifically the output power of STO increases from in-plane mode (Fig. 4A) to out-of-plane mode (Fig. 4B) with the clam-shaped mode (Fig. 4C) of oscillation having intermediate output power. Higher signal power, narrower linewidth and suppressed suprious modes in the red spectrum compared to the blue spectrum (in Fig. 4) is indicative of the locking of the STO across various modes by the AC strain $S$ from the mechanical resonator.

The amplitude of the AC strain needed to lock STOs in different modes of oscillations is shown in Table II. Out-of-plane mode needed 21ppm of AC strain, while in-plane mode

<table>
<thead>
<tr>
<th>Oscillation Mode</th>
<th>Locking AC Strain $S$ (ppm)</th>
<th>HBAR Drive (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Plane</td>
<td>107</td>
<td>4.9</td>
</tr>
<tr>
<td>Clam-Shaped</td>
<td>33</td>
<td>2.2</td>
</tr>
<tr>
<td>Out-of-Plane</td>
<td>21</td>
<td>1.4</td>
</tr>
</tbody>
</table>
required 107ppm. The huge difference in amplitude of strain $S$ is because the magneto-acoustic locking field $H_{st}$, with direction $\hat{t}$, couples into the effective field $H_{eff}$ as a dot product with magnetization $\hat{m}$. For the Z directional strain used in our simulations the amplitude of $H_{st}$ is weighted by the Z component of $\hat{m}$. In in-plane mode of oscillation the magnetization precessions is mostly along the X-Y plane with a very small Z component, so the locking field amplitude is greatly reduced when it is coupled into the LLGS equation. Hence larger amplitude of AC strain is required for locking in-plane mode compared to out-of-plane or clam-shaped modes, which have larger Z component of magnetization. Table II also shows the voltage drive needed by an AlN HBAR with a silicon substate to generate the locking strain amplitude at the frequency of the STO oscillation. The drive voltage values were calculated using an analytical model for HBAR assuming a quality factor of 1000 and AlN film thickness of half the wavelength [10]. Drive values calculated from the analytical mode are in agreement with finite element analysis done in Comsol.

The locking range of AC strain $S$ for the clam-shaped mode of oscillation with frequency 5.766GHz is shown in Fig. 5. The locking range is linear over an offset frequency range ($\Delta f_{lock}$) of $+/−$ 105MHz for 110ppm amplitude of AC strain $S$. The limit of locking range, observed as a saturation of the offset frequency, was not seen in our simulations; as we limited our simulations to a maximum strain amplitude of 110ppm, which is very difficult to achieve using HBAR based mechanical strain transducer.

IV. CONCLUSION

In conclusion, we have demonstrated acoustic locking of Spin Torque Oscillator using a realistic macro-spin model of MTJ with a magneto-elastic coupling term. The model used, takes into account effect of thermal noise and shows the enhancement in linewidth and amplitude of oscillation when subjected to AC locking strain. We have demonstrated locking for various oscillation modes of STO and observe that out-of-plane strain generated by the HBAR is more suited for locking clam-shaped mode or for out-of-plane mode. Acoustic locking illustrated here can be used for synchronizing large number of STOs by locking them to a single mechanical strain transducer. We believe using acoustic locking with HBAR-like transducer in feedback loop with STO an octave frequency tunable hybrid magneto-acoustic oscillator system can be developed.

ACKNOWLEDGMENT

The authors would like to thank the Cornell Center for Materials Research supported by National Science Foundation MRSEC program (DMR-1120296), whose generous grant made this research possible. We also would to thank Dr. Praveen Gowtham for providing measured data on magneto-elastic coupling coefficients of CoFeB-MgO layers and for discussions on acoustic locking and Spin Torque Oscillators.

REFERENCES

Piezoelectrically-Acutated Opto-Acoustic Oscillator

Siddhartha Ghosh, Jeronimo Segovia-Fernandez and Gianluca Piazza
Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, PA, USA
sidghosh@cmu.edu

Abstract—This paper describes the implementation of an opto-acoustic oscillator with the use of an integrated piezoelectrically-driven acousto-optic intensity modulator in aluminum nitride. Optical output collected from on-chip grating couplers is converted into the electrical domain via an avalanche photodiode, and used to drive the modulator by providing an input to its RF probing pads. Oscillation is observed at 35.2 MHz and 652.8 MHz, corresponding to peaks in the forward transmission of the modulator. An initial phase noise of -72 dBc/Hz at 10 kHz offset from the 652.8 MHz carrier is recorded, with a large contribution from thermal noise in the amplifier chain. Some proposals for future improvements in phase noise performance are provided.

Keywords—Opto-acoustic oscillator; piezoelectric actuation; optical modulator; RF-Photonics; Optical MEMS

I. INTRODUCTION

Highly stable reference oscillators producing spectrally pure signals in the range of MHz to GHz are required in numerous communication systems. Traditionally, quartz resonators have been applied in radio frequency (RF) systems thanks to extremely high Q-factors in the range of 10-100 MHz. Their use however suffers from the need to perform frequency multiplication to operate in higher bands, which also multiplies oscillator noise, as well as a lack of on-chip integration. Two promising technological alternatives, which have addressed different drawbacks associated with the state of the art, include the use of photonic-based and microelectromechanical systems (MEMS) based oscillators. In the case of the former, such as the opto-electronic oscillator [1]-[2], excellent phase noise performance has been achieved thanks to the use of a high-Q energy storage element, namely a long (km range) optical fiber loop. On the other hand, MEMS technology [3] benefits from compatibility with prevailing semiconductor fabrication techniques to enable single-chip oscillator solutions [4].

A combined approach, in which the electro-optic modulator is replaced with a MEMS-based acousto-optic modulator has also been demonstrated [5] for chip-scale miniaturization of the opto-electronic oscillator. This implementation, and other recent demonstrations [6] have involved the use of electrostatic actuation in the optical modulation scheme. Piezoelectric actuation however, presents a number of unique advantages, including lower motional impedances and the ability to interface with 50Ω electronics. To this end, aluminum nitride (AlN) contour mode resonators have generated a great deal of interest over the past several years, and have been successfully demonstrated in a number of oscillator implementations [7]-[9]. The AlN acoustic resonators are CMOS compatible and have fundamental frequencies set by lithographically defined dimensions. At the same time, AlN is an efficient optical material, well-suited for photonic applications with low losses over a wide range of wavelengths [10]-[11]. As a result, it is possible to consider the development of Opto-Acoustic Oscillators as well, by leveraging the strengths of piezoelectric actuation and integration in AlN.

In this work, we describe the demonstration of a piezoelectrically-actuated Opto-Acoustic Oscillator in AlN. The integrated acousto-optic modulator [12] used in the oscillator loop provides amplitude modulation for light in the telecommunications C-band. The optical output is then collected with an avalanche photodetector (APD) and converted back into the electrical domain to provide RF input to the device. In this manner, the loop functions on the same principle as the opto-electronic oscillator, with exception to the fact that the mechanical resonator forms the frequency-selective element. Thus, while the loop behaves similarly to a conventional MEMS oscillator, we will consider some strengths which may be leveraged through concurrent integration in the photonic domain.

II. ACOUSTO-OPTIC MODULATOR TRANSMISSION

The acousto-optic modulator forming the basis of the oscillator consists of two coupled AlN rings which function as an acoustic contour mode resonator and photonic whispering gallery mode resonator. RF excitation is supplied to the electrodes sandwiching the first ring, which generate lateral mechanical vibrations through the d_{31} piezoelectric coefficient. These vibrations are transferred into the second ring through a mechanical coupling spring. Simultaneously, the photonic ring supports a traveling wave optical mode confined to its periphery, which has been introduced through a coupled integrated photonic waveguide. Due to the mechanical vibrations acting on the photonic ring, its outer boundary undergoes a deformation, which alters the optical resonance condition. This produces a shift in the ring’s resonant wavelength (or optical frequency) proportional to the average amount of change in the ring’s outer radius. When the wavelength of the pump laser is detuned from the cavity resonance to a point corresponding to a maximal degree of change in the output transmission, the wavelength modulation can be detected as a modulated optical power transmission.
Relevant details of the modulator under consideration, and its fabrication process are provided in [12]. A scanning electron micrograph (SEM) of the device is shown in Figure 1.

![Fig. 1. Integrated AlN acousto-optic modulator.](image)

Prior to testing the modulator transmission, the optical response is first characterized by aligning the on-chip grating couplers to a fiber-array probe, which holds two single mode optical fibers spaced by 250 μm to send in transverse-electric (TE) mode polarized light from a tunable laser source, and collect the output with an optical power meter. Each grating introduces a loss of approximately 17.5 dB, which is partially attributed to sensitivity to alignment accuracy. In order to increase the output transmission power, +15 dBm of optical power is supplied from the laser source. A high optical quality factor (Qopt = 25,400) resonance at 1528.030 nm is selected for the testing. As a result of the thermo-optic heating in the ring from the increased input laser power, the resonant wavelength undergoes a red shift. Thus the bias-point for the cavity detuning is selected experimentally by monitoring the wavelength corresponding to a peak in output power modulation when RF excitation is provided. This wavelength is determined to be 1528.020 nm, and remains fixed for the subsequent testing.

Once the optical cavity is properly detuned from the laser source, the forward transmission (S21) of the optical modulator can be characterized with the use of a vector network analyzer (VNA). Input RF excitation is provided to the ground-signal-ground (GSG) pads of the modulator from Port 1 of the VNA. The modulated optical signal collected from the output grating of the device is then fiber-coupled to the APD (New Focus 1647). Since the optical modulation power is relatively low, it is possible to adjust the bias setting on the APD to correspond to a high conversion gain of 14,000 V/W. This electrically converted optical output is then supplied to Port 2 of the VNA. By measuring the electro-optomechanical response, it is possible to characterize the mechanical modes of the device, similar to the admittance response used in conventional acoustic contour mode resonators. The S21 output thus shows the occurrence of two prominent mechanical modes. The first is at 35.2 MHz, and has an associated mechanical quality factor (Qmech) of 120. The second corresponds to the main contour mode resonance in the first (electroded) ring at 652.8 MHz. This mode has an associated Qmech of 450. The transmission response of both modes is shown in Figure 2.

![Fig. 2. S21 transmission of two mechanical modes for the acousto-optic modulator at 35.23MHz (top) and 652.68MHz (bottom).](image)

III. OPTO-ACOUSTIC OSCILLATION

The acousto-optic modulator is then configured as an oscillator by closing the loop from the output of the APD. In order to satisfy the Barkhausen stability criteria, the loop must achieve unity gain and zero phase shift. The first requirement is met by using an RF amplifier to compensate for losses in the modulator transmission. In order to provide sufficient gain, two amplifiers (Mini-Circuits ZKL-1R5+) are cascaded in series. The second criterion is met by sending the amplified output through a phase shifter (Colby Instruments PDL-100A). This signal can then be routed through a power splitter to feed back into the modulator input, and monitor the oscillation simultaneously. While testing the oscillator response, the voltage bias of the two amplifiers is varied in order to achieve the necessary gain condition. In addition, the oscillator locks more readily to the low frequency mechanical mode. In order to observe the oscillation at 652.8 MHz, a high-pass filter (Mini Circuits SHP-400+) with a cutoff frequency of 395 MHz is introduced in the oscillator loop. A schematic corresponding to these elements in the opto-acoustic oscillator loop is shown in Figure 3.
Phase noises are provided in Figure 4. Plots for the carrier signal at each mode and the corresponding phase condition, the oscillation is monitored on a spectrum analyzer. For the 35.2 MHz mode, an output RF power of 7.75 dBm is measured. Likewise, for the 652.8 MHz mode, the carrier signal has an RF power of 6.51 dBm. Phase noise for the oscillator was also measured with a signal source analyzer.

The spectrum of the oscillation was recorded for both modes at 35.2 MHz and 652.8 MHz. The initial phase noise recorded for the oscillator appears to have a large component of thermal noise originating from the photodetector and amplifier chain. This can be attributed to the fact that a relatively low photocurrent is measured at the device output. While this limits the contributions of shot noise and laser intensity noise, it also generates a large noise to signal ratio [13]. Phase noise performance can be improved by increasing the photocurrent collected at the output. This may be achieved by implementing process-related improvements to increase the optical quality factor of the device, which dictates the transmission capability of the modulator. As a result, improvements in the forward transmission ($S_{21}$) should also eliminate the need for external amplifiers to compensate for losses in the loop.

IV. DISCUSSION

We have demonstrated an opto-acoustic oscillator with the use of piezoelectric actuation using a monolithically integrated AlN acousto-optic modulator. Oscillation is observed for two mechanical modes at 35.2 MHz and 652.8 MHz. The initial phase noise recorded for the oscillator appears to have a large component of thermal noise originating from the photodetector and amplifier chain. This can be attributed to the fact that a relatively low photocurrent is measured at the device output. While this limits the contributions of shot noise and laser intensity noise, it also generates a large noise to signal ratio [13]. Phase noise performance can be improved by increasing the photocurrent collected at the output. This may be achieved by implementing process-related improvements to increase the optical quality factor of the device, which dictates the transmission capability of the modulator. As a result, improvements in the forward transmission ($S_{21}$) should also eliminate the need for external amplifiers to compensate for losses in the loop.

ACKNOWLEDGMENT

This work was completed with support from the National Science Foundation under award NSF ECCS-1201659 and DARPA under the MESO program award # FA86501217624. The devices were fabricated in the Carnegie Mellon Nanofabrication Facility.

REFERENCES


UHF SiGe Push-Pull VCO MEMS Oscillators

HRL Laboratories, LLC, Malibu, CA, 90265 USA

Email: yyoon@hrl.com

Abstract—UHF quartz MEMS oscillators operating at 813 MHz, 920 MHz and 1.048 GHz carrier frequencies utilizing a push-pull topology based on MOSIS SiGe 7WL technology have been demonstrated. The 1048 MHz resonator has an unloaded Q of 6,920 and a motional resistance of 31.7 ohms, as measured in a vacuum. This yields an fQ product of 7.25x10^13, close to the expected limit for quartz devices of 1x10^15. Device phase noise was measured showing a minimum phase noise of -100dBc/Hz for an 813 MHz oscillator at 1 kHz offset. A tuning range of 80 ppm was demonstrated at 920 MHz.

Keywords— Quartz, MEMS, Resonator, Oscillator, Push-pull, Phase noise, tuning

I. INTRODUCTION

High Q UHF oscillators based on quartz MEMS resonators are highly desirable for commercial and military electronics applications such as high performance navigation, radar systems, and communication systems. Offering the potential for low phase noise, the resonators also have excellent stability over temperature and vibration, more so than competing surface acoustic wave (SAW) and bulk acoustic wave (BAW) counterparts. Additionally, SAW and BAW devices typically operate at frequencies <100 MHz, thus, frequency multiplication units for UHF operation are commonly required, unlike the UHF quartz technology.

Prior work [1] in UHF quartz technology focused on developing oscillators operating at fixed frequencies. This was done as a proof of concept for quartz oscillators using SiGe sustaining circuitry. Unfortunately, the temperature compensation requirement in many applications is not possible with a fixed frequency approach. This work extends our previous efforts by adding oscillator frequency tunability to enable temperature compensation. To do this we developed a voltage controlled oscillator (VCO) by adding MOS varactor tuning elements to our circuitry. For these designs, we continued to use the IBM 7WL SiGe technology provided by MOSIS. This technology offers the complimentary p-n-p devices necessary for push-pull operation that are modeled at UHF. The process also includes 0.18μm CMOS for control circuitry and the associated n-p-n HBT process has maximum unity gain frequencies, f, between 40 and 60 GHz.

The same push-pull approach was used to avoid the challenges with the large inductors (~30-60 nH) necessary at UHF. These inductors have a large footprint and low quality factors that can degrade circuit performance. In addition, they require substantial effort to model accurately. Using a push-pull topology eliminates the need for highly inductive RF chokes because the common emitter p-n-p device connected on top requires an AC ground as opposed to an RF choke.

We have designed different UHF push-pull voltage controlled oscillators (VCO) utilizing quartz resonators. These designs were targeted to operate from 0.8 to 1.1 GHz. The VCO’s are characterized in terms of phase noise and measured tuning range versus tuning voltage.

II. DESIGN OF QUARTZ MEMS RESONATOR

A. Resonator Design

UHF AT-cut quartz resonators operating in the fundamental thickness-shear mode at 1 GHz formed the basis for the VCO’s and are designed using COMSOL finite element analysis (FEA). The resonators fabricated at HRL consist of an AT-cut quartz plate with a quartz cut angle of 35°15’, and top and bottom side rectangular aluminum electrodes. 3-D FEA simulations are performed to evaluate modal energy confinement, and the stability over temperature of frequency, Q, motional resistance (Rm), capacitance (Cm) and inductance (Ln) of the resonator.

Experiments utilized three resonators of two different designs, operating at 813, 920 and 1048 MHz. The 920 MHz resonator has stress relieving spring tethers for de-coupling residual bonding/mounting stresses from the active electrode region. If stress couples into the active region of the device, the frequency-temperature (f-T) characteristics can become distorted and deviate from the ideal f-T excursion [2]. The in-situ springs reduce stress in the active region and the f-T curve is undistorted providing a minimal change in frequency over temperature. The remaining two resonators do not have spring mounts but rather a conventional rectangular mounting base and 255 x 85 μm aluminum electrodes.

The detailed quartz resonator fabrication process has been described in [3]. A layout and photograph of the S120 spring design is presented in Fig 1. The left image details the S120 spring design. A fabricated device is shown on the right, showing the stress relieving spring elements. The FEA

Fig 1. Schematic and photograph of a quartz S120 resonator with spring mounts.
simulated eigenfrequency analysis of a UHF AT-cut quartz spring coupled resonator shown in Fig. 2 indicates excellent modal energy confinement between the Al electrodes. In this case, the AT-cut quartz plate is 240 µm x 240 µm x 1.5 µm, and the aluminum electrodes are 120 µm x 120 µm x 0.04 µm thick. The predicted thermal stability of frequency and Q with the spring mounts is presented in Fig. 3, showing ~22 ppm frequency stability over temperature. Similarly, the V3 resonator design, photograph of fabricated device and f-T behavior are repeated in Figures 4 and 5 for completeness. The V3 resonator has Al electrode dimensions of 255 µm x 85 µm, and also 0.04 µm thick. The quartz blank for the 1 GHz V3 resonator is slightly thinner than for the 813 MHz device.

B. Electrical Model

Fig. 6 illustrates a Van Dyke model representation of a resonator in combination with an amplifier used to simulate an oscillator. Table 1 summarizes the associated component values. C0 corresponds to the static parasitic capacitance of the resonators. Motional inductance, L1, and motional resistance, R1, correspond to the effective modal mass, m, and damping, γ, in the analogous mass-spring system. For the typical linear case, motional capacitance, C1, is analogous to spring stiffness, k. In Table 1, the 1.048 GHz quartz resonator, vacuum sealed with a silicon cap, has an unloaded Q ~6.9 k and an fQ product 7.25x10^12. Similarly the 813 MHz and 920 MHz quartz resonators, vacuum sealed with a silicon cap, have an fQ product 6.61x10^12 and 6.64x10^12 respectively. The fQ products compare favorably with the quartz theoretical limit of 1x10^13.

III. DESIGN OF VOLTAGE CONTROLLED OSCILLATOR (VCO)

The IBM 7WL process enables push-pull performance at UHF because its process design kit (PDK) includes RF models of the vertical p-n-p transistors. The VCO push-pull topology

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Component Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>V3</td>
</tr>
<tr>
<td>Series fo</td>
<td>813 MHz</td>
</tr>
<tr>
<td>Electrode (µm)</td>
<td>255 x 85</td>
</tr>
<tr>
<td>R1 (Ω)</td>
<td>43.9</td>
</tr>
<tr>
<td>C1 (pF)</td>
<td>0.526</td>
</tr>
<tr>
<td>L1 (µH)</td>
<td>72.6</td>
</tr>
<tr>
<td>C0 (pF)</td>
<td>0.887</td>
</tr>
<tr>
<td>Q</td>
<td>8125</td>
</tr>
<tr>
<td>f x Q product</td>
<td>6.61x10^{12}</td>
</tr>
</tbody>
</table>
shown in Fig. 7 is based on the Pierce configuration depicted in [1] with the same common emitter buffer amplifier that has a high input impedance. The basic device lineup consists of an n-p-n transistor with a base width of 240nm, an emitter width of 800nm, and an emitter length of 12µm and a vertical p-n-p transistor has a base width of 800nm and an emitter length of 10µm. This was done to try to understand how different device input impedances affect oscillator output power and phase noise. CMOS transistors are used for the "top" current source and the scheme is exactly the same as that in [1]. The quartz resonator is external to the push-pull sustaining circuit and is connected with wirebonds.

Tuning is achieved by replacing the four phase shift capacitors in the original topology with MOS varactors (please note that the schematic in reference [1] is missing the top and bottom capacitors to the right of the devices). This allows for a greater tuning range than VCOs using a single diode approach. Quartz VCO tunability and its application to temperature compensation is discussed in further detail in [4]. The 1.8V NMOS varactors used to replace the 1pF fixed capacitors in the original design have a capacitive range of 0.98 to 2.87pF with -1 to 1 volt applied through 5KΩ resistors with DC blocking capacitors to isolate the varactor bias from the rest of the circuit. Fig. 8 shows the current source biasing push-pull VCO core. The reference current, Iref, in Fig. 8 varies up to 4 mA with typical operation being 2mA.

**IV. MEASUREMENT**

VCO sustaining ICs were assembled and wire bonded as depicted in [1] with Quartz MEMS resonators shown in Table I and then placed on a wafer probe station for measurement. Phase noise of the VCOs was measured with an Agilent 5052A. Fig. 9 shows an oscillation spectrum peak at 1.0485 GHz. VCO circuits were biased with 1.8 ~ 2.4 V with a fixed Iref of the current source. It is found that the VCO’s provide lower phase noise when biased with 2.0V as shown in Fig. 10. When the bias voltage was set to higher than 2.4V, the 920 and 1048 MHz VCOs ceased to oscillate. Output power for the 813 MHz VCO varies from -17 to -4.5 dBm depending on Iref as depicted in Fig. 10. Although the 1.048 GHz VCO provides output power ranging -24 to -12 dBm by varying Vcc from 1.8 to 2.2V and Iref from 1 to 4mA, the lowest phase noise was measured at Vcc = 2V. Overall power consumption is -25 mW with Vcc = 2V and Iref = 4mA with most of the power consumption occurring in the buffer amplifier which typically consumes 12-15mW.

Table II shows a phase noise comparison with different VCO frequency and Q of quartz MEMs resonators. The three resonators with phase noises range from -89dBc/Hz for the 813 MHz VCO to -92dBc/Hz for the 920 MHz VCO to -94dBc/Hz for the 1048 MHz VCO. Fig. 10 shows the current source biasing push-pull VCO core. The reference current, Iref, in Fig. 8 varies up to 4 mA with typical operation being 2mA.
1048 MHz resonator to -100 dBc/Hz for the 813 MHz resonator at a 1 kHz offset. Recent work using AlN contour-mode resonators at 1.16 GHz demonstrates a phase noise of -82 dBc/Hz at the same offset [5]. Fig. 11 shows the comparison of simulated phase noise with measured data, where excellent agreement can be seen due to accurate device models of 7WL and resonator parameters as described Table I.

Previous phase noise results from hybrid VCOs using quartz MEMS resonators built by HRL have been reported with -115 and -113 dBc/Hz at 1 kHz offset with respect to 705 MHz and 663 MHz carrier frequencies respectively [6,7]. The hybrids used commercially available packaged silicon bipolar transistors [7]. When the measured phase noise of the hybrid VCO at a 1 kHz offset is scaled up to 800 MHz or 1 GHz and compared with the IC oscillator results, the phase noise performance by the hybrids is better than the noise values shown in Table II. From phase noise simulations in Cadence, it was found that p-n-p transistors in push-pull core contribute more than 60% of overall phase noise amounts with an extra 20% from PMOS devices used in the CMOS current source. Thus, it is desirable to avoid the p-n-p transistor used in the push-pull topology in order to optimize phase noise.

Tuning range generally increases with the number of tuning elements. However theoretical tuning potential varies with oscillator Q and the ratio of C0/C1, the parasitic capacitance over the internal resonator capacitance. Thus, the tuning range may be limited by factors other than the number of varactors.

The 920 MHz VCO uses two tuning varactors, one on either side of the n-p-n transistor and denoted by control voltages, Vd1 and Vd2 in the schematic depicted in Fig. 7. The other two varactors tuning the p-n-p transistor were replaced by fixed 1pF MIM capacitors. One varactor bias voltage is independently varied while the other varactor voltage is connected to ground. As seen Fig. 12 overall tuning ranges are approximately 62 ppm with the variation of a single control voltage. When both Vd1 and Vd2 are varied simultaneously, the tuning range increases to ~80 ppm which is equivalent to 73 kHz. The 1048 MHz resonator was paired with an IC with four tuning elements and achieved an overall tuning range of ~65 kHz.

### Table II. Measured Phase Noise Comparison

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Component Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>V3</td>
</tr>
<tr>
<td>Series f0</td>
<td>813 MHz</td>
</tr>
<tr>
<td>R1 (W)</td>
<td>43.9</td>
</tr>
<tr>
<td>f x Q product</td>
<td>6.61 x 10^12</td>
</tr>
<tr>
<td>Phase Noise @ 1kHz (dBc/Hz)</td>
<td>-100.4</td>
</tr>
<tr>
<td>813 MHz VCO</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>S120 Spring</td>
<td>920 MHz</td>
</tr>
<tr>
<td>L1 (µH)</td>
<td>61.1</td>
</tr>
<tr>
<td>6.64 x 10^12</td>
<td></td>
</tr>
<tr>
<td>Phase Noise @ 1kHz (dBc/Hz)</td>
<td>-91.4</td>
</tr>
<tr>
<td>920 MHz VCO</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>1048 MHz</td>
</tr>
<tr>
<td>R1 (W)</td>
<td>31.7</td>
</tr>
<tr>
<td>7.25 x 10^12</td>
<td></td>
</tr>
<tr>
<td>Phase Noise @ 1kHz (dBc/Hz)</td>
<td>-89.3</td>
</tr>
<tr>
<td>1048 MHz VCO</td>
<td></td>
</tr>
</tbody>
</table>

Fig.11. Plot of measured and simulated phase noise for 920 MHz VCO.

Fig.12. Tuning range for 920 MHz VCO with independent variations of varactor control voltages.

V. Summary and Future Work

813 MHz, 920 MHz and 1.048 GHz VCOs using quartz MEMS resonators have been demonstrated and characterized in terms of phase noise and tuning range. With Vcc =2 V and Iref = 4mA, the 813 MHz VCO provides the lowest phase noise of ~100 dBc/Hz at 1 kHz offset. Output power under this bias condition is measured to be -4.5 dBm for the 813 MHz VCO and -11dBm for 1.048 GHz VCO. A tuning range of approximately 80 ppm is achieved. According to noise simulations, the elimination of the p-n-p transistor in the push-pull core and replacing the CMOS devices in current source with SiGe bipolar transistors may further improve phase noise. Future IC fabrication runs will include standard Pierce designs without p-n-p transistors and CMOS current sources. Also, hyperabrupt varactor diodes may be used in place of MOS varactors because they can be configured to use a single positive supply between 0 and 3V that is commensurate with
the voltage output of our external temperature compensation circuitry.

ACKNOWLEDGMENT

This work was sponsored by DARPA'S MTO office under the DEFYS program, issued by DARPA /CMO under contract No. HR0011-10-C-0109. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of DARPA or the U.S. Government.

REFERENCES

Ultra-low phase noise frequency synthesis chains for High-performance vapor cell atomic clocks

B. François and R. Boudot
FEMTO-ST CNRS
Université de Franche-Comté
26 chemin de l’Épitaphe
25030 Besançon
France
Email: rodolphe.boudot@femto-st.fr

C. E. Calosso
INRIM
Strada delle Cacce 91
10135 Torino
Italy
Email: c.calosso@inrim.it

J. M. Danet
LNE-SYRTE
Observatoire de Paris
CNRS-UPMC
75014 Paris
France

Abstract—This paper presents the development of a 9.192 GHz microwave frequency synthesis chain (absolute and residual noise characterization) dedicated to be a local oscillator (LO) in a high-performance cesium vapor cell atomic clock based on coherent population trapping (CPT). The key components are an ultra low noise oven controlled quartz crystal oscillator (OCXO), multiplied to 9.2 GHz using a non-linear transmission line. A dielectric noise oven controlled quartz crystal oscillator (OCXO), multiplied population trapping (CPT). The key components are an ultra low performance cesium vapor cell atomic clock based on coherent characterization) dedicated to be a local oscillator (LO) in a high-mid microwave frequency synthesis chain (absolute and residual noise functions, the expected Dick effect contribution to the atomic clock short term frequency stability is reported to $6.2 \times 10^{-14}$ at 1 s integration time. Main limitations are pointed out.

I. INTRODUCTION

Over the last years, state-of-the-art laboratory vapor cell atomic clocks, based either on optical-microwave double resonance technique [1] or coherent population trapping (CPT) [2], have demonstrated short-term frequency stability in the $1 - 3 \times 10^{-13}$ range [3], [4], [5], [6], [7]. Such performances demonstrate the capacity of vapor cell atomic clocks to be competitive with bulky hydrogen masers regarding short and mid-term fractional frequency stability.

A major contribution to the limitation of the short term frequency stability of an atomic clock is the local oscillator (LO) phase noise through the inter-modulation effect in continuous wave (CW) regime clocks [8] or Dick effect [9], [10] in pulsed clocks.

The frequency stability limitation of an atomic clock due to the Dick effect is given by [10]:

$$\sigma_{y,LO}^2 (\tau) = \frac{1}{\tau} \sum_{i=1}^{\infty} g_i^2 \sum_{i/T_c}^{} g_0^2 S_y (2i f_m) \quad \text{(1)}$$

The parameters $g_i$ and $g_0$ are defined by the sensitivity function $g$ [10]. $T_c$ is the cycle time of the atomic clock ($f_c = 1/T_c$) and $i/T_c$ are the harmonics of the interrogation frequency. Recently, it was demonstrated that a CPT-based atomic clock exhibits a more important sensitivity to high frequencies of the LO phase noise [4].

In the frame of the EURAMET Mclocks project [11], aiming at developing high performance vapor cell atomic clocks exhibiting a short term frequency stability at the level of $10^{-13}$, this paper describes a 9.192 GHz synthesis chain for a CPT-based Cs atomic clock.

II. ARCHITECTURE OF THE SYNTHESIS CHAIN

Figure 1 shows a simplified schematic of the synthesis chain. The pilot and the key component of the chain is a 100 MHz ultra low noise OCXO (Pascall OCXOF-E-100) [12]. The output power of the OCXO is + 19 dBm. The 100 MHz signal is power-split into two arms.

In the first arm, reference signals at 10 MHz (using a divider by 10 Zarlinks SP8401) are generated.

In the second arm, the signal is frequency-doubled to 200 MHz using a passive frequency doubler (Minicircuits FD2+). The output at 200 MHz is band-pass filtered, amplified to a level of 23 dBm using a power amplifier (ZFL-1000VH2+). The latter signal is used to drive a Non-Linear Transmission Line (NLTL, comb generator up to 20 GHz). The harmonic at 9.2 GHz is selected and band-pass filtered using a 50 MHz bandwidth microwave filter and then amplified to a level of 0 dBm using a low-noise microwave amplifier (AML612L2201). Microwave isolators and attenuators are added at the output/input of the NLTL components for adjusting the impedance matching in order to optimize the phase noise of the microwave signal [13].

The 9.2 GHz signal is mixed with a 9.192 GHz signal generated by a dielectric resonator oscillator (DRO MITEQ - DRO-G-09192-MT±140) to produce a frequency beatnote of 7.368 MHz. The latter is low-pass filtered and amplified using a high-isolation radio-frequency amplifier (UTC-573) and then compared to a 7.368 MHz signal generated by a DDS (Agilent 33220A). The DDS is driven by a 10 MHz signal. The DRO is phase-locked with a bandwidth of about 1 MHz.

The resulting signal is filtered and processed in a proportional integral (PI) controller. The operating point of the OCXO is set by summing previous signal with a bias voltage. High speed operational amplifiers (THS4011) are used to ensure proper and high-bandwidth lock of the DRO. A low phase

978-1-4799-8866-2/15/$31.00 ©2015 IEEE
noise 9.192 GHz signal with a power of 13 dBm is obtained at the output of the DRO. The signal is then split in two arms to produce 9.192 GHz or 4.596 GHz signals.

![Schematic of the synthesis chain architecture](image)

Fig. 1. Simple schematic of the synthesis chain architecture.

III. ABSOLUTE PHASE NOISE MEASUREMENTS

Measurements are performed with a signal source analyzer (Agilent E5052B) [14], [15]. Figure 2 plots the phase noise of key signals of the chain at 100 MHz, 200 MHz, 9.192 GHz (free-running DRO and phase-locked DRO) and the phase noise of the current LO used in SYRTE at 9.392 GHz. The phase noise spectrum of the 100 MHz is given in dBrad²/Hz by the power law \( S_b(f) = \sum_{i=0}^{4} b_i f^i \) with \( b_0 = -182, b_{-1} = -130 \) and \( b_{-2} = -82 \).

For \( f < 10 \text{ kHz} \), the 200 MHz signal phase noise is as expected 6 dB higher than the 100 MHz signal phase noise. For higher frequencies than 10 kHz, a degradation induced by the frequency doubler is measured regarding to the expected level of the 100 MHz signal ideally multiplied.

In free-running regime, the DRO exhibits a phase noise at the level of \(-83\) and \(-145 \text{ dBrad²/Hz} \) at 10 kHz and 1 MHz respectively. Its phase noise floor is measured to be \(-172 \text{ dBrad²/Hz} \) at \( f > 10 \text{ MHz} \). The phase locked DRO exhibits phase noise of \(-42, -100, -129 \) and \(-130 \text{ dBrad²/Hz} \) at \( f = 1 \text{ Hz}, 100 \text{ Hz}, 10 \text{ kHz} \) and 1 MHz respectively. The 4.596 GHz signal phase noise has been measured to be in good agreement with the theory and is 6 dB lower than the 9.192 GHz phase noise signal in locked regime.

IV. RESIDUAL PHASE NOISE AND LIMITATIONS

Residual phase noise characterization of key components has been realized to identify main limitations of the system. Figure 3 reports the residual phase noise of the frequency doubler and the NLTL-chain. All signals are compared to the 100 MHz phase noise in order to point out the limitations of the synthesis chain. Phase noise measurements are reported to 9.2 GHz. The synthesis chain noise is limited by the frequency doubler for \( f > 10 \text{ kHz} \) and by the NLTL-chain for 100 Hz \(< f < 1 \text{ kHz} \).

![Phase noise measurements](image)

Fig. 2. Absolute phase noise of key signals: a) 100 MHz OCXO, b) 200 MHz, c) free-running DRO, d) DRO phase-locked and e) current microwave 9.392 GHz signal used at LNE-SYRTE.

![Phase noise measurements](image)

Fig. 3. Phase noise signal reported at 9.2 GHz: a) absolute phase noise of the OCXO, (b) residual noise of the NLTL-chain and (c) residual noise of the frequency doubler.

V. DICK EFFECT CONTRIBUTION TO THE SHORT TERM FREQUENCY STABILITY

Figure 4(a) reports absolute phase noise performances of the microwave signal used to probe the atoms in a current state-of-the-art CPT clock developed in LNE-SYRTE and our present synthesis chain. Figure 4(b) shows the corresponding Dick effect contribution to the CPT-based clock short-term frequency stability in both cases. The latter has been estimated thanks to the use of Equation 1. Operating parameters and conditions of the pulsed sequence are reported in [4]. The Dick effect contribution is plotted as a function of the noise
integration bandwidth. Once this integrated noise stops to contribute to the Dick effect, $\sigma_y$ reaches a floor that is the total Dick effect contribution.

The Dick effect contribution has been estimated to $2.7 \times 10^{-13}$ at 1 s for the synthesis chain use in SYRTE and $6.2 \times 10^{-14}$ for the LO presented in this paper.

VI. CONCLUSION

We reported the development of an ultra low noise synthesis chain dedicated to the interrogation of the Cs atoms in a CPT-based vapor cell atomic clock. Regarding to the synthesis chain currently used in a state-of-the-art CPT clock developed in LNE-SYRTE, an improvement of $4-13$ dB is noticed for $f > 100$ Hz. Thanks to the phase noise improvement, the Dick effect contribution to the short term frequency stability of the clock could be reduced from $2.7 \times 10^{-13}$ to $6.4 \times 10^{-14}$.

ACKNOWLEDGMENT

This work was partly supported by LNE, LabEx FIRST-TF and EquipX Osc-Imp.

REFERENCES

Abstract—In this paper, we introduce a novel design length extension vibratory gyroscope to detect the angular velocity rotation about z-axis. The proposed gyroscope is a new type of a gyroscope which utilizes a length extension mode as a driving mode and a flexure mode as a sensing mode to detect the Coriolis force generated by the rotation of the system. The gyroscope was designed and gyro-characteristics were simulated using COMSOL, finite element method (FEM) software. Quartz and langatate crystals are used for gyroscopes and compared. The driving frequencies and sensing frequencies of each gyroscope are obtained by optimizing the geometries of each gyroscope using eigenfrequency analyses. Frequency response analyses were performed to simulate the gyro-characteristics of the gyroscopes which subjected to the angular velocity about z-axis. The results show that the length extension gyroscope can be used as a gyro-sensor. Moreover, we find that langatate crystals are more suitable materials for higher precision piezoelectric gyro-sensors than quartz crystal.

Keyword—Angular rate sensor, Finite element method, Quartz, Langatate, Length extension gyroscope, Piezoelectric vibratory gyroscope.

I. INTRODUCTION

The purpose of this study is to simulate and analyze a newly designed length extension gyroscope structure. The length extension gyroscope utilizes the first length extension mode of the drive arms of its sensor to detect the Coriolis force applied on the gyroscope. Therefore, it is referred to as a length extension gyroscope. The gyroscope is made of a single piezoelectric crystal. It is a flat and very small structure which can be easily mounted and manufactured.

Currently, the piezoelectric gyroscopes are widely used in many different industries such as aerospace, automobile, entertainment and consumer electronics industries. These gyroscopes need to overcome many technical issues, such as miniaturization, low sensitivity, complex geometry, and shock resistance. The length extension gyroscope on the other hand has very simple design with simple driving and sensing electrodes. Moreover, the gyroscope is very reliable for providing shock resistance since it is made from a single piezoelectric crystal.

The piezoelectric materials which are used in this study are quartz (SiO₂) and langatate (La₃Ga₅.5Ta₀.5O₁₄). The quartz crystals have been widely used in bulk acoustic wave (BAW) applications, whereas langatate is fairly new material for BAW applications which is fast gaining in popularity. Quartz and langatate are trigonal class crystals and belong to point group 32. Therefore, both materials are often compared. The langatate crystal has higher mass density and piezoelectric coefficient than a quartz crystal which makes langatate a more attractive material for vibratory gyroscope applications. The piezoelectric constants of each material are listed in table I. [1, 2, 3]

In this paper, we have used COMSOL 4.3a, the finite element method (FEM) software, to design and analyze the length extension gyroscope. In the first part of this paper, we study the operation of principle and geometric characteristics of the length extension gyroscopes. In the second part of this paper, we simulate and compare the gyroscopic sensitivity of both materials.

II. VERIFICATION AND VALIDATION OF THE SIMULATION

We have verified and validated our finite element model in a previous paper[4] with the experimental data by Sato et. al., [5, 6, 7] gathered using the doubled ended quartz tuning fork gyroscope. The eigenfrequency analyses and frequency domain analyses were performed to simulate the geometric characteristics and gyro-characteristics of the quartz doubled ended tuning fork rotation about x-, y- and z-axis. The driving frequency of the gyroscope was 30733.7 Hz and the sensing frequency rotation about z-axis was 30740.0 Hz. The gyroscope was optimized to detect the angular velocity rotation about z-axis. Hence, the sensitivity of the gyroscope rotation about z-axis.

<table>
<thead>
<tr>
<th>Property</th>
<th>Quartz</th>
<th>LGT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic constants (10⁹ N/m²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c₁₁ = c₂₂</td>
<td>86.74</td>
<td>189.4</td>
</tr>
<tr>
<td>c₁₂</td>
<td>6.98</td>
<td>108.4</td>
</tr>
<tr>
<td>c₁₃ = c₂₃</td>
<td>11.91</td>
<td>132</td>
</tr>
<tr>
<td>c₁₄ = -c₂₄ = c₅₆</td>
<td>-17.91</td>
<td>13.7</td>
</tr>
<tr>
<td>c₃₃</td>
<td>107.2</td>
<td>262.9</td>
</tr>
<tr>
<td>c₄₄ = c₅₅</td>
<td>57.94</td>
<td>51.25</td>
</tr>
<tr>
<td>c₆₆</td>
<td>39.88</td>
<td>40.52</td>
</tr>
<tr>
<td>Viscosity Constants (10⁻³ kg/m*s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c₁₁ = c₂₂</td>
<td>1.370</td>
<td>0.490</td>
</tr>
<tr>
<td>c₁₂</td>
<td>0.730</td>
<td>0.290</td>
</tr>
<tr>
<td>c₁₃ = c₂₃</td>
<td>0.710</td>
<td>0.184</td>
</tr>
<tr>
<td>c₄₄ = -c₂₄ = c₅₆</td>
<td>0.010</td>
<td>0.093</td>
</tr>
<tr>
<td>c₃₃</td>
<td>0.960</td>
<td>0.320</td>
</tr>
<tr>
<td>c₄₄ = c₅₅</td>
<td>0.360</td>
<td>0.185</td>
</tr>
<tr>
<td>c₆₆</td>
<td>0.320</td>
<td>0.100</td>
</tr>
<tr>
<td>Piezoelectric Coefficient (C/m²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c₁₁ = -c₁₂ = -c₂₆</td>
<td>0.171</td>
<td>-0.54</td>
</tr>
<tr>
<td>c₁₄ = c₂₅</td>
<td>-0.0407</td>
<td>0.07</td>
</tr>
<tr>
<td>Relative Dielectric Constants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eₚ₁₁= Eₚ₂₂</td>
<td>4.429</td>
<td>19.0</td>
</tr>
<tr>
<td>Eₚ₃₃</td>
<td>4.464</td>
<td>52.0</td>
</tr>
<tr>
<td>Mass density (Kg/m³)</td>
<td>2651</td>
<td>6126</td>
</tr>
</tbody>
</table>

Table I
y-axis was expected to be very small compared to the sensitivity rotation about z-axis. The sensitivity of the x-axis rotation was expected to be zero, since the vibration velocity is parallel to the angular velocity. The calculated $Q$ of the gyroscope was 218000. The sensitivity of the gyroscope also depended on the level of damping of the system. The two types of damping in this system were material damping and external damping. In this case, there was no external damping since we assumed that system was in vacuum condition. The material damping was dependent on the driving frequency. Higher driving frequency will increase material damping of the system. The simulation results are indicated as solid lines while experimental data are indicated as dash lines as shown in Fig.1. The gyro-sensitivity, $S$, of the gyroscopes rotation about x-, y-, and z-axis are $S_x$=-0.159µV/(deg/s), $S_y$=-1.49mV/(deg/s) and $S_z$=4.98mV/(deg/s), respectively. The measured sensitivities of the gyroscope from the experiment are $S_x$=-0.04mV/(deg/s), $S_y$=-1.48mV/(deg/s) and $S_z$=+5.54mV/(deg/s). Considering the amplification factors of the gyroscope in the experimental data, the simulation results are well matched to the experimental data. Therefore, the results of the simulated models when compared to the experimental data are reliable and can be used to design the piezoelectric gyroscope.

III. LENGTH EXTENSION VIBRATORY GYROSCOPE

The structure of the length extension vibratory gyroscope is shown in Fig.2 (a). The gyroscope consists of the driving part, the sensing parts, the connection parts, and the base.

A. Driving part

The driving part of the length extension gyroscope is located at the middle of the sensing arms, as shown in Fig.2 (a). In consideration of the crystallographic axis and electric polarization direction of each crystal, the electrodes for driving parts are placed on the sidewalls (y-z plane) of driving arms so that the driving mode shape is identical to the resonant frequency mode shape, the first length extension mode. The ±1 V is applied on one side of the electrodes whereas the opposite sides of the electrodes are grounded. The main advantage of using a length extension mode as driving mode is energy dissipation. The length extension mode is the symmetric mode that has zero displacement node at the center of the driving arms which has very small influence by the support condition. As a result, a high Q-factor of the gyroscope can be achieved.

B. Sensing part

The sensing part is the part that connects the driving arms and the support part while it measures the Coriolis force signals of the gyroscope. The sensing bridge is located at the end of the supporting arms, as shown in Fig. 2(a). Similar to the driving arms, the electrodes on the sensing bridge is located on the sidewalls (x-y plane) of the sensing bridge. The electrodes are placed so that the resonant frequency mode shape, the flexure mode, is similar to the sensing mode shape. In consideration of the crystallographic axis and electric polarization direction of each crystal, the sensing electrodes, which measure and output the Coriolis force signal of a gyroscope, are placed on one side of the sidewalls and the electrodes on the other side is grounded. The result is that higher sensitivity of the gyroscope can be achieved.

C. Support part and base

The supporting part is located between the base and sensing part, as shown in Fig. 2(a). The length of the support part, $L_2$, is determined by the length of driving arm, $L_D$. The small gap between the end of the driving arm and the supporting part is adequate, since the displacement of the driving arms is insignificant. The width of the length, $W_s$, has to be long enough so that displacement of the support part along the x-axis is minimized. The end of the base is fixed to the fixed end conditions.
IV. PRINCIPLE OF OPERATION OF A LENGTH EXTENSION GYROSCOPE

The principle of operation of a length extension gyroscope is shown in Fig.3(a). The two vibration modes are used respectively as driving mode and sensing mode. The sensing mode is used to detect the angular velocity rotation about z-axis. The driving mode is a first extension mode of driving arm in the x-y plane. The driving arm expands and contracts along y-axis by the external harmonic excitation force, as shown in Fig.3(b). The midpoint of the driving arm is the zero displacement node. The second vibration mode is the z-axis detection mode that is an anti-symmetric flexure mode in x-y plane, as shown in Fig.3(c). When the gyroscope is subjected to a rotation about z-axis, a pair of Coriolis forces, \( F_c \), on the driving arms is generated proportional to the angular rate of gyroscope, \( \Omega_z \), mass density, \( m \), and vibration velocity, \( V_y \), of the driving mode in Fig.3(b). A pair of equal and opposite Coriolis forces on each driving arm, which are generated by the rotation about z-axis, create the moment, \( M_c \), at the center of the driving arms. The moment, \( M_c \), induces the flexure mode of the sensing arms. The z-axis detection electrodes on the sensing arm measure the Coriolis force signals. The driving frequency, \( f_d \), and sensing frequency, \( f_s \), have to be tuned close to each other in order to achieve higher gyroscopic sensitivity.

<table>
<thead>
<tr>
<th>Conditions of FEM simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material</td>
</tr>
<tr>
<td># of elements</td>
</tr>
<tr>
<td># of degree of freedom</td>
</tr>
</tbody>
</table>

V. SIMULATION OF LENGTH EXTENSION GYROSCOPES

In this paper, we simulated and analyzed the length extension gyroscopes of two different materials: quartz and langatate. Both materials have the same crystal symmetry and therefore they can be easily compared. The eigen-frequency analysis and frequency domain analysis were performed to study the geometric characteristics and gyro-characteristics of each gyroscope. The following equations (1) to (4) were implemented in COMSOL finite element method software. Eqn.(1) is the strain-displacement relation and eqn.(2) is the electric field-potential relation:

\[
S_{ij} = \frac{1}{2} (u_{j,i} + u_{i,j}) 
\]

\[
E_i = -\Phi_i
\]

where \( S_{ij} \) is the mechanical strain, \( u_i \) is the mechanical displacements, and \( E_i \) is the electrical field and \( \Phi_i \) is the electric potential.

\[
T_K = -\varepsilon_{kk} E_k + c_e E^2_J S_J
\]
Equations (3) and (4) are the piezoelectric constitutive equations in stress-charge form. \( T_{ik}, e_{ik}, c_{ij}^p, D_i, \epsilon_{ik}^p \) are the mechanical stress, piezoelectric constants, linear elastic constants, electric displacements, and dielectric permittivity, respectively. [8, pp.129-130]

Figure 4 shows the FEM model of the proposed gyroscope. The 10 layers of the triangle shaped elements were used in mesh for all three materials. The total number of elements and degrees of freedom which were used in FEM are listed in table I. The drive and sense electrodes were also included in the FEM simulations, which are shown in Fig.5. All the electrodes are coated with 0.2\( \mu \)m thickness gold film.

In this study, we have set the target driving frequency of the gyroscopes to be around 1 MHz. The gyroscope used a length extension mode as a driving mode. The driving arm was symmetric along the x-axis. Hence, we considered a one-dimensional longitudinal vibration of elastic beam of fixed-free case to simplify the problem and estimate the resonant frequency. The equation of the motion for longitudinal vibration of elastic beam is

\[
m \frac{\partial^2 u}{\partial t^2} - EA \frac{\partial^2 u}{\partial x^2} = 0
\]

Applying the fixed-free boundary conditions to Eq.8, we can obtain the equation to calculate the frequency of a longitudinal vibration of beam when the structure is fixed at its one end.

\[
f = \frac{1}{2\pi} \sqrt{\frac{EA}{mL_D^2}} = \frac{1}{2\pi} \sqrt{\frac{E}{\rho}}
\]

The driving frequencies of both materials are dependent on the length of the driving arms, \( L_D \), elastic modulus, \( E \), and mass density. Since elastic modulus and mass density of both materials are constant, the driving frequencies depend on the length of the driving arms of the gyroscope. Therefore, the lengths of the drive arms of both gyroscopes are different; such that the length of the driving arms of langatate gyroscope is shorter than quartz gyroscope. Hence, the overall size of langatate gyroscope can be minimized. The eigenfrequency analyses were used to determine the length of the drive arms as well as overall size of each gyroscope. The specific sizes of each gyroscope are listed in Table 2.

<table>
<thead>
<tr>
<th>Dimension of the length extension gyroscope of quartz and langatate.</th>
<th>Quartz</th>
<th>Langatate</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W )</td>
<td>2.500</td>
<td>1.074</td>
</tr>
<tr>
<td>( W_0 )</td>
<td>1.258</td>
<td>0.574</td>
</tr>
<tr>
<td>( L )</td>
<td>3.000</td>
<td>2.000</td>
</tr>
<tr>
<td>( L_0 )</td>
<td>2.510</td>
<td>1.998</td>
</tr>
<tr>
<td>( L_2 )</td>
<td>1.400</td>
<td>0.300</td>
</tr>
<tr>
<td>( L_5 )</td>
<td>0.100</td>
<td>0.100</td>
</tr>
<tr>
<td>( t )</td>
<td>0.060</td>
<td>0.060</td>
</tr>
</tbody>
</table>

In this simulation, the cut angles of each crystal were different. It was realized that the z-cut crystals were not the optimum cut angle for all materials. Hence, we analyzed each material by rotating the geometry about the x-axis to find the geometry where the displacement along the z-axis of the driving arm was minimized. This not only reduced the out-of-plane vibration of the system but also reduced the energy dissipation that increased the sensitivities of the gyroscopes. The optimum cut angles of the quartz and langatate are respectively \( \theta = 9 \) and \( \theta = 24 \).

The sensitivity of the gyroscopes was dependent on the matching of the driving frequency to the sensing frequency. Therefore, parametric eigenfrequency analyses were performed to obtain the optimum length of drive arm of each material. Thus, the parameters which were used in this simulation were the length of the driving arms. The change in drive arm lengths not only affect the driving frequencies but also the sensing frequencies as shown in Figs.6 (a) and (b). The rates of change of the sensing frequencies were higher than driving frequencies. We found the optimum length of the drive arms of quartz gyroscope and langatate gyroscope were respectively 2.5 mm and 1.998 mm.

The driving and sensing frequencies could be adjusted by changing the length of the sensing arm, \( W_s \) or width of the sensing arm, \( L_s \). However, changing those two parameters to match the driving and sensing frequencies could not be accomplished due to the limitation of its geometry. Hence, it

![Fig.6(a) Frequency spectrum of the quartz gyroscope as a function of the length of the drive arm, \( L_D \).](image-a)

![Fig.6(b) Frequency spectrum of the langatate gyroscope as a function of the length the drive arm, \( L_D \).](image-b)
The mechanical coupling between the driving mode and sensing mode of the quartz gyroscope was very small compared to the langatate gyroscope, which made it easier to match the driving and sensing frequencies as close to each other as possible when the length of the driving arm was around 2.515 mm as shown in Fig. 6(a). Since the piezoelectric coupling factor of langatate was large the driving and sensing modes never cross each other as shown in Fig. 6(b).

We discussed previously that the matching of driving and sensing frequencies would increase the gyro-sensitivity of the gyroscope. However there was a caveat that the system became unstable when the driving frequency got too close to the sensing frequency. This forced us to require a minimum frequency separation of driving and sensing modes. The frequency separations of quartz and langatate gyroscopes were determined to be 29.4 Hz and 519.3 Hz respectively.

### A. Gyroscopic Sensitivities of the Length Extension Gyroscopes.

We further analyzed the length extension gyroscope with frequency response analyses using the optimized geometries of the gyroscopes of both materials. The equations of the motion of the element of the rotating system rotating about z-axis are implemented in FEM simulation and they are

\[
f_x = F_x + \rho \omega^2 u_x + j2\rho \Omega_x \omega u_y + \rho \Omega_x^2 x \tag{7a}
\]

\[
f_y = F_y + \rho \omega^2 u_y - j2\rho \Omega_x \omega u_x + \rho \Omega_x^2 y \tag{7b}
\]

\[
f_z = 0 \tag{7c}
\]

where \( f, F, u, \Omega \) and \( \omega \) are respectively the total force acting on the element, external forces, element displacement, and angular velocity and angular frequency.\[^{[9]}\] The \( x \) and \( y \) are the coordinates of the element. The first term and second terms of the right hand side of (7a) and (7b) are the external force and inertia forces induced by vibration. The third terms on the right hand side of (7a) and (7b) are the Coriolis force terms that we are interested in this study. Finally, the last terms of the equations are the centrifugal forces. The Coriolis forces are applied as body forces, \( F_b \), on the entire gyroscopes in frequency response analyses to simulate the gyroscope rotation about z-axis and they are

\[
F_{bx} = 2\rho \omega \Omega_x u_y \tag{8}
\]

\[
F_{by} = -2\rho \omega \Omega_x u_x \tag{9}
\]

\[
F_{bz} = 0 \tag{10}
\]

where \( F_{bx}, F_{by} \) and \( F_{bz} \) are respectively the body forces applied to the entire structure in \( x, y, \) and \( z \)-directions.\[^{[9]}\] Since the vibration direction of the drive arms of the gyroscope is parallel to the angular velocity, \( z \)-axis, the Coriolis force in the \( z \)-direction is zero. Similarly, we can also apply the Coriolis forces for rotation about \( x \)- and \( y \)-axis. The absolute value of charge on the driving and sensing electrodes as a function of the driving frequency, \( f_{dr} \), are shown in Fig. 7.

The gyroscopic response of each gyroscope to angular velocity of 60 degrees/s about the \( x \)-, \( y \)- and \( z \)-axis are shown in Figs. 8. These results show that the charges are the maximum at the resonance frequency, \( f_{dr} \), of the driving mode where the sensitivity at the resonance will get amplified the most. The sensitivities of each gyroscope were calculated by frequency response analyses. The electric potential was used as output.
The gyro-sensitivities of the gyroscopes of both materials are shown in Fig.9. The sensitivities of the quartz and langatate gyroscopes are $S_x = 2.78 \times 10^{-4} \text{V/(deg/s)}$ and $S_z = -5.76 \times 10^{-3} \text{V/(deg/s)}$. The sensitivity of the quartz gyroscope was calculated to be very small as expected. The sensitivities of the langatate gyroscope were about 20 times more sensitive than the quartz gyroscope. We have simulated the gyro-sensitivities of the gyroscopes rotation about the x-and y-axis. Both gyro-sensitivities were very small as expected. The quality factor, Q, was also calculated. The summary of the gyro-sensitivity of the gyroscopes are listed in Table III. The results clearly show that the length extension gyroscope performed well as a gyro-sensor.

<table>
<thead>
<tr>
<th>Units</th>
<th>Quartz</th>
<th>Langatate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{dr.}$ Hz</td>
<td>1034888</td>
<td>960028.3</td>
</tr>
<tr>
<td>$f_{sens.}$ Hz</td>
<td>1034854</td>
<td>960547.6</td>
</tr>
<tr>
<td>$S_x$ V/(deg/s)</td>
<td>8.46e-6</td>
<td>2.18e-4</td>
</tr>
<tr>
<td>$S_y$ V/(deg/s)</td>
<td>1.70e-6</td>
<td>-1.12e-4</td>
</tr>
<tr>
<td>$S_z$ V/(deg/s)</td>
<td>2.78e-4</td>
<td>-5.76e-3</td>
</tr>
<tr>
<td>Q</td>
<td>1</td>
<td>230048</td>
</tr>
</tbody>
</table>

VI. SUMMARY AND CONCLUSIONS

A novel design of the length extension gyroscope was proposed and analyzed using COMSOL finite element method software. Our model for analyzing and designing the gyroscope was validated by experimental data. The calculated optimum cut angle for a singly rotated quartz crystal was $\theta = -9$, and for a singly rotated langatate crystal was $\theta = 24$. The sizes of the gyroscopes varied due to the material properties and length of the driving arms. The overall size of the quartz gyroscope was larger. The langatate gyroscope was smaller than the quartz. The gyro-sensitivity rotations about z-axis of quartz and langatate gyroscopes were $2.78 \times 10^{-4} \text{V/(deg/s)}$ and $-1.6 \times 10^{-2} \text{V/(deg/s)}$.

The newly designed length extension gyroscope could be used as a gyro-sensor. The langatate gyroscope provided stronger sensitivity to angular velocity than the quartz gyroscope. The gyro-sensitivities of the gyroscopes were a function of the geometry of the gyroscope. The length extension gyroscope was able to detect the angular velocities about the other two axes (x- and y-axes).

REFERENCE

Remote Atomic Vapor Magnetometer with Sub-pT Resolution Operating at Ambient Temperature

Janet W. Lou
C4ISR Division
Sotera Defense Solutions, Inc.
Columbia, MD

Fredrik K. Fatemi and Geoffrey A. Cranch
Optical Sciences Division
U.S. Naval Research Laboratory
Washington, DC

Abstract—We have designed and characterized a Cesium vapor cell magnetometry system specifically aimed at meeting the challenges of remote interrogation. Standard single-mode telecom optical fibers are used for optical delivery and collection. The additional noise contribution due to optical polarization drift during propagation is mitigated by polarization and modal walk-off. We show that it is possible to obtain as low as 0.1 pT/Hz1/2 resolution at 1 Hz without temperature control over the ambient temperature range of -5 to 30 C.

Keywords—atomic vapor magnetometer, remote operation, vapor cell

I. INTRODUCTION

Atomic vapor magnetometers are of interest for a range of applications from medicine to geological and military surveying. Because they are compatible with all-optical techniques, these systems offer several advantages for remote underwater applications in addition to high sensitivity. Optical fiber remoting is a natural extension of optical pumping and interrogation schemes and offers high reliability due to the elimination of electronics in the water.

For several reasons, it is desirable to operate the magnetometers at the ambient temperature of the environment. A system requiring a fixed operating temperature may be prone to long-term degradation due to component drifts. Temperature control is often achieved using electrical current and resistive elements. This not only requires electrical power at the remote location, but the electromagnetic field induced by the current in the heating element may be a major contributor to both short-term noise and long-term drift. Short-term noise poses a limitation on the resolution of the magnetic field measurement. High resolution magnetometers at ~1 Hz would enable detection of small changes in the magnetic field on the order of seconds. Long-term variations in the measurements degrade the accuracy with which the background magnetic field can be characterized and limit the ability to detect small changes within that background.

We have designed and characterized a magnetometer specifically aimed at meeting the challenges of a remote system operating without thermal control. Standard single-mode telecom optical fibers (>1.5 km) are used for the delivery and collection of the optical signals. In addition to the degradation due to optical losses (~2-3 dB/km), the fact that the optical polarization is not maintained in these fibers poses an even greater challenge. However, polarization and modal walk-off can be used to mitigate the additional noise contribution. The system operates without temperature control and high resolution can be achieved at the ambient temperature range of -5 to 30 C.

II. CESIUM VAPOR CELL CHARACTERIZATION

A. Lifetime Measurement Setup

The shot noise limited performance of the magnetometer is dependent on the spin polarization lifetime. Thus, we use a pump-probe setup, as shown in Fig. 1, to characterize the upper state lifetimes of various Cs vapor cells. All of the cells are cylindrical with 7.5 mm diameter and 2 cm length. Two lasers are frequency locked using the DAVLL technique [1]. The pump laser is locked to the D2 transition (S1/2 (F=4) → P3/2 (F=4)) and amplitude modulated by either an optical chopper or fast shutter. The probe laser is locked to the D1 transition (S1/2 (F=3) → P1/2 (F=4)). The circularly-polarized pump and linearly-polarized probe are spatially overlapped and pass through the length of the vapor cell. At the output, the pump is removed with a bandpass filter. Following the NMOR technique [2], the probe signal is sent through a Wollaston prism and the two orthogonal polarizations are detected by a balanced amplified detector (ThorLabs PDB210A).

The vapor cells are characterized inside three layers of magnetic shielding, which reduces the ambient magnetic field to ~400 nT. A solenoid coil inside the chamber generates a controlled magnetic field parallel to the direction of beam propagation. The balanced output of the detector is recorded by an oscilloscope and the measured signal is numerically fit to the sum of two exponentials to obtain the relaxation time [3].

Fig. 1: Setup for characterization of Cs vapor cells. BPF=bandpass filter.
B. Measured Results

A range of vapor cells with different Ne buffer gas pressures is measured to determine the relaxation time [4]. We find that the fitted time constant can be affected by the precision of the initial balance between the two detectors. In order to get the range of values, we rotate the half-wave plate to each of the four “balance points” and record the relaxation signal. The fitted time constants are shown in Fig. 2. For pure Cs, the lifetime is only ~50 μs, dominated by dephasing caused by wall collisions. With the addition of Ne buffer gas, the lifetime increases as the diffusion time increases. For 10, 37.5, and 200 Torr Ne, the fitted time constants are approximately 1 ms, 2.5 ms, and 12 ms, respectively.

III. TWO-LASER MAGNETOMETER

The magnetometry system is a slight modification of the previously described pump-probe setup. The vapor cell is optically pumped at the D1 transition, and the pump laser amplitude is modulated at a fixed frequency with an acousto-optic modulator (AOM). The probe laser is locked to the D2 transition and operated continuous wave. The output from the balanced detector is sent into a lock-in amplifier and the modulation frequency of the AOM is used as the reference frequency. A resonance signal is detected from the lock-in amplifier when the modulation frequency matches the half-Larmor frequency of the corresponding magnetic field [5]. The Cs cell filled with Ne at 200 Torr is used because its long lifetime implies the lowest shot noise limit and thus, it is expected to be capable of the highest resolution.

A. Detection of the Magnetic Field

An example of the lock-in output is shown in Fig. 3. Fig. 3(a) shows the varying voltage (current) driving the solenoid coil, which effectively sweeps the magnitude of the magnetic field at 0.2 Hz. Figs. 3(b) and 3(c) are the in-phase and out-of-phase signals from the lock-in, respectively. The lock-in is set for a 1 ms time constant with 12 dB/octave rolloff. A signal is detected for the co-propagating as well as the counter-propagating magnetic field, differing by a polarity change in the in-phase signal.

B. Principle of Resolution Measurements

To further reduce the magnetic field noise inside the magnetic shielding, variations in the residual magnetic field are partially compensated to the resolution of the magnetometer with a feedback loop. A DC drive is applied to the solenoid coil to set the magnetic field at the zero crossing of the dispersive (in-phase) component of the signal, as demodulated by the lock-in. Then, the dispersive signal is used as the input to a feedback loop that drives the coil so as to maintain the magnetic field at its set bias point.

The intrinsic noise of the magnetometer is represented by the power spectral density (PSD) around the modulation tone. The balanced detector signal is sent through a low-noise amplifier with a 10 kHz to 100 kHz filter (SRS SR560). The demodulation of the signal is performed digitally with a Labview™ program. As can be seen in the example result in Fig. 4, we can measure the PSD down to ~2 Hz. The signal from the balanced detector has a white noise component, which results in the 20 dB/decade frequency dependence exemplified by the dashed red line.

---

The work is supported by an ONR 6.2 Base Program.
data acquisition program is insufficient to give reliable data. The multiple peaks that can be seen are due in part to the time window of the demodulation algorithm and also noise imposed by the function generator on the AOM. For the Cs gyromagnetic constant of ~3.5 Hz/nT and for a PSD of ~65 dB re Hz/Hz\(^{1/2}\) (at about 2 Hz), it means that it is possible to resolve down to ~0.16 pT/Hz\(^{1/2}\).

C. Measured Resolution Results

1) Temperature Dependence: To ensure that this system is robust for a range of temperatures, meaning that no temperature control of the vapor cell would be necessary, the resolution is measured as a function of the ambient temperature. The entire magnetic shielding and vapor cell assembly is cooled in a freezer. Once removed from the freezer, the assembly is allowed to return to room temperature on its own. During this time, a thermocouple inserted into the chamber records the ambient temperature while the Labview\(^{TM}\) program measures the resolution limit of the system. For higher than room temperature, warm air is blown into the chamber. Fig. 5 shows the measured resolution dependence on temperature for average pump power of ~250 \(\mu\)W and modulation frequency of 50 kHz with 50% duty cycle, and probe power of ~400 \(\mu\)W. Over the range from -5 to 30 \(\degree\)C, the resolution at 2 Hz improves from ~4 to ~0.15 pT/Hz\(^{1/2}\). As will be shown in the following section, higher pump power can improve the resolution.

2) Power Dependence: It is necessary to understand the effect of power on the resolution of the system because of the relatively high loss of the optical fibers. The resolution is measured over a range of two orders of magnitude of pump power and the results are shown in Fig. 6, for room temperature (23 \(\degree\)C), and pump and probe parameters as in Section III.C.1. Over this power range, the resolution at 2 Hz improves from ~7 to ~0.1 pT/Hz\(^{1/2}\). On the other hand, the resolution is not significantly affected by the probe power because of the high sensitivity of balanced detection and the use of a low-noise amplifier after the detector.

IV. REMOTE INTERROGATION

To demonstrate a practical remotely-interrogated system, the input and output signals are delivered through long lengths of commercially-available Corning SMF-28 DS fiber. The input fibers are 0.7 km and 1.6 km lengths for the probe and pump lasers, respectively. To facilitate coupling the output to optical fibers, the Wollaston prism is replaced with a polarizing beam splitter. Then, each beam is individually coupled into a 2.2 km length of fiber. The outputs of the collection fibers are sent to the balanced detector.

A. Use of Single-Mode Fiber

Optical polarization is not maintained during propagation in single-mode fibers and is subject to environmental perturbations. For a magnetometry system that inherently relies on fixed input polarizations, this causes enormous changes in power. Large fluctuations in power can translate to variations in the signal from the lock-in amplifier, making phase-locked-loop feedback difficult. Active polarization controllers may be used to compensate for the polarization drift. However, they are complex, expensive, and their use would most likely require electrical power at the remote location.

B. Addition of Multi-mode or Polarization-maintaining Fiber

Instead of trying to maintain a fixed polarization from SMF, we consider techniques to reduce power fluctuations by using optical fibers with additional degrees of freedom. Short lengths (2-5 m) of multi-mode (MM) fiber or polarization-maintaining (PM) fiber connected to the end of the SMF (i.e., entrance to the magnetometry system) are tested. We find that either fiber can be used to improve the performance of the magnetometer. With MM fiber, the polarization evolution of the modes as well as modal walk-off average out the power at the desired polarization axis. With PM fiber, fluctuations generally occur on a time scale much shorter than the resolution measurement time and we are continuing to investigate the mechanism by which this assists the magnetometer performance. MM input fibers require MM fibers at the outputs to capture all the signal power; thus
creating a MM fiber system. We choose to use the PM fiber in order to build a single-mode fiber system.

C. Resolution of Remote System

Fig. 7 shows the measured PSD of the remotely-interrogated magnetometer operating at room temperature. The average pump power is ~225 μW and the beam size is 2.5 mm full width at half maximum. The modulation frequency is 50 kHz with 67% duty cycle. The probe power is ~100 μW and the beam size is 2.3 mm full width at half maximum. The minimum detectable resolution is ~1 pT/Hz\(^{1/2}\) at 1 Hz.

![Graph showing measured power spectral density of the demodulated signal from the remotely-interrogated magnetometer.](image)

V. SUMMARY

We have developed a remotely-interrogated magnetometer and addressed several issues that potentially limit its performance. One of the major factors that can lead to long-term drift and measurement inaccuracies is the requirement for a fixed temperature vapor cell. However, we have shown that it is possible to achieve a resolution of 0.15 – 4 pT/Hz\(^{1/2}\) while operating at the ambient temperature range of -5 to 30 C. This magnetometer uses two lasers and a Cs vapor cell filled with Ne at 200 Torr pressure. Reducing noise in the function generator and using an amplifier with 10 dB better signal-to-noise ratio for the balanced detection would enable the resolution of the remote magnetometer to reach ~0.1 pT/Hz\(^{1/2}\) at 1 Hz at room temperature. Remote operation of the magnetometer despite polarization fluctuations in standard single-mode fibers is enabled by the use of a short length of polarization-maintaining fiber.

REFERENCES

Ultra high frequency attenuation limits in synthetic IIa diamond single crystal were studied. The investigations were provided by High-overtone Bulk Acoustic Resonator “Al/AlN/Mo/(100) diamond”. It was shown, that the change of Akhiezer’s attenuation regime to Landau-Rumer one takes place at \( f \approx 1 \) GHz. The experimental results on the quality parameter showed \( Qsf \approx 1.8 \times 10^{15} \) Hz for Akhiezer’s regime, and \( Qsf \approx 1.5 \times 10^{14} \) Hz for Landau-Rumer regime. As a consequence on a higher frequencies diamond single crystal has more preferred acoustic parameters including a lower acoustic attenuation in comparison with known acoustoelectronic materials.

**Keywords**—diamond; HBAR; acoustic attenuation; UHF; AlN, damage layer, Kikuchi lines

I. INTRODUCTION

In recent years there is a trend in a materials search for crystals with low acoustic damping at UHF. At such high frequencies (up to 10 GHz) there are widely used acoustic resonators with elastic energy concentrated in the substrate material. There are some previously studied High-overtone Bulk Acoustic Resonators, with different substrate materials: single crystal and fused quartz and silicon [1], sapphire [1, 2], YAG [3]. The search of the substrate material is a separate task: it is necessary to consider a lot of parameters, such as the availability of the crystal, the quality of processing, low attenuation of the acoustic waves at UHF, etc. Thus, the diamond crystal is a perspective material with largest bulk acoustic wave (BAW) velocities, with low damping, high radiation stability and thermal conductivity. For example, it was shown, that diamond based HBAR can operate at frequencies up to 20 GHz [4].

HBAR structures may be used as special device for the determination of thin films and substrates physical properties [5]. Features of energy trapping in HBAR were studied in [6]. The mechanism of phonon-phonon damping for relatively low frequencies (\( \omega \tau < 1 \)) was first studied in [13]. The attenuation mechanism has purely quantum effect and three phonon interactions should be taken into account. Further study of this high frequency regime was carried out in [12, 14-17], more detailed study phonon selection rules was carried out.

There are many other energy dissipation mechanisms in crystals and devices [18]. The influence of damage layer as a destruction of single crystalline subsurface layers was studied in [19], while the contribution of roughness losses in HBAR resonators was calculated in [21].

The aim of this work is to study the mechanisms of dissipation of acoustic energy in the BAW resonators based on synthetic diamond single crystal at UHF.

II. UHF ATTENUATION

A. Akhiezer attenuation

Limiting acoustic attenuation in HBAR can not be lower than the values of phonon-phonon attenuation. Let’s consider the Akhiezer attenuation regime, when \( \omega \tau > 1 \). The first ever approach of acoustic attenuation in solids was made in [7]. Here we will use the convenient expression for the calculation of acoustic attenuation \( a_{ph-ph} \) (dB/cm) and quality parameters \( Qsf \), which were obtained in [12]:

\[
a_{ph-ph} = \frac{8.688nCTV^2\gamma\tau}{\rho V_a^2 \left[1 + (2\pi f\tau)^2\right]} f^2, \tag{1}
\]

\[
Qsf = \frac{\rho V_a^2 \left[1 + (2\pi f\tau)^2\right]}{CTV^2\gamma\tau} = const. \tag{2}
\]

Here \( C_v \) is volumetric heat capacity, \( T \) is the ambient temperature, \( \rho \) is the density of the medium, \( \gamma \) is Grüneisen parameter of a studied crystal, and \( V_a \) is the BAW velocity. As one can see from (1), during Akhiezer attenuation regime
\( \alpha_{\text{ph-ph}} \sim f^2 \), while the quality parameter has no frequency dependence: \( Q \approx \text{const} \) [12].

**B. Landau-Rumer attenuation**

In Landau-Rumer regime, which takes place at \( \omega \tau > 1 \), the attenuation is purely quantum effect, and was first studied in [13]. The convenient expressions for \( \alpha_{\text{ph-ph}} \) and \( Q f \) were also obtained in [12]:

\[
\alpha_{\text{ph-ph}} = \frac{8.68 \pi^3 k T \gamma}{30 p V^\gamma} f, \quad (3)
\]

\[
Q f = \frac{30 \rho V^\gamma h^2}{4 \pi^3 k T} f. \quad (4)
\]

Here \( k \) and \( h \) are Boltzmann and Planck constants, respectively. Comparing (1, 2) with (3, 4), we can notice, that in Landau-Rumer regime attenuation increases slower than in Akhiezer regime. It results in quality parameter increase while increasing frequency, or quality factor \( Q \) has no frequency dependence at UHF. It was shown in [12], that the attenuation regime transformation for diamond takes place at approximately 50 MHz.

**C. Grüneisen parameter**

It is clear that the parameter and acoustic attenuation both significantly depends on the anharmonicity Grüneisen parameter \( \gamma \) (1-4). The accurate determination of such parameter is a separate task. As it was shown in [10], \( \gamma \) is a second-rank tensor and can be calculated for each acoustic mode \( i \) as phonon frequency shift due to applied stress:

\[
\gamma^{(i)}_{jk} = - \left[ \frac{\partial \omega^{(i)}_{jk}}{\partial \sigma_{jk}} / \omega^{(i)} \right] \]

Here \( \sigma_{ij} \) is stress tensor. In order to get proper data for the studied acoustic mode, the results on attenuation and quality parameter (1-4) should be modified by replacing \( \gamma \) with \( \gamma^{(i)}_{jk} \). It was shown in [20], that there is a number of equations for Grüneisen parameter determination. Such variety of equations for \( \gamma \) determination results in uncertainty of its value: for diamond crystal one can obtain values from 0.74 up to 1.5, and for Grüneisen parameter determination. Such variety of equations was shown in [20].

\[
\alpha_{\text{ph-ph}} = 2 \pi \cdot 8.68 (k^2 \eta_{\text{AlN}}^2 + k^2 \eta_{\text{diam}}^2) f^2 \quad \text{MHz.}
\]

**D. Roughness losses**

In addition to phonon scattering losses, many other dissipation sources are also should be taken into account: attenuation in piezoelectric layer and electrode materials, elastic energy dissipation due to surfaces roughness and damage layer, boundary effects, etc. For example, the contribution of roughness losses on diamond and AlN surfaces can be calculated as [21]:

\[
\alpha_{\text{ph-ph}} \sim \omega^2 \eta_{\text{AlN}}^2 + \omega^2 \eta_{\text{diam}}^2 f^2 \quad \text{MHz.}
\]

Here \( \eta_{\text{AlN}} \), \( \eta_{\text{diam}} \), \( \eta_{\text{AlN}} \), \( \eta_{\text{diam}} \) are wave vectors and rms roughness values for AlN and diamond, respectively. As one can see from (6), roughness losses \( \alpha_{\text{ph-ph}} \) have squared frequency dependence. As \( \alpha_{\text{ph-ph}} \) in Landau-Rumer regime has linear frequency dependence, it is rather easy to separate phonon-phonon and roughness attenuations: removing squared frequency dependence in attenuation one can obtain \( \alpha_{\text{ph-ph}} \). Precise study of acoustic energy dissipation on roughness of diamond surface was provided in [20]. It was shown, that the roughness attenuation of (100) diamond substrate with \( R_e \leq 10 \text{ nm} \) is an order less than phonon-phonon scattering even at high frequencies such as 10 GHz. The main attention should be paid on the quality of AlN piezoelectric film.

Because HBAR has a sandwich structure, the attenuation in all layers should be taken into account. However, the thickness of metal electrodes and piezoelectric film are much lower than the thickness of a diamond substrate, and their contribution on total acoustic attenuation can be neglected.

**III. HBAR PREPARATION**

**A. Materials quality**

The substrate materials were thin plates of IIa synthetic diamond single crystal with [100] orientation grown by HPHT method at FSBI TESNCM (Moscow, Troitsk, Russian Federation). All the specimens were double-sized polished, faces were prepared with better than 5° deviation from [100] crystalline direction, were polished to obtain the roughness \( R_e \leq 10 \text{ nm} \). The typical HBAR had sandwich structure, as in Fig. 1. The thickness of polished diamond substrates, used in this study, varies from 0.4 mm up to 4 mm.

In order to control the depth of damage layer in diamond single crystal, the study of Kikuchi lines was produced. The presence of Kikuchi lines during experiments on Electron Back-Scattering Diffraction (EBSD) results in low values for damage layer depth [23]. The EBSD study of polished (100) diamond substrates was carried out on FEI Quanta 200 environmental SEM with EDAX microanalysis at Moscow Institute of Physics and Technology (Dolgoprudny, Russian Federation). A series of diamond plates with different polishing levels was studied. The energy of electron beam varied from 3 up to 30 keV. It was shown, that the damage layer for all studied crystals regardless of the polishing level is lower than 30 nm [20]. The presence of Kikuchi lines is shown on Fig. 2, presenting the studied crystal which was polished by a diamond powder 2/1 with weight load up to 300 g.
Fig. 1. Typical HBAR structure: 1, 3 are top and bottom electrodes, produced of Al and Mo, respectively; 2 is piezoelectric film (AlN); 4 is diamond substrate.

Fig. 2. The observation of Kikucki lines during 3 keV EBSD in (100) synthetic diamond single crystal.

Such results mean that the thickness of damage layer in synthetic diamond single crystal is extremely small. So, the influence of acoustic energy dissipation in this subsurface layer on total attenuation is insignificant. For example, as was shown in [23], the damage layer for sapphire crystal was sometimes hundreds nm, and a contribution into total acoustic attenuation should be taken into account.

B. Sputtering technique

The conducting electrodes and piezoelectric AlN layers were produced with magnetron sputtering technique. Studied HBARs were sintering by means of AJA ORION 8 magnetron rf sputtering equipment intended for precise deposition of conducting and isolating thin films. First, thick Mo conducting layer was deposited on polished and cleaned diamond surface. Second, synthesis of AlN piezoelectric film deposited on conducting layer was produced in the flow of Ar/N₂ gas mixture. To achieve almost pure longitudinal mode excitation, the orientation of sputtered AlN film was controlled to be (00 ⋅ 2). Finally, top electrode of Al in a required configuration was formed. In order to have a precise study of a crystal, a series of resonators with various top electrode configuration and size was produced on the same studied crystal. A typical studied HBAR with a number of resonators is represented on Fig. 3. Special attention was paid to the quality of AlN piezoelectric film. The roughness of aluminum nitride film was \( R_s \approx 10 \text{ nm} \) and main orientation was (00-2). For more detailed information see [24].

As the thickness of metal electrodes was up to 200 nm, the electrical conductivity of thin film Al and Mo electrodes could differ from bulk one. The conductivity investigations were carried out by 4-point Van der Pauw method. The electrical conductivity of a thin molybdenum film nonlinearly depends on thickness and is many times greater than conductivity of a bulk material. It means that the quality of molybdenum thin electrode is not optimal, and future investigation on its quality and morphology will take place.

IV. EXPERIMENTAL RESULTS

Microwave acoustic properties were tested by E5071C Network Analyzer (Agilent Tech.) and M-150 Multipurpose Probing System (CASCADE Microtech.), operating at frequencies from 300 MHz up to 20 GHz (Fig. 4). The study of quality factor frequency dependence in a wide frequency band was provided by the reflection \( S_{11} \) coefficient measurements. The quality factor \( Q \) was calculated as

\[
Q = \frac{f_r}{\Delta f},
\]

where \( f_r \) is resonant frequency, and bandwidth \( \Delta f \) was measured at -3 dB level. In order to improve the quality of experimental results, quality factor \( Q \) was also measured as

\[
Q_{\tau_d} = \frac{1}{2} \frac{d\phi}{d\omega} = \pi \tau_d f_r.
\]

Here \( \tau_d \) is the group delay time, and \( \phi \) is the phase of \( S_{11} \). Both ways of \( Q \) determination show almost the same results, sometimes varying less than 10\%. The set of \( Q \) at different resonant frequencies was obtained for each of the studied HBARs.

The features of studied structure behavior were discussed in [25] and can be associated with AlN film resonant excitation. Maximums and minimums of \( Q \) should be associated with max and min of form-factor \( m \) (Fig. 5). The experimental results on frequency dependence of quality factor \( Q \) for different electrode configurations are represented on Fig. 6. There is a trend in decreasing of quality factor at frequency increasing, while at relatively high frequencies one can say that \( Q \) tend to be constant. The maximal values of \( Q \approx 35,000 \) were obtained at low frequencies \( \approx 0.7 \text{ GHz} \), and relatively high value \( Q \approx 10,000 \) was observed about 9-10 GHz.

Fig. 3. A set of studied HBARs with different electrode configuration and lateral size, based on “Al/AlN/Mo/(100) diamond” layered structure (164 nm/624 nm/169 nm/392 μm).
Fig. 4. Experimental technique: 1 - Network Analyzer, 2 - Multipurpose Probing System, 3 – studied HBAR.

Fig. 5. Frequency dependence of form-factor $m$ in “AlN/AlN/Mo/(100) diamond” layered structure.

Fig. 6. Experimental results on the frequency dependence of quality factor $Q$ for HBARs based on “AlN/AlN/Mo/(100) diamond” layered structure with pentagon electrodes of different area: 1a – 40000 $\mu$m$^2$, 1b – 22500 $\mu$m$^2$, 1c – 10000 $\mu$m$^2$.

Fig. 7. Experimental results on the frequency dependence of $Q\times f$ quality parameter for HBARs based on “AlN/AlN/Mo/(100) diamond” layered structure with pentagon electrodes of different area: 1a – 40000 $\mu$m$^2$, 1b – 22500 $\mu$m$^2$, 1c – 10000 $\mu$m$^2$. 


In order to compare experimental results with calculated values for different attenuation regimes, it is convenient to operate with \( Q_\mu f \) quality parameter instead of quality factor \( Q \) or acoustic attenuation \( \alpha \). So, the results on quality parameter \( Q_\mu f \) are represented on Fig. 7. The best results on quality parameter for investigated resonators, based on “Al/AlN/Mo/(100) diamond” structure are about \( Q_\mu f \sim 9 \times 10^{13} \) Hz at both 5 and 9 GHz. In terms of acoustic attenuation this means that \( \alpha = 4.3 \) and 14 dB/cm at 5 and 9 GHz, respectively.

In order to get a proper data on diamond’s quality parameter and attenuation, only the best results should be taken into account. Also it should be noticed, that the location of minimums of \( Q_\mu f \) and maximums of \( \alpha \) are both associated with HBARs structure, depending on the Al, AlN, Mo thicknesses and so on. In order to obtain the proper experimental results on quality parameters, a series of HBARs with different electrodes and piezoelectric film widths and lengths was produced, and the study of quality factor frequency dependence was provided. The best results on quality factor for a set of different diamond based HBARs are represented on Fig. 8. The estimated results on diamond’s quality factor taken from [12] are also represented on Fig. 8.

To obtain pure material quality parameters the subtracting of roughness losses calculated from experimental data as in Eq. (6) was made. Analyzing Fig. 8, we can describe frequency dependence of \( Q_\mu f \) quality factor. From our experimental results there is no significant frequency dependence of quality parameter up to \( \sim 1 \) GHz, and \( Q_\mu f \) remains constant. This a classic behavior of quality factor for Akhiezer attenuation regime. Starting from 1 GHz there is a linear increasing of quality parameter, which is typical of Landau-Rumer attenuation regime. So, the bend of experimental curve can be associated with attenuation regime transformation from Akhiezer to Landau-Rumer about 1 GHz. The fitting curve, calculated from Eqs (2), and (4), represents transition from Akhiezer to Landau-Rumer regimes.

Previously, in [12], such transition in the diamond was estimated about 50 MHz, but there are no any calculations or experimental data proving this result. In terms of thermal relaxation time, that means \( \tau = 3.5 \times 10^{-9} \) s (for transition at 50 MHz), while the value \( \omega \tau = 1 \) at 1 GHz in this work results in \( \tau = 1.6 \times 10^{-10} \) s.

Dividing attenuation regimes at 1 GHz, one can get experimental results on \( Q_\mu f \) for two different attenuation regimes. There is no frequency dependence for quality factor in Akhiezer regime, while attenuation has squared frequency dependence: \( Q_\mu f \approx 1.8 \times 10^{13} \) Hz and \( \alpha \approx 0.9 \) dB/cm-GHz\(^2\). In Landau-Rumer regime both quality factor and acoustic attenuation have linear frequency dependence: \( Q_\mu f \approx 1.8 \times 10^{9} \mu f \text{Hz}, \) and \( \alpha \approx 0.9 \) dB/cm-GHz. The best fit of Eqs (2) and (4) with experimental results from Fig. 8 occurs in assumption of \( \gamma_{100}^L = 0.85 \). To make calculations in Eqs (2) and (4), the values of diamond’s BAW velocities and density were taken from [26]. All the represented experimental and calculated results are performed for room temperature.

The origin at such unexpectedly low frequency of transition to Landau-Rumer dissipation mechanism can be associated with high Debye temperature (\( \theta_D \approx 2200 \) K) of a diamond single crystal. The ambient room temperature is much lower than \( \theta_D \). It results in achieving Landau-Rumer attenuation regime at rather low frequencies (1 GHz), and attenuation becomes linear frequency function. Note, that for almost all crystals, which are widely used in acoustoelectronic devices, Akhiezer attenuation regime has been observed. On the higher frequencies the further increasing of acoustic attenuation in diamond becomes rather slow, and total phonon attenuation in synthetic diamond single crystal happens to be lower than in other known crystals, such as SiO\(_2\), Si, LiNbO\(_3\), etc [20].
V. RESULTS AND DISCUSSION

The synthesis of a number of HBARs with “Al/AIN/Mo/(100) diamond” structure has been produced. Precise investigation of materials quality, acoustic losses, experimental and theoretical study of attenuation regimes in diamond single crystal were carried out at room temperature. The set of (100) IIa diamond based HBAR’s substrates was produced with roughness $R_g \leq 10$ nm, controlled by AFM method, and the depth of subsurface damage layer was obtained less than 30 nm, which was proved by Kikuchi lines observation. It was experimentally shown, that the peculiarities of frequency dependences for the quality parameter $Q\mu f$ and acoustic attenuation $\alpha$ can be explained by the transition between Akhiezer and Landau-Rumer regimes about 1 GHz under thermal relaxation time $\tau = 1.6 \cdot 10^{-10}$ s. Experimental results show the absence of frequency dependence for quality parameter $Q\mu f$ in Akhiezer regime: $Q\mu f \approx 1.8 \cdot 10^{13}$ Hz, and $\alpha \approx 0.9$ dB/cm GHz², while in Landau-Rumer regime there are linear frequency dependences for both quality parameter and attenuation: $Q\mu f = 1.8 \cdot 10^{13} f$, and $\alpha = 0.9$ dB/cm GHz². These results can be compared with theoretical equations in assumption of Grüneisen parameter for diamond single crystal as $\gamma_{100} = 0.85$. The change of the acoustic attenuation mechanism at relatively low frequencies for IIa synthetic diamond single crystal means lower values of $\alpha$ compared to other known acoustoelectronic materials.

ACKNOWLEDGMENT

This work has been supported by the Ministry of Education and Science of the Russian Federation (No 14.574.21.0074, unique ID project RFMEFI57414X0074) with the use of Shared-Use Equipment Center “Research of Nanostructured, Carbon and Superhard Materials” (FSBI TISNCM).

REFERENCES

M. Allani1,2, X. Vacheret, A. Clairet, T. Baron, JJ Boy
1FEMTO-ST Institute, UFC, CNRS, ENSMM, UTBM
25000 Besançon, France
2 INSAT, BP 676, 1080 Tunis cedex - Tunisie
C. Reibel, O. Cambon
Institut Charles Gerhard, Université de Montpellier,
34000 Montpellier, France
J.M. Lesage
DGA MI - 35998 Rennes Cedex 9
E-mail: maroua.allani@ens2m.fr

Abstract— Before using any piezoelectric crystal to realize acoustic devices (sensors, transducers, actuators or ultra-stable resonators) and beyond its mechanical properties, the crystal material itself has to be characterized. Whether the very interesting properties of the LGT crystal make it the best candidate to substitute quartz crystal for frequency output devices, we must take into account the crystal quality. Indeed, applications require homogeneous crystals with reproducible physical properties.

The presence of structural defects and inhomogeneity of the crystal composition significantly affect the physicochemical properties (optical, electrical, piezoelectric…), and can be revealed by chemical, optical, and electrical analytical methods. So, before manufacturing acoustic devices, we perform different analyses as: ESR, IR, and UV-Vis spectrometry, ICP-MS coupled with a femtosecond laser ablation, electrical resistivity…

After that, we have fabricated small Bulk Acoustic Wave resonators working on a thickness shear mode of vibration to tentatively know the influence of the crystal quality on the acoustic properties.

Keywords—LGS crystals family; optical spectrometry; ESR; ICP-MS; electrical resistivity, BAW resonator

I. INTRODUCTION

Whether the very interesting properties of the LGT crystal (high acoustic quality, very good thermal stability and high electromechanical coupling factor) make it the best candidate to substitute quartz crystal for frequency output devices, we must take into account the crystal quality. Indeed, applications require homogeneous crystals with reproducible physical properties. Furthermore, it is necessary to know whether some defects do not change over time. If this is the case, it is not impossible that the resonant frequency of the device changes too. Our experience on the use of the quartz crystal indicates that a very small amount of chemical impurities in the lattice can lead to a drift of the resonant frequency, inducing so a bad aging. It forces us elsewhere to choose very accurate technics to control the quality of the crystal, particularly to analyze the impurities levels.

II. WHAT KIND OF ANALYSES?

A. The samples

The crystals belonging to the Langasite (LGS) family grown by the technique of Czochralski pulling which is the most common method for big dimensions “boules”. The benefits of this technique compared to the hydrothermal one (used for quartz) are less restrictive synthesis conditions and its higher industrialization potential. The Czochralski method is used for the congruent compounds (it means the defined crystal has the same composition than the melt). It is the case for LGT (La3Ga5.5Ta0.5O14), LGN (La3Ga5.5Nb0.5O14) and LGS (La3Ga5SiO14) but not for all compounds of the family [1]. On the other side, as the congruent melting composition does not correspond to the stoichiometric one, during growth, this results in continuous radial and longitudinal shifts of the crystal composition [2]. If authors have proved that the generated striations induce SAW velocity variations and so frequency shifts, their effects in BAW devices is still unclear [3]. The most common method used to reveal striations are chemical etching which is sufficiently sensitive to any crystal lattice perturbation.

In the same way, the presence of any structural defect or inhomogeneity of the crystal composition will significantly affect all the physicochemical properties (optical, electrical, piezoelectric…) and can be revealed by chemical, optical and/or electrical analytical methods.

Each defect can have one of the three following origins:
- oxides constituting the melt (La2O3 / Ga2O3 / Ta2O5),
- The growth process including the atmosphere of the furnace during pulling,
- The annealing process necessary to anneal the stress induced by the growth process and the mechanical processing to prepare samples.

As presented previously [4], we have selected different ingots, coming from different suppliers: (Fo and Fn from FOMOS (Russia) – CI3 and CI6 from Cristal Innov (France) and CK as C. Klemenz [3]. Details of growth, post-growth
treatment conditions and color are tabulated in the TABLE 1 (to the extent of our knowledge) and the Fig. 1 presents our samples. Furthermore, some measurements have been achieved on CI6 before annealing (CI6ba) and after annealing (CI6aa). Annealed LGT CI6aa is heterogeneously pale orange colored. To distinguish between zones having different color intensities, we note CI6aa-Y and CI6aa-Z.

**Fig. 1.** Photography of different samples: 1. CI3 – 2. CI6ba – 3. CI6aa – 4. Fn – 5. CK – 6. Fo and the last one (n°7) being the only sample of LGS (n° 7).

<table>
<thead>
<tr>
<th>No</th>
<th>Supplier</th>
<th>Growth axis</th>
<th>Growth atm</th>
<th>Annealing</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cristal Innov</td>
<td>X</td>
<td>N₂ + O₂</td>
<td>- PgA*: Air, 1320°C, 24 h</td>
<td>heterogeneous light orange</td>
</tr>
<tr>
<td>2</td>
<td>Cristal Innov</td>
<td>X</td>
<td>N₂ + O₂</td>
<td>Before annealing</td>
<td>colorless</td>
</tr>
<tr>
<td>3</td>
<td>Cristal Innov</td>
<td>X</td>
<td>N₂ + O₂</td>
<td>- IsA*: - PgA*: N₂, 1400°C, 24 h</td>
<td>heterogeneous light orange</td>
</tr>
<tr>
<td>4</td>
<td>Fomos (HM)</td>
<td>Z</td>
<td>?</td>
<td>Air, 1250 °C, several days</td>
<td>Bright orange</td>
</tr>
<tr>
<td>5</td>
<td>CK</td>
<td>Z</td>
<td>N₂</td>
<td>unannealed</td>
<td>colorless</td>
</tr>
<tr>
<td>6</td>
<td>Fomos</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>green</td>
</tr>
<tr>
<td>7</td>
<td>Fomos</td>
<td>X</td>
<td>?</td>
<td>PgA*: Air, 1250 °C, several days</td>
<td>Bright orange</td>
</tr>
</tbody>
</table>

*PgA*: Post-growth Annealing of the ingot.
*IsA*: In situ Annealing in Czochalski furnace, just after growth and before cooling.

TABLE I. Growth and post-growth treatment conditions of LGT (6 samples) and LGS (1 sample) from different suppliers.

- to tentatively attribute the color of the crystals to a particular defect, we have added to the previous experiments the Electron Spin Resonance (ESR) spectrometry realized on a few mg of the crystal pounded in a very fine powder.
- At least, we have measured the electrical conductivity of different samples in order to compare them.

C. Chemical analyses

**EPMA**

The atomic compositions of the samples were analyzed by Electron Probe Micro-Analysis (EPMA). This microanalysis is performed on small (a few hundred square microns) polished samples having a thickness of about fifty microns. They are bonded to resin or metal support.

Samples were analyzed on a CAMECA SX100 device with five wavelength dispersion spectrometers.

The monochromator crystals are Lithium Fluoride for Ga (K₀), Pentaerythritol for La (La₃) and Thallium acid Phthalate for Ta (Ma). The used standards are GaAs for Gallium, REE3 glass containing rare earths (Y, La, Ce, Pr) for Lanthanum and pure Ta for Tantalum. Each measurement corresponds to the average of a few tens of scans. We compare, then, between the mass of the majors analyzed by EPMA and fs laser ablation coupled with ICP-MS (TABLE II).

**LA-ICP-MS**

All measurements were carried out on a DRC2 quadrupole ICPMS (Perkin Elmer) instrument coupled to an Alphamet femtosecond laser ablation system (Nexeya SA, Canejan, France). This laser machine is fitted with a diode-pumped Yb:KGW crystal laser source (HP1, Amplitudes Systèmes, Pessac, France) delivering 360-480 fs pulses at 1030 nm. The laser beam is focused with a 25-mm objective providing 15 μm diameter spot size. Further details of a previous similar model of this laser ablation system are described elsewhere [5, 6]. The final ablated zone is of 500 μm x 500 μm in 50 s. Each LGT was ablated 10 times under these conditions.

Measuring conditions were adjusted for maximum sensitivity, stability and plasma robustness using a transparent glass certified reference material (NIST 612) [7].

Trace metals concentrations were determined after calibration of the fsLA/ICPMS coupling using certified glass reference material (NIST 612 and NIST 610). In order to counterbalance ablation yield bias from glass to LGT, as well as sample transport an internal standard (La) was used here considering La concentration in LGT of 37.4%.

D. Spectroscopic analyses

**UV-Vis and IR spectroscopy**

The presence of point defects and inhomogeneity of the crystal composition significantly affect the optical properties which can be revealed by optical transmission spectra.

The transmission spectra of LGT and LGS samples with plane and parallel polished surfaces (thickness 4 to 5 mm) were measured in the wavelength range from 200 to 800 nm.
Electron Spin Resonance (ESR) spectroscopy

All measurements of the Electron Spin Resonance (ESR) signal were made using a Bruker ER100 spectrometer in the X-band (at 9.8 GHz) domain.

First, LGT and LGS samples were powdered manually using a small agate mortar and pestle. Then, each powdered sample was transferred into ESR sample tubes. A reference material consisting of DPPH crystal, with an ESR signal at g=2.00 is employed mainly for system performance check.

Each tube was then placed along the axis of the front resonator of the spectrometer.

F. Electrical resistivity

Electrical resistivity of LGS Fs and LGT CI, Fo and CK was measured at room temperature. The samples had the form of Y-cut plates – 15 × 15 mm in lateral dimensions and ~1 to 1.3 mm in thickness. First, we placed on the both sides of the plate an Au/Cu electrode. Then, these two electrodes are connected to a dc-voltage (50 V) and a picoamperemeter. The electrical resistivity of the sample was calculated from the measured current value.

G. Macroscopic defects revealed by chemical etching

The surfaces of the 3 different faces of a cube (X, Y and Z) revealed by chemical etching in a solution of H₃PO₄, during 2h at 130°C are presented below, Fig. 6. After etching, they reveal particular shapes observed by optical microscope.

III. RESULTS AND DISCUSSION

A. Chemical analyses

Stoichiometry

The following table (TABLE II) presents the contents of the majors (La, Ga and Ta) measured in various samples by EPMA and ICP-MS, these 2 techniques having been presented above.

<table>
<thead>
<tr>
<th>Element</th>
<th>La</th>
<th>Ga</th>
<th>Ta</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>theory</td>
<td>37.4</td>
<td>34.4</td>
<td>8.1</td>
<td>20.1</td>
</tr>
<tr>
<td>EPMA</td>
<td>35.3</td>
<td>35.4</td>
<td>28.9</td>
<td>7.5</td>
</tr>
<tr>
<td>ICP-MS</td>
<td>34.1</td>
<td>32.2</td>
<td>7.9</td>
<td>7.8</td>
</tr>
<tr>
<td>CI6ba</td>
<td>35.1</td>
<td>37.4</td>
<td>8.1</td>
<td>7.8</td>
</tr>
<tr>
<td>CI6aa</td>
<td>35.2</td>
<td>37.4</td>
<td>8.1</td>
<td>7.8</td>
</tr>
</tbody>
</table>

For La-ICP-MS analysis, we note that the mass percentage values of Lanthanum are normalized. The mass percentage values of Gallium are under-estimated because our glass standards contain Barium Ba which may create a bias of over 10% on our Ga standard curve.

According to EPMA analysis, we note for all the samples the presence of GaTa′′ defect (Gallium excess of about 2% and Tantalum deficiency 0.3%) and VaLa′′ defect (Lanthanum deficiency of about 1.8%). The gallium excess is due to gallium oxide excess in charge composition.

Impurities level

From the results of fs LA-ICP-MS analysis (Table III), we note that the impurities of the raw material and especially the lanthanides (from La₂O₃) such as Ce and Gd and metals transition (from Ta₂O₅) such as Fe and Ti contaminate crystals while there is no contamination due to Iridium crucible.

Comparing the concentration of Fe and Ti in the crystal LGT Fm and LGT CI6aa, we notice that the intensity of color increases with the concentrations of Fe and Ti. We deduce that these impurities can cause coloration.

After annealing, colorless LGT (CI6ba) became orange (CI6aa). We deduce annealing modifies the valence state of Fe and Ti impurities. Indeed, the valence state of an ion exerts a strong influence on both the hue and the color intensity [8].

<table>
<thead>
<tr>
<th>Sample</th>
<th>Na</th>
<th>Mg</th>
<th>Al</th>
<th>Ti</th>
<th>Fe</th>
<th>Ni</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI6ba</td>
<td>17</td>
<td>1</td>
<td>59</td>
<td>14</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>CI6aa</td>
<td>59</td>
<td>3</td>
<td>165</td>
<td>10</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Fm</td>
<td>81</td>
<td>9</td>
<td>139</td>
<td>25</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>CI6ba</td>
<td>4</td>
<td>8</td>
<td>&lt;0.1</td>
<td>16</td>
<td>1</td>
<td>170</td>
</tr>
<tr>
<td>CI6aa</td>
<td>1</td>
<td>8</td>
<td>19</td>
<td>16</td>
<td>1</td>
<td>171</td>
</tr>
<tr>
<td>Fm</td>
<td>1</td>
<td>4</td>
<td>22</td>
<td>17</td>
<td>2</td>
<td>176</td>
</tr>
</tbody>
</table>

Aluminum impurity can occupy empty interstitial sites or substitutes cations of the crystal lattice. It is likely that Al could replace Gallium and Tantalum because their ionic radiuses are similar. In addition, Al and Ga are in the same group of the periodic table.

The crystals contain alkali metal impurities Na⁺ and K⁺ that can participate in the conservation of electrical neutrality after the appearance of point defects such as GaTa′′ and FeTa′′.

B. Spectroscopic analyses

UV-Vis spectroscopy

The characteristic transmission spectra in UV-VIS ranged from 200 to 800 nm of LGT crystals from the different suppliers are shown in Fig. 2.

The LGT CI3 and CI6aa spectra exhibit a pronounced intrinsic absorption edge around 256 nm. For CK, a shift of the intrinsic absorption edge to a higher wavelength of 270 nm is observed.

The absorption band at 280 nm in the spectra of LGT CK, CI3 and CI6 was recorded by [9, 10] and attributed to “La vacancies” VaLa′′, which correlates also with our EPMA chemical analysis results (TABLE II).
Indeed, the band at 280 nm is more intensive for the sample CI in which lower “La” content is found.

The characteristic transmission spectra in UV-VIS ranged from 200 to 800 nm of LGT crystals from the different suppliers are shown in Fig. 5.

The transmission spectra of the LGT CI3 and CI6aa exhibit a pronounced intrinsic absorption edge around 256 nm. For CK, a shift of the intrinsic absorption edge to a higher wavelength of 270 nm is observed.

The absorption band at 280 nm in the spectra of LGT CK, CI3 and CI6 was recorded by [9, 10] and attributed to “La vacancies” \( V_{\text{La}}^{\text{″}} \), which correlates also with our EPMA chemical analysis results (TABLE II). Indeed, the band at 280 nm is more intensive for the sample CI in which lower “La” content is found.

The weak and wide absorption band at around 330-340 nm observed in the transmission spectra of CK was attributed [9, 10] to oxygen vacancies \( V_{O}^{\text{••}} \).

A strong absorption band at \( \lambda \approx 490 \) observed in LGT Fn and LGS spectra was previously attributed to F centers (\( V_{O}^{\text{••}}, 2é \)) and to orange coloration [10].

A spread and a shift to higher wavelength \( \lambda \approx 350-380 \) nm of the intrinsic absorption edge are observed in the transmission spectra of LGS Fs, LGT Fn and Fo.

The defect responsible for the absorption band at 450 nm can be at the origin of the green color of LGT Fo.

**IR spectroscopy**

The characteristic transmission spectra in IR range from 7000 to 2000 cm\(^{-1}\) of LGT crystals from different suppliers are shown in Fig. 3.

The 3 crystals from FOMOS (Fo, Fn and Fs), colored in deep orange or yellow / green, exhibit an absorption peak at 5370 cm\(^{-1}\) whose intensity increases with that of the color.

This sharp absorption peak seems to be linked to a defect responsible to the crystal coloration. The variation of the intensity (with conservation of the energy, is linked to the variation of the defect density.

Also, the absorption at 3430 cm\(^{-1}\), we attribute it to the link Ga-OH after the lecture of the IR spectrum of the Gallium Oxide (not presented here).

**Electron Spin Resonance spectroscopy**

All ESR spectra of LGT and LGS samples, Fig. 5, except colorless and unannealed (LGT CI6ba), show a spin transition at \( g \) between 4.2 and 4.3, characteristic of Fe\(^{3+}\) in tetrahedral environment and a spin transition at \( g \) between 2.4-2.1 which is characteristic of Fe\(^{2+}\) in octahedral environment. This is correlates with ICP-MS chemical analyses which show the presence of Fe impurities.

The ESR spectrum of green LGT Fo exhibits an additional spin transition at \( g=3.8 \).

ESR spectrum of colorless and annealed LGT CI6ba-Y exhibits two signals at \( g=2.11 \) and \( g=1.94 \) characteristic of Fe\(^{2+}\) in octahedral environment, Fig. 5. After annealing (CIAaa-Y spectrum), the signal at \( g=1.94 \) disappears and two signals appear at \( g=3.16 \) characteristic of Fe\(^{3+}\) in octahedral environment and at \( g=4.19 \) characteristic of Fe\(^{3+}\) in tetrahedral environment.

So, we can conclude that during annealing Fe\(^{2+}\) is partially oxidized into Fe\(^{3+}\).
Comparing the spectra of the two zones of annealed LGT Claa which have different intensities of color (Fig. 1), we note that the spectrum of LGT Claa-Y whose color is more intense exhibits an additional spin transition at $g = 3.16$ characteristic of Fe$^{3+}$ in octahedral environment which is not the most stable for Fe$^{3+}$. It tends to move in tetrahedral environment (with $g = 4.19$) if the annealing is well finished, Fig. 4.

C. Electrical resistivity

In the following table (TABLE V), we note that electrical resistivity measurements of the samples correlate with their band gap energies calculated from the edge of the intrinsic absorption in UV-vis transmission spectra. Indeed, the less resistive LGT Fo and LGS samples have an edge of the intrinsic absorption at higher wavelengths at $\lambda = 350-380$ nm compared with those of LGT CI and CK at $\lambda = 250-270$ nm.

The edge of the intrinsic absorption determines the width of the band gap which is at 4.84 eV for CI and at 4.59 for CK.

In [11], the authors have calculated the band gap of defect-free LGT crystal at 5.279 eV.

D. Vickers Micro Hardness

TABLE IV. Electrical resistivity of LGT and LGS samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$I$ (nA)</th>
<th>$\rho$ (GΩ.cm)</th>
<th>$E_g$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI</td>
<td>5.12</td>
<td>93.6</td>
<td>4.84</td>
</tr>
<tr>
<td>CK</td>
<td>4.55</td>
<td>83.3</td>
<td>4.59</td>
</tr>
<tr>
<td>Fs</td>
<td>5.33</td>
<td>64.2</td>
<td>3.54</td>
</tr>
<tr>
<td>Fo</td>
<td>5.39</td>
<td>45.3</td>
<td>3.26</td>
</tr>
</tbody>
</table>

During grinding and polishing of the samples, we note that they are machined at different speeds. In fact, we have observed that the removal rate for LGT CK is the fastest, as shown on the Vickers micro-hardness measurements.

E. Macroscopic defects revealed by chemical etching

The chemical etching of the surfaces of the crystal leads creation of pits or hillocks for which the shape are different, depending on the crystallographic orientation. The observation of the 3 orientations (X, Y and Z), shows that the mechanism is strongly anisotropic. But, beyond the lattice, our photographs reveal also presence of defects, which can be due, in most cases, to the emergence of dislocations.

If the roughness of the surfaces of the cube is of the order of a few nm, it is, on average, more than 1 µm after etching.

III. THE RESONATORS

To test the acoustic properties of a “new” piezoelectric material, it is finally necessary to manufacture bulk acoustic waves (BAW) resonators working at a given frequency and, if possible, not too sensitive to external parameters. Indeed, the Q-factor defining the quality of a given resonant frequency depends on crystal quality, of course, but also on the geometry of the resonator. We have to take into account the losses due to the mounting structure, the surface quality, the deposited electrodes (thickness and diameter), the atmosphere surrounding the resonator… and the viscoelastic behavior of the crystalline material itself.

To minimize the structure parameters effect, we have chosen a resonator design for which the energy trapping is inherently linked to the electrodes and so working at a reasonable high frequency: around 40 MHz, on the 3rd overtone for example. And finally, to guarantee the best quality material control, we will compare the results to those obtain on a “pure” quartz crystal resonator having the same design and working at almost the same frequency. We hope so to eliminate the “noise” measurement representing the design and to obtain results characterizing only the material effect.

In summary, our resonators are discs of 8.2 mm of diameter, about 100 µm thick with electrodes of different diameters. The 2 parallel surfaces are polished, after different steps of grinding alternating with light chemical etching. The electrodes are deposited by evaporation (gold with a sub-layer in Chromium). At least, the crystallographic orientation is chosen to minimize the thermal effect. It is a single rotated cut, close to the Y-cut which its frequency-temperature curve exhibits a turnover point around 70°C.

The conditions of the trapping being not the same as for quartz, before manufacturing, we have calculated the position...
of the anharmonic modes with respect to the main mode, with a soft developed previously at FEMTO-ST [12]. For that, we choose 2 different diameters: 3.5 mm (as for quartz resonator) and 1 mm. The following table (TABLE VI) shows that there is a great variation between quartz and LGT with the 2 chosen diameters of the electrodes. The LGT resonator with a greater diameter of the electrodes does not seem interesting, the anharmonic modes (called also “spurious mode”) being too close to the main mode.

TABLE VI. frequency spectra for quartz and LGT resonators: we have noted the number of kHz above the main mode.

<table>
<thead>
<tr>
<th>mode</th>
<th>AT-cut SiO2 Diam: 3.5 mm</th>
<th>Y-cut LGT Diam: 3.5 mm</th>
<th>Y-cut LGT Diam: 1 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>300  (main mode)</td>
<td>39.7 MHz</td>
<td>38.473 MHz</td>
<td>38.497 MHz</td>
</tr>
<tr>
<td>320</td>
<td>+78 kHz</td>
<td>+5 kHz</td>
<td>+73 kHz</td>
</tr>
<tr>
<td>302</td>
<td>+153 kHz</td>
<td>+6 kHz</td>
<td>+155 kHz</td>
</tr>
<tr>
<td>340</td>
<td>-</td>
<td>+14 kHz</td>
<td>-</td>
</tr>
<tr>
<td>322</td>
<td>+20 kHz</td>
<td>340 kHz</td>
<td></td>
</tr>
<tr>
<td>304</td>
<td>+33 kHz</td>
<td>460 kHz</td>
<td></td>
</tr>
</tbody>
</table>

We have so fabricated a dozen of resonators in crystal coming from the different suppliers. It is very important to manufacture together a relatively high number of blanks to guarantee a very good parallelism of the surfaces during the grinding process.

The 2 figures above illustrate the table. Indeed, the 2 frequency-spectra whose the span is equal to 100 kHz show clearly that the quality of the 3rd overtone of the CK resonator is worse than the resonant frequency of the AT-cut.

Unfortunately, we cannot present comparative results between families of LGT crystals, except 3 interesting values of the Q-factor of the 3rd ov. of 3 resonators measured with non-deposited electrodes with a diameter equal to 2.5 mm.

The table VII summarizes the values that are waiting to be confirmed. It stays now to confirm this result and to hope that we can compare the crystals together.

TABLE VII. Q-values of the 3rd overtone of 3 resonators (AT-cut on quartz, Y-cut on Fin and CK).

<table>
<thead>
<tr>
<th>Resonator</th>
<th>AT quartz</th>
<th>Y LGT CK</th>
<th>Y LGT Fin</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd overtone</td>
<td># 40 MHz</td>
<td># 41.9 MHz</td>
<td>32.9 MHz</td>
</tr>
<tr>
<td>Q-val (x10^-3)</td>
<td>250</td>
<td>?</td>
<td>348</td>
</tr>
<tr>
<td>Q,f product</td>
<td>10^13</td>
<td>-</td>
<td>1.15x10^13</td>
</tr>
</tbody>
</table>

IV. CONCLUSION

We have qualified LGT crystals from three different suppliers, submitted to different growth and annealing process. Many experiments have been used to compare their quality that correlate with each other. Our first conclusions are that:

- Iron impurities is responsible for LGT color,
- Post growth annealing changes the charge state of Iron impurity and then the color,
- All the samples contain an excess of Gallium, which substitutes the Tantalum and deficiency of Lanthanum
- Point defects decrease electrical resistivity of samples
- We have defined a small resonator which will help us to define the best material for frequency and time applications. Our measurements are in progress.

ACKNOWLEDGMENT

This work was funded by the “Agence Nationale de la Recherche” in the framework of a French research program ASTRID (supported by French MOD through DGA) with the project entitled “ECLATEMS2012”.

Dr B. Gauthier-Manuel (from FEMTO-ST) is gratefully acknowledged for his help and advice for IR and UV-Visible spectra. Our thanks also go to Dr C. Klementz Rivenbark for very profitable talks about crystal growth of LGT.

REFERENCES

Langasite family crystals as promising materials for microacoustic devices at cryogenic temperatures

A. Sotnikov¹,², E. Smirnova², H. Schmidt¹, M. Weihnacht¹, J. Götze³, S. Sakharov⁴
¹) IFW Dresden, SAWLab Saxony, Dresden, Germany
²) A.F. Ioffe Physical-Technical Institute, St. Petersburg, Russia
³) Freiberg University of Mining and Technology, Freiberg, Germany
⁴) JSC Fomos-Materials, Moscow, Russia

Abstract — Using the pulse-echo ultrasonic technique, the velocities of the longitudinal and the shear bulk acoustic waves propagating in five crystallographic directions of LGS and SNGS single crystals were measured in a wide temperature range from 4.2 to 300 K. Elastic constants are derived and treated in the frame of Varshni approach. Some characteristic parameters of the theory are obtained. Strong piezoelectric excitation survives down to liquid helium temperature.

Keywords—Langasite family crystals; elastic constants; cryogenic temperatures.

I. INTRODUCTION

Piezoelectric single crystals of the langasite (LGS) family belonging to the same trigonal crystal class 32 as quartz are of current interest in microacoustics as promising materials for various bulk- and surface acoustic wave devices and sensors. Langasite was grown for the first time in Russia in 1980s [1]. Like quartz, LGS shows temperature compensated cuts for bulk- and surface waves, but compared with quartz it has considerably higher electromechanical coupling coefficients. Other important features of langasite is the absence of a structural phase transition between liquid helium temperature and its melting point and commercial availability of large size crystals grown by the well-developed Czochralski technique. Nowadays high quality LGS boules with diameter up to 100 mm, length up to 102 mm and weight up to 5 kg are easily available (see for example [2]). However it was found that acoustic and dielectric losses as well as electric conductivity in LGS increase with increasing temperature at temperatures higher that about 600° C. These features are explained by the langasite disordered structure [1]. As a result of further successive attempts, now the langasite family includes not only disordered crystals like such as LGS, LGN and LGT, but also the ordered compounds such as Sr₃NbGa₃Si₂O₁₄ (SNGS), Sr₃TaGa₃Si₂O₁₄ (STGS), Ca₃NbGa₃Si₂O₁₄ (CNGS) and Ca₃TaGa₃Si₂O₁₄ (CTGS) with clearly better properties at elevated temperatures [3 - 5]. Notice that the use of piezoelectric crystals in microacoustic devices is not only a challenge for high temperature applications. In addition to the normal and high temperature conditions, also the ability for low (cryogenic) temperature operation is of special attractivity. Obviously, for successful application the crystals should preserve their physical properties at lowest temperatures without considerable worsening. Additionally, specific features of low temperature elastic properties of the langasite family crystals are also important for basic understanding of the crystal lattice dynamics. To our best knowledge the only paper on LGS material parameters at low temperatures is [6]. Notice also [7] which is devoted to low temperature properties of quartz. In the present contribution, we report on the dielectric and elastic properties of LGS and SNGS piezoelectric crystals at temperatures from 4.2 K to 295 K. In contrast to [6, 7] where resonance ultrasound spectroscopy were used, our data were obtained with the traditional pulse-echo ultrasonic method.

II. SAMPLE PREPARATION

LGS and SNGS single crystals were grown by the Czochralski technique at FOMOS-Materials (Moscow, Russia) and Freiberg University of Mining and Technology (Freiberg, Germany), respectively. For the bulk acoustic wave velocity measurements cubes of 7 x 7 x 7 mm³ were cut from the boules in three different orientations: (i) with the edges parallel to the X, Y and Z crystallographic axes and (ii) rotated by ±45° around the X axis. Components of the dielectric constant tensor were obtained using X- and Z-cuts plates with the averaged dimensions of 10 x 10 x 0.5 mm³. Orientation accuracy was within 0.1° of the X, Y, Z and ± 45° YZ directions for all cases. All the samples were carefully ground followed by fine polishing to achieve parallelism (about 0.5 μm/mm) and flatness (within 0.5 μm) of the opposite faces. Gold electrodes of 300 nm thickness have been deposited on the big opposite faces of the plates for dielectric measurements.

This work was supported by DFG (grant SCHM 2365/11-1), BMBF (grant InnoProfile-Transfer 03IPT610Y), and by RFBR 14-02-91330 NNIO grant.
III. EXPERIMENTAL PROCEDURE

Measurements of the bulk acoustic wave velocities were carried out by a RITEC Advanced Ultrasonic Measurement System RAM-5000. The system realizes the pulse-echo method of time propagation measurements with an accuracy of about 10⁻⁴. To generate longitudinal and shear ultrasonic waves, Y+36° and X41° LiNbO₃ transducers of 10 MHz central frequency were used. A special attention was paid to the couplant materials to bond transducers on the surface of the samples under study, especially in the case of shear modes. Dielectric measurements were performed using a Solartron SI 1260 Impedance/Gain-Phase Analyzer. The relative dielectric constants ε₁₁ and ε₃₃ were obtained by measuring the capacitance of X- and Z-cut plates, respectively, at a frequency of 1 kHz which is far below any frequency corresponding to the electromechanical resonances. All temperature measurements were done using a continuous flow Oxford cryostat in the 4.2 to 300 K range with an accuracy and stability of 0.1 K.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Temperature dependences of the dielectric constant ε₃₃ are shown in Fig. 1 for LGS and SNGS crystals. As is seen from Fig. 1, ε₃₃ for LGS increases with temperature decreasing and saturates at very low temperatures while ε₃₃ for SNGS decreases only slightly with decreasing temperature. The first behavior is typical of so called incipient ferroelectrics while the second is usual for ordinary dielectrics. Notice that ε₁₁ for both crystals behaves in similar manner i.e. decreases with decreasing temperature.

The elastic constants Cᵢⱼ for 32 symmetry class crystals can be derived using a system of relations between elastic constants and sound velocities measured at different directions of propagation for different modes (see for example [8]). For instance, measurements of the longitudinal and shear sound velocities propagating along Z direction yield Cₑ₃₃ and Cₑ₄₄, respectively. The combination of two shear modes propagating along X-axis gives (Cₑ₆₆ + Cₑ₄₄) and therefore Cₑ₆₆ taking into account already known elastic constant Cₑ₄₄. Cₑ₁₁ can be derived from the velocities of quasi-longitudinal and quasi-shear modes propagating along Y direction, Cₑ₁₂ can be easily obtained from the relation Cₑ₆₆ = ½(Cₑ₁₁+Cₑ₁₂). The remaining Cₑ₁₄ and Cₑ₁₃ elastic constants can be obtained using velocities along X and ±45°YZ directions and before determined elastic constants. Figs. 2, 3 show as example temperature dependences of Cₑ₁₁ and Cₑ₃₃ for SNGS [9] while Figs. 4, 5 represent the same constants for LGS. It is seen from Figs. 2-5, the common feature of the temperature dependences of the elastic constants is a gradual increase with decreasing temperature followed by saturation at low temperatures (T < about 50 K). In contrast to results obtained in [6] for LGS, we cannot find for both SNGS and LGS any specific softening of Cₑ₁₁, Cₑ₃₃ and Cₑ₁₂ elastic constants at temperatures around 150 K.

Fig. 2. Temperature dependence of the elastic constant C₁₁ for SNGS

Fig. 3. Temperature dependence of the elastic constant C₃₃ for SNGS

Figs. 6 and 7 represent temperature dependences of the elastic constant Cₑ₆₆ for SNGS and LGS, respectively. Notice the
“usual” temperature behavior for SNGS [9], but quite differently a gradual decreasing of $C_{66}$ with temperature decreasing for LGS. Besides, there exists a turnover point around room temperature. The usual temperature behavior of all elastic constants (the only exclusion is $C_{66}$ for LGS) for both SNGS and LGS can be successfully fitted by using the Varshni function [10]:

$$C_i(T) = C^0 - \frac{s}{\exp(\frac{t}{T}) - 1},$$

where $C^0$ is the zero-temperature elastic constant, $t$ relates to the Einstein temperature $\Theta_E$ and $s$ is the lattice anharmonicity parameter. In simple cases $\Theta_E = \frac{3}{4} \Theta_D$, where $\Theta_D$ is the Debye temperature.

Later, on the base of Einstein oscillator model Ledbetter [11] showed that

$$s = \frac{3k\gamma(\gamma+1)\Theta_E}{V_a},$$

where $k$ is the Boltzmann constant, $\gamma$ the Grüneisen parameter and $V_a$ the atomic volume. The solid lines in Figs. 2-6 show least square fitting of the function (1) to the experimental results. The fitting parameters are presented in Table I.
It is important that the coupling coefficients of both SNGS and LGS single crystals remain still high enough with decreasing temperature down to 4.2 K. To demonstrate this remarkable piezoelectric activity in langasites at cryogenic temperatures we used the direct excitation of the acoustic waves by the internal piezoelectrically active shear mode propagating along Y direction in LGS and SNGS at room and liquid helium temperatures. It is clearly seen that a strong signal is still present at 4.2 K.

V. CONCLUSION

Dielectric and elastic properties of LGS and SNGS piezoelectric crystals have been studied at temperatures from 4.2 K to 300 K. The obtained results for the elastic constants (with the exception of \( C_{66} \) for LGS) have been treated using Varshni approach based on the Einstein oscillator model. In LGS, the elastic constant \( C_{66} \) versus temperature shows a turnover point close to room temperature followed by a gradual decreasing with decreasing temperature down to 4.2 K. It has been also demonstrated that high piezoelectric activity of the crystals keeps down to 4.2 K which predestine clearly LGS and SNGS as promising materials for applications at cryogenic temperatures.
ACKNOWLEDGMENT
The authors would like to thank Dr. S. Biryukov for helpful discussions.

REFERENCES
As-doped Si’s Complex Permittivity and its Effects on Heating Curve at 2.45 GHz Frequency

Siddharth Varadan, George Pan, Member, IEEE
Department of Electrical Engineering
Arizona State University
Tempe, AZ, USA
Email: siddharth.varadan@asu.edu

Zhao Zhao, Terry Alford, Member, IEEE
School of Matter, Transport and Energy
Arizona State University
Tempe, AZ, USA
Email: TA@asu.edu

Abstract—An analysis to determine the complex permittivity of arsenic-doped silicon wafer at 2.45 GHz is presented based on closed-form analytical expressions for cylindrical symmetry. Experimental results in support with the numerical analysis and simulation results are also presented. This analysis will further help analyze the capacitive heating of doped and undoped silicon wafer at microwave frequency; hence, this paper is a precursor to elucidation of capacitive heating of silicon substrates placed between susceptors. This study indicates that when the dopant is added to the silicon the loss tangent decreases with increase in concentration but upon annealing the loss tangent becomes constant with respect to concentration of the dopant.

Keywords—Complex permittivity, microwave annealing, silicon wafer dielectric constant, temperature curve.

I. INTRODUCTION

Silicon device features are being scaled down rapidly. Although faster device performance is achieved, this has created many processing challenges. Doping concentration needs to be increased significantly. Ion implantation is the most commonly used method for the placement of dopants into the Si substrate [1]. But such implantations into silicon at high concentrations and low energies damage the silicon near the surface. The amount of the damage is directly proportional to the nuclear energy loss deposited by kinematic scattering events implanted dopants [1]. Microwave (MW) annealing has been proposed by Alford, et al. in [2] as one of the methods to repair the damage created by ion implantation and to electrically activate dopants. MW processing presents a more uniform, volumetric heating of the wafer mainly due to high penetration depth of the microwave radiation compared to conventional annealing [1]. The complex permittivity of the silicon is one of many properties that influences the MW heating. Therefore the knowledge of complex permittivity of doped silicon is necessary.

The complex permittivity is an important electromagnetic property used in designing, modeling, fabricating and testing of MW processed Si. Numerous measurement techniques for estimating the complex permittivity of a material have been proposed in the past. These include cavity resonator, transmission line, free space and open-ended coaxial probe [3]-[10]. The measurement technique used in this paper was proposed by Keam and Williamson in [11] and was later extended by Guo, et al. [12].

Junctions between coaxial cable and waveguide are widely used in microwave devices like power dividers, rectangular waveguide and power combiners. The junction considered for this analysis is a coaxial line with center conductor, entering the cylindrical waveguide through its flat face. The radial line and rectangular waveguide has been separately considered by Otto [13] and by Williamson [14], where the accurate current distribution and input admittance at coaxial-line input are presented.

Analyses previously presented [11-15] were for low-loss dielectric (i.e., loss tangent of 0.02) whereas the investigation done here estimates the complex permittivity of a highly lossy (i.e., loss tangent close to 50) dielectric (arsenic-doped silicon wafer). The effect of annealing on the complex permittivity was compared by measuring the permittivity before and after annealing. Experiments were conducted to validate the numerical and simulation results. The reflection coefficients were obtained from the coaxial cable and were matched with the reflection coefficients obtained from the simulation results obtained from HFSS (High Frequency Structural Simulator) simulation software and analytical solutions obtained by solving field equations for boundary conditions in the experimental setup. Three samples with varying doping concentration were analyzed and presented in this paper.

Sample 1 is an undoped Si substrate, samples 2 and 3 have Arsenic (As) implanted in the Si substrate with two different dose of 1×10¹³/cm² and 4×10¹³/cm², respectively. Each wafer has a 0.5 mm hole that was laser drilled in the center of the wafer. Heating curves are obtained during each anneal, one susceptor-assisted anneal and the other is conventional MW anneal. Rate of temperature change is calculated in both cases and is compared to the one acquired analytically using the bulk properties of arsenic-doped silicon wafer.

The MW (2.8×10⁴ cm² cavity) oven was used for the post-anneal with the frequency centered at 2.45 GHz, the radiation was generated with the use of a 1200 W magnetron.
source [16]. The wafer temperature was monitored with the use of a Raytek Compact MID series pyrometers.

II. FORMULATION

Fig. 1 shows the cylindrical waveguide, which is fed by a coaxial cable of radius $r = b$. The central conductor is of radius $r = a$. The height of this waveguide is $h$. The radial line region is filled with silicon up till $r = c$. The region beyond silicon, $r > c$, is assumed to be filled with air with relative permittivity $\varepsilon_r = 1$.

![Fig. 1. Radial transmission line used in the experimental setup along with its dimensions, where $a = 15 \text{ mm}$, $b = 25 \text{ mm}$, $c = 101.6 \text{ mm}$, $d = 152.4 \text{ mm}$ and $h = 0.6 \text{ mm}$. The center of junction between coaxial cable and the cylindrical waveguide is assumed to be the origin.](image)

![Fig. 2. Laboratory setup of radial transmission line (cylindrical cavity) [17]. The dimensions of the coaxial line are set such that it supports only quasi-TEM waves at the desired frequency. It is also assumed that all conductors have perfectly conducting surfaces.](image)

The analysis of a lossless dielectric is given in [13] and this paper extends the idea to a lossy dielectric case. There are six boundaries, namely, $r = a$, $r = b$, $r = c$, $r = d$, $z = 0$ and $z = h$, where the boundary conditions must be enforced so as to determine the field distributions uniquely through this setup. The dimensions were selected to satisfy (1), therefore no cavity modes are excited.

$$h << \frac{1}{2f \sqrt{\mu \varepsilon}}.$$  

The following equations are the fields inside the coaxial cable ($z < 0$), obtained by solving the above boundary conditions. Equations (2) and (3) are electric and magnetic fields in the region $a \leq r \leq b$:

$$E_z = -\frac{q'^2}{2n^2} \frac{\eta}{\varepsilon_0} f(\alpha) I_0(q'a) K_0(q'r) - \frac{2}{\ln(b/a)} [I_0(q'b) K_0(q'r) - I_0(q'a) K_0(q'r)] + A I_0(q'r)$$

$$H_\phi = \frac{q'^2}{2n^2} f(\alpha) I_0(q'a) K_1(q'r) - \frac{2}{\ln(b/a)} [I_0(q'b) K_1(q'r) - I_0(q'a) K_1(q'r)] + j \frac{k'}{\eta} A I_0(q'r)$$

where, $E_z$ and $H_\phi$ are the electric and the magnetic field in $z$ and direction, respectively, $f$ represents the current flowing through the central conductor of the coaxial cable, given in (15). $I_0$, $I_1$, $K_0$ and $K_1$ are modified Bessel functions; $q'$ is a constant defined as $q' = \sqrt{\frac{\varepsilon_0}{\varepsilon_0'}}$; $k'$ is the wavenumber within the dielectric. $\eta$ is the intrinsic impedance of the dielectric; $\alpha = m\pi/h$, where $m$ is a natural number and is obtained by applying the boundary condition that the tangential electric field at $z = 0$ and $z = h$ is 0; $A$ is a variable dependent on various factors given by (9). Equations (4) and (5) are fields in the region $b \leq r \leq c$ and $0 \leq z \leq h$.

$$E_z = \frac{q'^2}{2n^2} f(\alpha) I_0(q'a) K_0(q'r) - \frac{2}{\ln(b/a)} [I_0(q'b) K_0(q'r) - I_0(q'a) K_0(q'r)] + A I_0(q'r)$$

$$H_\phi = \frac{q'^2}{2n^2} f(\alpha) I_0(q'a) K_1(q'r) + \frac{2k'}{q' \eta^2} [K_1(q'r) I_0(q'b) - L_0(q'a)] + j \frac{k'}{\eta} A I_0(q'r)$$

Equations (6) and (7) represent fields in the region $c \leq r \leq d$

$$E_z = E(K_0(q'r) + S I_0(q'r))$$

$$H_\phi = -\frac{jk'}{q' \eta} E(K_1(q'r) - S I_1(q'r))$$

where, $S$ is the environment factor that depends on the surrounding that the dielectric is kept. In this study the surrounding is a cylindrical waveguide and hence $S = -K_0(qd)/I_0(qd)$ as derived in [11]; $k$ is the free space wavenumber; $q$ is defined as $q = \sqrt{\frac{\mu}{\varepsilon}}$ and $\eta$ is the intrinsic impedance of free space.

Equations (8) gives the
coefficient $E$ used in equation (6) and (9) give the coefficient $A$ used in the equations (2-5)

\[
E = \frac{1}{S_0} \left\{ -\frac{q'q''}{2\eta k^2} I(\alpha)I_0(q'a)K_0(q'c) - \frac{2}{\ln(b'/a)} K_0(q'b) \right\} + A I_0(q'c)
\]

(8)

\[
A = \frac{j k'}{q'q''} I_0(q'c) + \frac{j k'}{q'q''} I_0(q'c)
\]

(9)

\[
2 \left[ I_0(q'b) - I_0(q'a) \right] \left( \frac{k'}{q'q''} K_0(q'c) - \frac{k'}{q'q''} K_1(q'c) \right)
\]

where $S_0 = K_0(qc) + SI_0(qc)$ and $S_1 = K_1(qc) - SI_1(qc)$. The reflection coefficient is analytically obtained from the input admittance of the coaxial cable, which is given by

\[
Y_m = \frac{2\pi}{\ln(b'/a)} \int_a^b H_1(\rho,0) d\rho
\]

(10)

Here $Y_m$ is the input admittance calculated at the $z = h_2$. For convenience, it is assumed that the excitation is 1V and (10) becomes:

\[
Y_m = \frac{2\pi}{\ln(b'/a)} \int_a^b H_1(\rho,0) d\rho
\]

(11)

From Equs. (11) and (3) we get,

\[
Y_1 = \frac{q'q'' I(\alpha)}{\ln(b'/a)} \left[ I_0(q'a) \{ K_0(q'b) - K_0(q'a) \} \right]
\]

(12)

\[
Y_2 = \frac{j \pi k'}{q'q'' \ln(b'/a)} \left[ \frac{1}{q} - I_0(q'a) \right] \{ K_0(q'b) - K_0(q'a) \}
\]

(13)

\[
Y_3 = \frac{j \pi k'}{q'q'' \ln(b'/a)} A \{ I_0(q'b) - I_0(q'a) \}
\]

(14)

The total input admittance, $Y_m = Y_1 + Y_2 + Y_3$. Equation (15) is the current flowing through the central conductor in the coaxial cable, while equations (16) and (17) are constants used in equation (15)

\[
I = \frac{-j \pi k'}{q'q'' \ln(b'/a)} \left\{ \frac{q'q''}{q'q''} K_0(q'c) I_0(q'a) - K_0(q'c) I_0(q'a) + K_0(q'c) I_0(q'a) \right\}
\]

(15)

\[
\gamma = q'q'' S_0 (K_0(q'c) I_0(q'a) - K_0(q'c) I_0(q'a))
\]

(16)

\[
\psi = q'q'' S_0 (I_0(q'c) K_0(q'a) + K_0(q'c) I_0(q'a))
\]

(17)

III. RESULTS

Fig. 3 shows the input port voltage reflection coefficient (S11 in dB), experimental result vs HFSS software simulation result, for annealed undoped silicon wafer from 1-4 GHz. The HFSS simulation result, in the Fig. 3, is obtained by placing an arbitrary material in a setup that closely resembles the cylindrical waveguide and tuning the relative permittivity and loss tangent of this material to a point when the curve closely traces the experimentally obtained reflection coefficient curve.

Fig. 4 compares S11 in dB, obtained from the experimental reflection coefficient for annealed doped (all doses) sample with the analytical formulation given in equations (2-17). Figs. 5 and 6 show the reflection coefficient of samples 2 and 3, respectively, having post annealing dose of $1 \times 10^{15}$/cm$^2$ and $4 \times 10^{15}$/cm$^2$.

Fig. 5. Analytical solution and experiment observation compared for annealed doped (all doses) silicon wafer from 1-4 GHz.
Fig. 5. Experimental result of reflection coefficient from 1-4 GHz of sample 2 (doped silicon with doping concentration $1 \times 10^{15}$/cm$^2$). This is reflection coefficient obtained post annealing the samples.

Fig. 6. Experimental result of reflection coefficient from 1-4 GHz of sample 4 (doped silicon with doping concentration $4 \times 10^{15}$/cm$^2$). This is reflection coefficient obtained post annealing the samples.

Figs. 7 and 8 show reflection coefficient comparison, experimentally obtained using Vector Network Analyzer (VNA) vs simulation results from HFSS, of unannealed doped samples, for frequency range 1-4 GHz.

Fig. 7. $S_{11}$ in dB of As-doped Si wafer with a dose $1 \times 10^{15}$/cm$^2$ for frequency range 1-4 GHz. This is reflection coefficient of silicon obtained prior to annealing.

Fig. 8. $S_{11}$ in dB for As-doped Si with a dose $4 \times 10^{15}$/cm$^2$ for the frequency range of 1-4 GHz. This is reflection coefficient of silicon obtained prior to annealing.

IV. DISCUSSION

The experimentally obtained values of the relative permittivity and the loss tangent values of the pre-annealed silicon samples are shown in Table I. The loss tangent of undoped silicon substrate is high. When the silicon is doped with As dopants, prior to annealing, the loss tangent drops considerably. However when it is annealed and activated the loss tangent increases higher than the undoped silicon, this implies that the conductive losses are higher [17]. This theory is supported by the Hall measurement and the four-point-probe measurements as shown in Table III and IV and the discussion following them.
The relative permittivity and the loss tangent values in Table I was obtained by curve fitting in both methods, namely analytically using equations (2-17) and in the HFSS simulation. In [2], authors proposed that the power absorbed from the projected microwaves can be related to the heat energy required for annealing. Equation (18) directly relates to the formula proposed by Alford, et al., [2], where the rate of temperature change is the ratio of heat energy to the absorbed microwave energy.

\[
\frac{\Delta T}{\Delta t} = \frac{\rho \varepsilon_p \tan \delta |E|^2}{\rho_{max} C_p}
\]

(18)

where, \(\rho_{max}\) is the mass density of silicon and \(C_p\) is the heat capacity of silicon. Table II compares the results obtained by substituting the values in Equ. (18) with the rate of change of temperature obtained from the heating curves, shown in Fig. 9. The pyrometer used has a detection range from 200-2000 °C. The lower dose As-doped silicon quickly rises to a higher temperature and then slowly settles at the saturation temperature; while, the higher dose As-doped silicon takes more time to reach its relatively higher saturation temperature.

Table II indicates that the assumption made in [1], that the heating curve is driven by the sheet properties of the sample and not the bulk properties, is correct. This also is valid assumption considering the fact that in a good conducting material, the electromagnetic (microwaves) waves do not penetrate beyond the skin depth and this forms the sheet of this sample. To further validate this fact Hall effect and four-point-probe measurements (before and after annealing) the observations from the annealed samples are listed in Table III and Table IV below. Table III and IV indicate that the change of temperature for heating curve from susceptor-assisted annealing with heating curve from annealing with silicon base and heating curve calculated from (18).

### Table I: Complex Permittivity of Arsenic-Doped Silicon

<table>
<thead>
<tr>
<th>Dose (cm(^{-2}))</th>
<th>Relative Permittivity</th>
<th>Loss tangent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate (Undoped)</td>
<td>11.9</td>
<td>55.5</td>
</tr>
<tr>
<td>1 × 10(^{15})</td>
<td>11.9</td>
<td>5.0</td>
</tr>
<tr>
<td>4 × 10(^{15})</td>
<td>11.9</td>
<td>4.8</td>
</tr>
<tr>
<td>Annealed (all doses)</td>
<td>11.9</td>
<td>64.7</td>
</tr>
</tbody>
</table>

### Table II: Rate of Temperature Change Comparison

<table>
<thead>
<tr>
<th>Dose (cm(^{-2}))</th>
<th>Rate of temperature change during susceptor-assisted annealing ((K , s^{-1}))</th>
<th>Rate of temperature change during annealing with silicon as base ((K , s^{-1}))</th>
<th>Rate of temperature change obtained from ((17) (K , s^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 × 10(^{15})</td>
<td>5.5</td>
<td>1.6</td>
<td>4.9 × 10(^{-6})</td>
</tr>
<tr>
<td>4 × 10(^{15})</td>
<td>5.4</td>
<td>3.4</td>
<td>4.7 × 10(^{-6})</td>
</tr>
</tbody>
</table>

Interestingly, the dose of arsenic in silicon wafer affects the heating curve when annealed; where as in susceptor-assisted annealing it has, relatively, no effect. The temperature attained in the latter case is also higher and hence simultaneously activating the dopant.

There are 2 assumptions made in the calculations presented in Table II. First, the rate of temperature change is linear. Since the rate is high, this assumption is not totally incorrect. Second, the loss tangent is constant with respect to temperature change which from [18] it is known that it is not true. But in this presentation, the second assumption, gives satisfactory results.

Fig. 10 shows the heating curve of arsenic-doped silicon annealed with silicon base for doses 1 × 10\(^{15}\) /cm\(^2\) and 4 × 10\(^{15}\) /cm\(^2\). Comparing figs. 9 and 10 silicon has a higher saturation temperature when it is annealed with susceptor, which is more than the activation temperature while without susceptor the saturation temperature is much lower and hence it is not activated immediately. Table II shows the comparison of rate

---

Fig. 9. Heating curve of As-doped Si with a dose 1 × 10\(^{15}\) /cm\(^2\) and 4 × 10\(^{15}\) /cm\(^2\) measured during susceptor-assisted annealing process.

Fig. 10. Heating curve of As-doped Si with a dose 1 × 10\(^{15}\) /cm\(^2\) and 4 × 10\(^{15}\) /cm\(^2\) measured during annealing with silicon base.
readings are consistent with the expected results [1] and agree with the previous experiment results presented in this paper.

<table>
<thead>
<tr>
<th>TABLE III</th>
<th>HALL MEASUREMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose (cm²)</td>
<td>Mobility (cm²/V.s)</td>
</tr>
<tr>
<td>1 × 10¹⁵</td>
<td>2.0 × 10⁴</td>
</tr>
<tr>
<td>4 × 10¹⁵</td>
<td>2.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE IV</th>
<th>FOUR-POINT-PROBE MEASUREMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dose (cm²)</td>
<td>Sheet Resistance (Ω)</td>
</tr>
<tr>
<td>1 × 10¹⁵</td>
<td>127</td>
</tr>
<tr>
<td>4 × 10¹⁵</td>
<td>37.8</td>
</tr>
</tbody>
</table>

V. Conclusion

The complex permittivity of Arsenic-doped silicon wafer before and after annealing at 2.45 GHz were measured and presented in this investigation. Analytical solutions and HFSS simulation in support of the measurement were presented. These results are further used to analyze the heating curve during the process of susceptor-assisted annealing. It was concluded that the bulk properties of the material have nil contribution in these annealing processes and it only depends on the sheet properties of doped silicon wafer.

Acknowledgment

The authors would like to thank Mr. C. Birtcher for the help with the experiments and also for lending his laboratory equipment when needed. The authors would also like to thank Dr. J.T. Aberle for his inputs and advise. This research was partially funded by the Intel Corporation.

References

Sputtered Al\((1-x)\)Sc\(_x\)N thin films with high areal uniformity for mass production

V. Felmetser, M. Mikhail
OEM Group Inc.
Gilbert, AZ, USA
valeriy.felmetser@oemgroupinc.com

GMME-CEMDATC-ETSIT
Universidad Politécnica de Madrid
Madrid, SPAIN

Abstract— In this work, we describe a sputter technique enabling deposition of AlScN thin films with homogeneous thickness and composition on production size wafers (150-200 mm) and present some preliminary results on the assessment of the structural and piezoelectric properties of the films with Sc content of about 6.5 at.%. The technique is based on the use of pure Sc ingots embedded into the Al targets of the dual-target S-gun magnetron enabling reactive sputtering with high radial thickness and composition homogeneity. Rutherford backscattering spectrometry was carried out to obtain the film composition. The microstructure and morphology were assessed by X-ray diffraction. Density was determined by X-ray grazing angle reflectometry. Electroacoustic properties and dielectric constant were derived from the frequency response of BAW test resonators. 1 μm-thick films showed wurtzite structure with pure c-axis orientation and rocking curves of the (00-2) diffraction peak with FWHM as low as 1.5°. Film properties appear to be uniform across 150-mm wafers. The material electromechanical coupling factor reached 9%, although the sound velocity of longitudinal mode was relatively low (around 8500 m/s).

Keywords— Aluminum nitride, Doped AlN films, AlScN, BAW, Sputtering, Piezoelectric, Electroacoustic devices

I. INTRODUCTION

Aluminum nitride (AIN) is the preferred piezoelectric material for a broad range of thin film devices and applications requiring piezoelectric actuation, such as resonators for telecommunication, RF filters [1], sensors [2], or microelectromechanical systems [3]. Due to its good piezoelectric properties and chemical stability, AIN is appropriate for operating in moderately harsh environments too [4]. However, new advanced applications require further improvement of piezoelectric properties of AIN.

Since Akiyama and coworkers demonstrated the remarkable effects of doping AIN with scandium [5], achieving Sc-doped AIN films has become a challenge for scientists and device manufacturers searching for a technology suitable for mass production.

It has been verified that the addition of Sc produces a significant increase of the overall piezoelectric activity of the Al-Sc-N alloy [5], although other parameters, such as the acoustic losses or the sound velocity, worsen [6]. In fact, as the Sc content increases, the e\(_{33}\) coefficient decreases to values around 320 GPa at a Sc concentration of 12 at.%, considerably lower than the 395 GPa of pure AIN [7]. As the material density does not vary much with addition of Sc (because the unit cell volume increases roughly in the same proportion as its mass), the sound velocity decreases accordingly [7]. This reduction of the sound velocity together with the softening of the material, which implies more acoustic losses, makes it less attractive for high frequency devices. However, for other devices working at lower frequencies, such as SAW type resonators [8] or mechanical resonators containing cantilevers or bridge architectures, the increase of \(d_3\) could be very beneficial [9]. One more useful feature of Sc-doped AlN is its larger dielectric constant, reaching up to 14\(_{10}\) compared to 10\(_{10}\) of pure AlN [8,9].

Reactive magnetron sputtering is the most appropriate technique to synthesize AlN-based ternary compounds due to good controllability of film morphology and suitability for industrial implementation. However, sputter process performance is substantially tied to the magnetron and sputter target design and composition. Although, at first glance, using Al-Sc alloy targets seems to be an obvious solution [9,10], the metallurgy of the Al-Sc system impedes obtaining alloys with Sc content greater than 12%. Moreover, this approach has limited controllability of the film composition during target lifetime, since the ratio between the two metallic components in sputtered flux may change as a result of target erosion. A possible solution is co-sputtering of two elemental targets [5, 11]. The disadvantage of this approach is the inability to ensure homogeneity of the film composition along the substrate, especially on large size wafers employed in mass production. A more versatile method for obtaining Al-Sc-N films with well-defined composition is to use Sc ingots placed on an Al target [12-14]. The distribution and quantity of these ingots on the target enables managing the film composition and uniformity. In this way, any composition can be obtained, which is important for research purposes. So far, this method has been used with small targets (76 mm in diameter or less) and on small substrates for assessing the Al-Sc-N properties.

In this communication, we describe a novel sputter technique enabling deposition of Sc-doped AlN thin films with homogeneous thickness and composition on production size wafers (150-200 mm) and present some preliminary results on the assessment of the structural and piezoelectric properties of the films. The technique is based on the use of pure Sc ingots embedded into the Al targets of the dual-target S-gun magnetron for AC reactive sputtering implemented in an Endeavor-AT™ cluster tool from OEM Group Inc.
II. EXPERIMENTAL TECHNIQUES

A. Film deposition

1-μm thick AlScN films were deposited on 150-mm diameter thermally oxidized Si (111) wafers (oxide layer thickness 500 nm). The wafers intended for piezoelectric characterization, contained an acoustic reflector previously deposited. The mirror alternated four Al and Mo layers, serving as low and high acoustic impedance materials, with thickness tuned to achieve a reflection band centered at 2.5 GHz. These substrates allowed fabricating bulk acoustic wave (BAW) test resonators with satisfactory quality for extracting AlScN acoustic properties. A 300-nm thick high-quality (110)-oriented Mo film acted as the bottom electrode.

All the films were deposited in an Endeavor-AT™ cluster tool from OEM Group equipped with dual-target S-gun magnetrons [15]. Briefly, the S-gun consists of two ring-shaped targets (diameter of 178 and 280 mm) mounted concentrically on the same vertical axis. Each target has its own magnetic array with opposite polarities. An alternating current power of 40 kHz applied between the two targets creates a plasma discharge at the conical face of each target. The targets are made of pure Al with a series of 12.27 mm in-diameter holes located in the center of the target erosion zone that enable to lodge Sc or Al pellets with a simple cylindrical shape. The pellets precisely fit to the holes to ensure good thermal and electrical contact during sputtering. A specific target arrangement is illustrated in Fig. 1.

With this target arrangement, a desired film composition can be easily attained by varying the number of the Sc inserted pellets. Composition uniformity can be also controlled by adjusting the ratio of the number of pellets in the inner and the outer targets. For this study, the number of Sc pieces was kept constant (10 in the inner and 14 in the outer target).

Prior to AlScN deposition, the substrate surface (Mo bottom electrode) was treated with low energy Ar ions in a separate planarized RF (13.56 MHz) etch module. Deposition processes were performed without external heating. Maximal temperature during AlScN reactive sputtering did not exceed 300°C.

B. Film characterization

The composition of AlScN was assessed by Rutherford backscattering spectrometry (RBS). Figure 2 shows the RBS spectrum of a typical AlScN/Mo/Al/Mo/Al/Mo/SiO₂/Si stack. Fitting the spectra of Fig. 2 with the SIMNRA software tool [15] allowed us to determine the composition of the film, yielding a Sc content of 7% atomic.

X-ray diffraction (XRD) patterns of the Al₀.5₅Sc₀.4N₀.5 films were measured in conventional Bragg-Brentano geometry in a high intensity Supratech XPert MRD diffractometer between 2θ = 10° and 2θ = 80° using the Kα1,2 doublet of a β-filtered Cu anode radiation. Because all the samples were strongly c-axis oriented, the rocking curves (RC) around the (00-2) peak were also measured to assess the quality of the films through the value of their full width at half maximum (FWHM). The XRD pattern of a representative Al₀.5₅Sc₀.4N₀.5 (x = 0.075) is depicted in Fig. 3.

For assessing the electroacoustic performance of the material, bulk acoustic wave (BAW) test devices were fabricated by depositing a Mo top electrode on top of the AlScN film, deposited on the acoustic reflector described above. The electroacoustic properties of Al₀.5₅Sc₀.4N₀.5 films were assessed by measuring the electrical reflection coefficient (S₁₁) at frequencies ranging from 10 MHz to 10 GHz using an Agilent PNA N5230A network analyzer. The electrical impedance of the resonators, derived from S₁₁, was fitted with Mason’s model. The non-optimized reflector used had an acoustic transmittance of -30 dB at the central frequency, which yielded resonators with a quality factor (Q) of more than 300. Despite the low value of Q, an accurate fitting of the impedance with frequency was achieved. All the geometrical parameters of the resonators, thicknesses of all the layers and device area, as well as all the material properties of reflector components and electrodes must be perfectly known in order to
obtain reliable values of the dielectric constant \((\varepsilon)\) and longitudinal acoustic velocity \((v_L)\) of the piezoelectric layer. The thicknesses of the reflector, layers, the electrodes and the piezoelectric film were carefully measured with a Veeco DekTak 150 profilometer in each sample (after the electrical characterization) by sequentially patterning all the layers, which includes the wet etching of the \(\text{Al}_0.5\text{Sc}_0.5\text{N}_{0.8}\) films in warm KOH using the Mo top electrode as hard mask. The area of each device was also carefully measured with an optical microscope with accuracy better than 1% to assess possible undercutting effects in the top electrode patterning. Mass density was derived from X-ray grazing angle reflectometry using RCRef-Sim (IHP) software package for fitting.

III. RESULTS AND DISCUSSION

Among all the possible process parameters that can be varied to optimize the sputter process, we have investigated the effect of the AC power applied to the targets and the ionic bombardment the film received during deposition. We will analyze two pairs of samples: (i) two samples deposited at low AC power under high and low ionic bombardment (H-IB and L-IB) respectively, and (ii) two samples deposited without ionic bombardment at high and low AC power (H-Pw and L-Pw), respectively. The first pair of samples was additionally pre-heated before film deposition.

A. Composition

The Sc content does not vary significantly neither with the different sputter conditions nor with the position on the wafer, displaying a mean value of \(7.5\pm0.5\%\) (Fig. 4). RBS measurements do not provide better accuracy.

B. Structure

Figure 5 shows the diffraction angle of the \((00\ 2)\) planes for the four samples analyzed as a function of their position on the wafer. It is worth noticing that the \((00\ 2)\) diffraction angle shows a significantly different value and evolution across the wafer for the four samples analyzed. Two of them, H-Pw and L-Pw, show a double peak in the zone near the center of the wafer. This indicates that, although the Sc concentration is approximately the same, the way this element is bonded in each sample is different. The ionic bombardment seems to homogenize the Sc binding. As for the rocking curves, except those of the sample deposited at higher power, they are very narrow with FWHM in the range of 1.5° to 1.6°, which are comparable to those obtained in pure AlN.

C. Electroacoustic response

The electrical response of the test resonators described in the experimental section were unambiguously fitted with Mason’s model, taking as variable parameters the electromechanical coupling factor \((k^2)\), the dielectric constant \((\varepsilon)\), and the longitudinal propagation velocity \((v_L)\) of the \(\text{Al}_0.5\text{Sc}_0.5\text{N}_{0.8}\). The mass density derived from X-ray reflectometry is identical for all samples and around 3150 kg/m³, 95% of that of pure AlN. For the rest of the materials (Mo, Al, SiO₂) we have used the physical parameters deduced in previous works through the assessment of BAW devices based on pure AlN.

The incorporation of Sc atoms into the AlN lattice has a strong influence on \(k^2\), \(\varepsilon\) and \(v_L\). The relative dielectric constant increases to values ranged between 11.7 and 12.2, which are the same as reported in \([6, 9, 12]\). The values of \(v_L\) vary from 10300 m/s to 10600 m/s. The determination of these two material constants is directly affected by the measured thickness of the piezoelectric layer, which has a considerable error (~5%). Therefore the observed variations inside the wafer and from wafer to wafer are not significant.

The material \(k^2\) is not so dependent on other parameters like the thickness. Therefore, its value can be calculated from the fittings more reliably. In Fig. 7, the variations of the \(k^2\) values...
as a function of the position on the wafer for the four analyzed samples are shown.

![Figure 7](image)

**Figure 7.** $k^2$ value as a function of the position on the wafer for the films analysed.

The expected value for this parameter is about 7.4% [6, 12]. It can be seen that this value is only reached by two samples in the edge of the wafer. This anisotropy in $k^2$ can be related to the anomalous binding of Sc to the AlN network as sensed by the XRD 00-2 reflection shown above. A non-homogeneous distribution of Sc into the material could originate inversion domains, which reduce the piezoelectric activity. In order to explore this issue, we have carried out several annealing processes in the samples to try to homogenize the Se distribution and type of bonds. This improvement in the film characteristics was previously observed by Wang et.al. [8], although they did not explain the origin of such improvement. The variation of $k^2$ of the sample (L-1B) before and after a 650°C, 10 min annealing is represented in Fig. 8.

![Figure 8](image)

**Figure 8.** $k^2$ values of some films after a heat treatment after deposition.

We think that the heat treatment helps the Sc atoms to reorganize into the AlN matrix allowing a better piezoelectric response. More work in this way is needed to better understand and optimize the deposition process.

IV. CONCLUSIONS

Sc-doped AlN films with Sc contents around 7.5% atomic have been deposited by AC reactive sputtering of Al targets with Sc pieces inlaid into them. We have studied the uniformity of the material properties on 150-mm Si wafers widely used in mass production of electroacoustic devices. Good film homogeneity in thickness and composition was confirmed. We found that ionic bombardment affects the way how Sc atoms incorporate into the AlN matrix, leading to radially heterogeneous film properties. Deposition at elevated temperatures as well as post-deposition heat treatment seems to be effective solutions to solve this issue.

ACKNOWLEDGMENT

This work was partially supported by the Ministerio de Economía y Competitividad del Gobierno de España through project MAT2013-45957-R.

REFERENCES


Micro OXCO EWOS-0513: A 20 years space odyssey up to 67P/Churyumov-Gerasimenko

Philippe Guillemot, Gilles Cibiel, Toulouse, France
CNES – French Space Agency
Toulouse, France
Philippe.guillemot@cnes.fr

Yves Richard, Jean-Marie Tarot, Guy Richard
Syrlinks
Rennes, France

Abstract—The story probably started somewhere on the planet Mars in the middle of the 90’s when the Sojourner rover demonstrated that it was possible to make great space missions with ‘low cost’ systems. It was the beginning of the ‘Better / Faster / Cheaper’ period. It led space agencies and space industry to develop space equipment using ‘professional’ systems instead of full space qualified ones.

In 1995, the CNES, the French Space Agency, decided to develop a family of ‘low cost’ space miniaturized OCXOs. A so called EWOS-0500 micro OCXO, used up to then for distress beacons, was selected for this purpose. The EWOS-0500 was a very small size (DIL 14, 1.5 cm³) and low power (150 mW) OCXO with a short-term stability in the 10⁻¹¹ range (A-Dev) and a frequency stability of 0.2 ppm in the temperature range [-30 °C; +60°C]. A specific qualification program was set up, to demonstrate the performances and the capability of this OCXO to fulfil space missions. The EWOS-0513 micro OCXO was born.

The EWOS-513 was embarked on numerous space missions, mainly in low earth orbit. But some of them were also embarked on the Rosetta mission for a 10 years journey through the solar system up to the 67P/Churyumov-Gerasimenko comet. Once arrived, they contributed to the success of the mission, allowing telecommunication between the orbiter and the lander and being involved in the science program through the CONSERT instrument.

This paper redraws the main steps and the main performances of this program that led an oscillator initially intended for distress beacons to contribute to the success of the Rosetta mission.

Keywords—ocxo; space applications

I. INTRODUCTION

In the middle of the 90’s, in the track of success of the Sojourner mission, the space industry changed his mind with a renewed interest for smaller satellites (in opposition to the large telecommunication payloads). That led all space companies to a great effort in order to miniaturize space equipment and to reduce weight, power consumption and cost.

The principal way explored was the use of professional components rather than space qualified ones. That gave access to more integrated technologies authorizing significant profits mainly on costs and delays of development. That also brought to reconsider the methods of qualification of space equipment. Traditional methods, based on the individual qualification of each component or process and on the traceability of the suppliers were not applicable any more. Taking into account what was done in industry, in particular aeronautical and automobile, it was decided to use the Highly Accelerated Tests method [1] to qualify space equipment.

First products to be developed at CNES from such an approach were miniature OCXO derived from those produced on a large scale for ARGOS/SARSAT beacons. Since then, 2 generations of micro-OCXO were developed and qualified. They fly on several space missions, in low Earth orbit as well as through the Solar System.

II. HIGHLY ACCELERATED TESTS

The Highly Accelerated Tests gather a set of methods which main objective is seeking where and how failures occur, investigating real design margin and using that information to improve product reliability or make better component selections within the shortest delay and at the lowest cost.

Fig. 1. The Highly Accelerated Tests and operating domain.

The method of the Highly Accelerated Tests made its entry in space field to answer, at lower cost and within the shortest deadlines, to the double objective:

- To explore the limits imposed by technologies, which makes it possible to determine the limits of operation of the device (i.e. beyond the theoretical design margins – Fig. 1) and, if necessary, to push them back by new technological choices when the margins appear insufficient compared to the specifications;
To detect as soon as possible (so that they can be corrected) the chargeable causes of defects, these causes being inherent in errors of design or in an insufficient control of the manufacturing processes, and not to technologies.

These objectives are achieved while applying to the product increasing levels of constraint, up to the failure or nearby. Constraints are determined thanks to failure mode analysis and in regards with the use of the product and its environment. They can be divided into 4 main groups:

- Mechanical constraints: To stress the structure,
- Electrical constraints: To stress components,
- Thermal constraints: To stress the assembly,
- Life test: To test the reliability.

In the case of space applications it is also advisable to take into account a fifth point which are space radiations.

That led us to define a test routine as presented Fig. 2 [2]. The levels of constraints are defined according to the missions requirements, without necessarily arriving up to the failure of the oscillator. It is sufficient to demonstrate the existence of sufficient margins.

![Typical Highly Accelerated Tests flow.](image)

III. THE EWOS 513

First products to be developed at CNES from such an approach were miniature OCXO derived from those produced on a large scale for ARGOS/SARSAT beacons by SYRLINKS (Formerly SOREP, then THALES Microelectronics and TES).

The EWOS 0513 is a very small size and low power OCXO especially designed for space application [3]. It is mainly made up of a quartz crystal resonator and a specific integrated circuit (ASIC). The crystal is a 10 MHz, fundamental mode AT cut resonator with a Q factor of at least 30 000, housed in a SMD ceramic package. Quartz crystal is not swept. In addition to being the oscillator loop and the output amplifier, the ASIC, directly stuck on the resonator’s package, is used for heating and controlling the temperature of the crystal unit and the electronic. The EWOS 0513 is housed in a hermetic metallic DIL package.

![Micro-OCXO EWOS 513](image)

A first generation of EWOS 0513 was developed at the end of the 90s’, with 12 OCXO involved in the qualification process [2]. Tests were successful, expect for radiations: Tests at high cumulative dose (50 kRad) with high dose rate showed a strong degradation of the short term noise. The origin of this degradation was identified as being related to the technology used for the ASIC. This degradation was however compatible with applications such as frequency reference for receivers and transmitters for remote control and telemetry.

<table>
<thead>
<tr>
<th>Item</th>
<th>Requirement</th>
<th>Type</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>10</td>
<td>10</td>
<td>MHz</td>
</tr>
<tr>
<td>Tuning</td>
<td>±5</td>
<td>±7</td>
<td>ppm</td>
</tr>
<tr>
<td>Power supply</td>
<td>5</td>
<td>-</td>
<td>V</td>
</tr>
<tr>
<td>Consumption</td>
<td>300</td>
<td>220</td>
<td>mW</td>
</tr>
<tr>
<td>Stability (vacuum)</td>
<td>Allan (1s)</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>In [-30/ +60 °C]</td>
<td>0.10</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Slope/mn (with 0.5 °C/mn)</td>
<td>±10</td>
<td>±6</td>
</tr>
<tr>
<td></td>
<td>Per day</td>
<td>1</td>
<td>10^-6</td>
</tr>
<tr>
<td></td>
<td>Per year</td>
<td>0.25</td>
<td>10^-9</td>
</tr>
<tr>
<td>Phase noise @ 10 MHz</td>
<td>10 Hz</td>
<td>-109</td>
<td>dB/Hz</td>
</tr>
<tr>
<td></td>
<td>100 Hz</td>
<td>-136</td>
<td>dB/Hz</td>
</tr>
<tr>
<td></td>
<td>1000 Hz</td>
<td>-148</td>
<td>dB/Hz</td>
</tr>
<tr>
<td>Mass</td>
<td>4.2</td>
<td>g</td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>1.6</td>
<td>cm^3</td>
<td></td>
</tr>
<tr>
<td>Operating Conditions</td>
<td>Temperature</td>
<td>[-30, +60]</td>
<td>°C</td>
</tr>
<tr>
<td></td>
<td>Random vibrations</td>
<td>0.732</td>
<td>g^2/Hz</td>
</tr>
<tr>
<td></td>
<td>Radiation</td>
<td>100</td>
<td>kRad</td>
</tr>
</tbody>
</table>

A second generation was developed at the beginning of the 2000s’: The main difference between the two generations of EWOS 0513 is the ASIC. Because of the obsolescence of the current one, a new ASIC was developed, with a new design. Even if this new design was not really rad hardened, results
from the radiation tests held on the first generation of OCXO were taken into account resulting in a less sensitive OCXO. That led us to start a new qualification process on the new EWOS 0513/B. Twenty four oscillators were implied [4]. Once again tests were successful, this time, including radiations.

![Graph showing Δf/f Irradiations vs. time for different radiation rates.](image)

**Fig. 4.** EWOS 513B (0515/018) : Global behavior (top) and short term stability (bottom) during radiation tests up to 100 krad.

IV. FROM LOW EARTH ORBIT TO THE SOLAR SYSTEM

A. Micro-OCXO for Micro Satellite

A soon as they became available, EWOS 513 have been used on all CNES micro satellites (Myriade Platform). Since 2004, date of DEMETER launch, more than 100 flight models of EWOS 0513 are flying in low Earth orbit, either

- As clock for the on-board computer and to generate the on-board time. For some mission (PICARD, TARANIS, thanks to some specific ground-to-space synchronization loop, it allows on-board time accuracy lower than 1 µs (usual value is rather 1 ms).
- Or as frequency reference for the S-Band receivers/ transmitters for remote control and telemetry.

<table>
<thead>
<tr>
<th>Mission</th>
<th>Launch</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARASOL</td>
<td>2004</td>
<td>Polarization and Anisotropy of Reflectance for Atmospheric Science coupled with Observations from a Lidar.</td>
</tr>
<tr>
<td>ESSAIM</td>
<td>2004</td>
<td>Constellation of 4 microsatellites for analysis of the electromagnetic environment (military use).</td>
</tr>
<tr>
<td>SPIRALE</td>
<td>2009</td>
<td>Preparatory System for Infrared</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mission</th>
<th>Launch</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PICARD</td>
<td>2010</td>
<td>Monitoring of the solar diameter, the differential rotation, the solar constant, etc... and their variations.</td>
</tr>
<tr>
<td>ELISA</td>
<td>2011</td>
<td>Constellation of 4 microsatellites for analysis of the electromagnetic environment (military use).</td>
</tr>
<tr>
<td>SSOT</td>
<td>2011</td>
<td>Sistema Satelital para la Observación de la Tierra (Satellite for Earth Observation)</td>
</tr>
<tr>
<td>VNRedSat</td>
<td>2013</td>
<td>Vietnam Natural Resources, Environment and Disaster Monitoring Satellite</td>
</tr>
<tr>
<td>TARANIS</td>
<td>2016</td>
<td>Tool for the Analysis of RAdiations from lightNings and Sprites</td>
</tr>
<tr>
<td>MICROSCOPE</td>
<td>2016</td>
<td>fundamental physics experiment to test the general theory of relativity.</td>
</tr>
</tbody>
</table>

B. Through the Solar System

The EWOS 0513 made a first raid through the Solar System between 2005 and 2010 with the NASA space probe Deep Impact. Micro–OCXOs were used to drive the Rx/Tx system that allowed the communications between the probe and the impactor.

But the most impressive journey is the one done by EWOS-0513 with the ESA probe ROSETTA. Launched in 2004, the probe reached the 67P/Churyumov-Gerasimenko comet in 2014. After of few months in orbit around the comet, the PHI LAE lander was dropped towards the comet surface where she lands successfully for a first exploration of a few days.

EWOS 0513 are used into 2 equipment of the mission, both of them operating between the PHI LAE lander and the ROSETTA probe in orbit around the comet:

- CONSERT (Comet Nucleus Sounding Experiment by Radio wave Transmission) is a time domain transponder [5]. A radio signal is transmitted from the orbiting component of the instrument and passes through the comet nucleus to the component on the comet surface. The signal is received on the lander, where some data is extracted, and then immediately re-transmitted back to the orbiter, where the main experiment data collection occurs. The variations in phase and amplitude that occur as the radio waves pass through different parts of the cometary nucleus will be used to perform tomography of the nucleus and determine the dielectric properties of the nuclear material. CONSERT operations imply some specific requirements of the oscillators. The main one is that...
the 2 oscillators, on the lander and on the orbiter, shall have frequencies as close as possible with the same long term drift, either powered on or not, such as the system is still able to operate after a 10 years journey. That leads us to conduct some specific tests and sorting on ground in order to select good candidates for the mission.

- S-Band Rx/Tx is a receiver / Transmitter that will allow the communication between the lander and the orbiter all along the comet phase of the mission. It plays a major role in the mission: Without it, the PHILAE module would be deaf and silent. Here also the main constraint is that, after a 10 years journey, Rx and Tx are still able to synchronize themselves, that is to say that the relative frequency of their OCXO is still within +6 / -10 ppm [6].

Even if analysis of data is still on going, Rosetta mission is already a great success. It is also a success from the OCXO point of view: Both equipment, Rx/Tx and CONSERT instrument, worked perfectly, the first one allowing communication between the orbiter and the lander all along the descent, the landing and the few days of operation on the comet (the relative frequency drift between Rx and Tx clock is less than -8 Hz) and the second one providing a huge amount of data about the internal structure of the comet (the relative frequency drift between the ROSETTA module and the PHILAE module is negligible, within the measurement noise).

V. CONCLUSION

The EWOS 0513 program demonstrates the possibility to use a low cost / Highly Accelerated Tests approach to develop space equipment. Two generations of OCXO were developed and more than 100 of them are flying around the Earth of farther in the Solar System. But this success is clearly limited to Science Mission, where high reliability and long lifetime are not the key factor.

To go farther, SYRLINKS and CNES decided, almost 20 years after the development of the EWOS 0513, to develop a third generation of micro-OCXO. This new OCXO, named EWOS 0530, will this time use full “space qualified” parts and processes in order to address markets more demanding in terms of level of quality. In the meantime, micro-OCXO should also fly with the new Myriade Evolution platform.

REFERENCES

[1] “Guide for defining and performing highly accelerated tests”, RG.AERO.000.029, BNAE, August 2003
Noise modeling methodology of an integrated circuit for quartz crystal oscillator

Vorobyev Nikolay\textsuperscript{a,b,\dagger}, Imbaud Joël\textsuperscript{a}, Baron Thomas\textsuperscript{a}, Cibiel Gilles\textsuperscript{b} and Galliou Serge\textsuperscript{a}

\textsuperscript{a}Time and Frequency Dept., FEMTO-ST Institute (UMR 6174, CNRS, UFC, ENSM, UTBM), Besançon, France

\textsuperscript{b}Microwave and Time Frequency Dept., CNES, Toulouse, France

\textsuperscript{\dagger}Labex FIRST-TF

E-mail: nikolay.vorobyev@femto-st.fr

Abstract—This paper reports the methodology and results of noise modeling and tests of a quartz crystal oscillator based on a dedicated integrated circuit (ASIC). Unfortunately, the phase noise of the inner sustaining amplifier is unknown. The challenge is to evaluate the performance limits of this ASIC in an oscillator circuit.

The oscillator phase noise computation results from the knowledge of individual intrinsic noise of the sustaining amplifier and the resonator respectively are presented. Moreover the simulation results are compared with the oscillator phase noise measurements.

Keywords—Noise; ASIC; quartz crystal oscillator; simulation; amplifier

I. INTRODUCTION

The design of a new generation of micro-OCXO oscillators for space applications requires to use reliable components. The ASIC designed for the new oscillator generation has validated its radiation-resistant quality within many space missions. Also it has a good thermal stability due to the integrated heating transistors and the temperature-control unit of crystal and electronics [1-2]. Furthermore this ASIC secures a well-known phase noise figure for the case of 10 MHz AT-cut quartz crystal resonators. It has been decided to clarify the possibility of using this ASIC with another resonator type and at a different frequency.

Referring to Leeson’s model [3-5], we can express the relationship between the frequency-stability floor of the oscillator (\(\sigma_y(\tau)_{floor}\)), the resonator loaded quality factor (\(Q_L\)) and the amplifier flicker noise coefficient (\(b_1\), the value of the \(f^1\) part of \(S_1(f)\) at \(f = 1\)Hz, where \(f\) is the offset frequency):

\[
\sigma_y(\tau)_{floor} = \frac{1}{Q_L} \sqrt{\frac{(b_1)_{floor} + 3}{10^{10} \ln 2} \cdot \frac{2}{2}} \tag{1}
\]

A good 10MHz SC-cut quartz crystal resonator usually exhibits an unloaded quality factor of 1.3 million. In the best cases, the loaded value of quality factor is about 50% from the unloaded one. This means that to obtain a frequency stability floor at the level of \(10^{-12}\), the amplifier flicker noise coefficient has to be better than -110dBc. For a 10MHz oscillator it is easy to reach these values, but in our case we fix the objective to increase the frequency. Therefore, a commercial 20MHz quartz crystal resonator with an unloaded quality factor of 0.5-0.6 million has been chosen. In this condition, to achieve the same frequency stability of level \(10^{-12}\), the amplifier flicker noise has to be better than -120dBc.

The first step of the study was to create a realistic noise model of the sustaining amplifier in the CAD system. There are many approaches to obtain the noise model of a single transistor or amplifier [6-9]. But in our case the sustaining amplifier is an ASIC whose internal design is just partially known and with very restricted electrical inputs-outputs. This fact considerably increases difficulties of the parametric characterization process. The analysis of the core of the internal circuit shows that its noise modeling can be performed by defining four noise parameters. Thus, a set of four external configurations is designed to extract experimental data. These data allow to compare the measured noise to the simulated one and to identify the unknown parameters by an engineer tool like Advanced Design System (ADS).

Moreover the noise of usual grade BAW (Balk Acoustic Wave) resonators, commercially available, has been measured. Passive measurements of the quartz crystal resonator phase noise are conducted in our laboratory by using an interferometric bench [10-11]. The resonator noise model is matched according to the measurements and implemented in the simulation oscillator circuit [12].

II. ASIC NOISE MEASUREMENT AND SIMULATION

A. Unknown parameters determination

Fig. 1 presents the known part of the ASIC. This is a Colpitts oscillator type [13-14] with a Cascode amplifier [15]. The cascode type amplifier increases the amplifier cut-off frequency (by decreasing the Miller feedback capacitance). The circuit consists of a power supply \(V_{CC}\), two BJT transistors \(VTT1\) and \(VTT2\), collector resistance \(R_c\), one emitter resistance \(Re\) and two load capacitances \(C1\) and \(C2\). Two current sources \(I_{bias1}\) and \(I_{bias2}\) ensure the forward active operation region of transistors.

There are mainly five flicker noise sources in this circuit as described below:
a) Two transistors VTT1 and VTT2.

The SPICE BJT transistor model is known but, it does not include the noise model constants. The power spectral density of the base current of the bipolar transistor corresponds to the following equation:

\[ S_b(f) = K_f \frac{f^{rac{1}{2}}}{f} + 2I_b \]  

(2)

This model involves three coefficients \(K_f\), \(A_f\) and \(F_{fe}\) to adjust the flicker noise. The constant \(F_{fe}\) is taken equal to one (flicker noise 1/f). The two constants \(K_f\) and \(A_f\) remain unknown parameters.

b) Power supply Vcc

The phase noise of the internal power source can be described as:

\[ \phi(f) = K_{Vcc} \frac{f}{f} \]  

(3)

The flicker noise constant \(K_{Vcc}\) is unknown.

c) Two bias current sources \(I_{bias1}\) and \(I_{bias2}\)

\(I_{bias1}\) and \(I_{bias2}\) are derived from a common reference voltage, internal to the ASIC, which justifies the hypothesis of the same noise source for both. Therefore, the phase noise spectrum described by the same law like in the previous case (3) with a noise constant \(K_{bias}\).

Finally, to obtain the realistic noise model of sustaining amplifier, we need to know four constants.

B. Amplifier noise measurements

The available input/output ports of the ASIC are limited (Fig. 1). So the direct measurement of the noise figure of each element is inaccessible. Nevertheless, four measurement set up have been used (Fig. 2):

a) A feedback closed configuration. \(C_L\) and \(R_e\) are connected. \(B_{VT2}\) is the input and \(R_i\) is the output.

b) A feedback-free configuration. \(C_L\) and \(R_e\) are disconnected. \(B_{VT2}\) is the input and \(R_i\) is the output.

c) A feedback free configuration with an AC gain increased. \(C_L\) and \(R_e\) are connected. \(B_{VT2}\) is the input and \(R_i\) is the output.

d) A direct feedback loop noise measurement. \(C_L\) and \(R_e\) are connected. \(C_L\) is the input and \(R_i\) is the output.

The experimental set up, shown in Fig. 3, has been used to measure the amplifier phase noise. This is an interferometric bench measurement with carrier suppression. The signal source frequency \(\nu_0\) is a 10 MHz reference oscillator. Depending on amplifier configurations (DUT), inputs and outputs are matched at 50 Ohms. Topology of this kind of bench measurement implies the use two identical DUTs simultaneously. Therefore, -3 dB has been subtracted from the measurement results to obtain the noise of one DUT.

C. Amplifier noise modeling

Four amplifier measurement configurations have been assembled in the same ADS simulation. It gives four unknown parameters and four equations to solve. ADS harmonic balance simulations allow the calculation of the phase noise. The simulation goal for each circuit is implemented from the measurements. The optimization process automatically finds the values of the unknown parameter within predetermined
limits.
To check the uniqueness of computation results, the initial values of seeking parameters have been changed several times for different limits.

III. OSCILLATOR NOISE PREDICTION

Based on the obtained amplifier model, a 10 MHz oscillator has been simulated in order to confirm this model. The oscillator model includes not only the amplifier noise model, but a realistic state of the art 10 MHz resonator model [12] as well, which has been previously measured and implemented in simulation.

The oscillation loop has been isolated from the load to match impedance and to purify the output signal. In addition to the amplifier noise and the resonator noise, thermal noise sources of all components are enabled. The flicker noise of transistors used in the buffer stages has been measured previously and implemented in the simulation model. Fig. 4 shows the simulated phase noise PSD and the measured one of the 10 MHz ASIC oscillator. The simulated curves correspond to actual measurements.

Second simulation and realization with a commercial 20MHz resonator have been carried out. Fig. 5 presents the phase noise measurements of a commercial 20 MHz SC-cut resonator. This phase noise measurements are conducted by classical method described in several publications [11, 16-17]. The quality factor of the measured resonator is about 0.5 million and the Leeson’s frequency equal to 50 Hz which give a loaded quality factor of 0.2 million. The Leeson’s frequency is masked by the spurious spikes caused by the reference source (HP8643A synthesizer) and 50 Hz perturbation. The phase noise of the passive measured resonator is \( L(1Hz) = -120 \) dBc. In the case of a 50% loaded quality factor, the phase noise of a 20 MHz resonator coupled with the ASIC amplifier to make an oscillator, give a \( L(1Hz) = -90 \) dBc after the \( f' \) to \( f'' \) slopes conversion caused by Leeson effect.

This has been confirmed by ADS phase noise simulation and by a direct measurement realized oscillator. Fig. 6 presents the results of this simulation and measurement. It should be noted, that the simulation results correlate very well with the measurements. The difference between them for offset Fourier frequency lower than 5 Hz is caused by thermal drift which is not included in the theoretical resonator model.

IV. CONCLUSION

The theoretical computation results are in very good agreement with the oscillator phase noise measurements: their difference is less than about 3 dB. As a consequence the developed model allows parametric and structural optimization of the electrical environment of the ASIC as well as resonator-type selection.

ACKNOWLEDGMENT

The authors would like to thank the Centre National d’Etudes Spatiales (CNES) and the LABEX First-TF network.
(funded by the Agence Nationale de la Recherche (ANR) Programme d’Investissement d’Avenir (PIA)) for funding the PhD student salary. We also thank Fabrice Sthal, professor at the Ecole Nationale Supérieure de Mécanique et des Microtechniques, for his knowledge of the interferometric bench and Ahmed Bakir, engineer at FEMTO-ST for the resonator noise measurements.

REFERENCES


The Prediction, Simulation and Verification of the Phase Noise in Low-Phase-Noise Crystal Oscillator

Xianhe Huang*, Junjie Jiao, Fuyu Sun and Wei Fu
School of Automation Engineering
University of Electronic Science and Technology of China
Chengdu, China
xianhehuang@uestc.edu.cn

Abstract—In order to achieve the prediction of the phase noise of low phase noise crystal oscillator, based on the classic phase noise model of Leeson, the load Q value ($Q_L$) is calculated according to the selected oscillator circuit parameters. Thus, on the basis of Leeson phase noise formula, the predicted results of the phase noise of low phase noise crystal oscillators are obtained. Then, the nonlinear transistor model is constructed to simulate the phase noise of low phase noise crystal oscillator by using the ADS (Advanced Design System) simulation software of Agilent and obtain the simulated curve of the phase noise. At last, practical measurement has been performed on these low phase noise crystal oscillator prototypes. The measured results show that: the predicted phase noise of the oscillators and the ADS simulation results obtained by using nonlinear transistor model are both close to the actual measured phase noise, which are at 100Hz and far away offset the carrier frequency. After that, the existence of the deviation, which is near carrier frequency, is analyzed. The prediction and simulation methods given by this paper might be beneficial to simplify the design progress of the low phase noise crystal oscillator.

Keywords—prediction; simulation; phase noise; crystal oscillator

I. INTRODUCTION

Phase noise is very important parameter for crystal oscillator, however, the specific phase noise will not be know for sure before actual oscillator is made and measured. In the engineering design process, it is meaningful to predict and simulate the phase noise of crystal oscillator. This paper shows the prediction and simulation of phase noise of three crystal oscillators.

II. THE PREDICTION, SIMULATION AND VERIFICATION OF THE PHASE NOISE OF LT10.7 MHZ CRYSTAL OSCILLATOR

LiTaO$_3$ crystal has good electro-optical, pyroelectric and piezoelectric properties. The unloaded quality factor of LiTaO$_3$ resonator is much higher than LC tank, and it has zero temperature coefficient cut type. Its electromechanical coupling coefficient is as high as 60%. So LiTaO$_3$ resonator has low capacitance ratio and it can be used in wide frequency offset voltage controlled oscillator [1].

By the analysis of Leeson formula derived from the phase noise model of Leeson, we can find that phase noise within the half-bandwidth of the loop decreases with increasing $Q_L$. The Leeson formula is as follows [2]-[7]:

$$S_{\phi\phi}(f_m) = \left[1 + \frac{1}{f_m^2} \left(\frac{f_m}{2Q_L}\right)^2\right] S_{\phi\phi}(f_m^0)$$

(1)

Where $f_m$ is offset frequency, $f_0$ is center frequency, $Q_L$ is loaded quality factor, $S_{\phi\phi}(f_m^0)$ is the power spectral density of the output phase noise, $S_{\phi\phi}(f_m)$ is the power spectral density of the oscillator input phase noise. So a high $Q_L$ is beneficial to reducing phase noise within the half-bandwidth of the loop.

The main parameters of LiTaO$_3$ crystal, adopted in this experiment, are illustrated in table I. The unload Q is about 1240.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shunt capacitance $C_0$</td>
<td>4.0×10$^{-12}$ F</td>
</tr>
<tr>
<td>Dynamic inductance $L_q$</td>
<td>3.8×10$^{-4}$ H</td>
</tr>
<tr>
<td>Dynamic capacitance $C_q$</td>
<td>5.791×10$^{-13}$ F</td>
</tr>
<tr>
<td>Equivalent resistance $R_q$</td>
<td>20.6 Ω</td>
</tr>
<tr>
<td>Frequency $f$</td>
<td>10.727 MHz</td>
</tr>
<tr>
<td>$C_0/C_q$</td>
<td>6.9</td>
</tr>
</tbody>
</table>

Based on the calculation in [9], the load Q of this sample in Butler oscillating circuit is approximately 413. According to Leeson formula, the predicted phase noise data of this LT Butler crystal oscillator is obtained, as shown in figure 1.
Fig. 1 Prediction phase noise graph of LT Butler crystal oscillator

ADS software of Agilent was adopted to simulate the phase noise, which includes the HB ActiveX to analyze nonlinear circuits. By using this ActiveX, we can simulate the phase noise of oscillators. As it is incorrect to use the transistor directly from ADS component library, which will lead to an erroneous phase noise slope, a nonlinear transistor model is established by our own during the simulation. After the nonlinear transistor model is employed, the phase noise simulation curve of this LT Butler crystal oscillator can be shown in figure 2.

Under the instruction of prediction and simulation, we measured the LT Butler crystal oscillator prototype. By using the E5052B Signal Source Analyzer of Agilent, the measured phase noise curve is shown in figure 3.

Compared figure 1-3, the measured phase noise characteristic is near the simulated curve and predicted result.

III. THE PREDICTION, SIMULATION AND VERIFICATION OF THE PHASE NOISE OF 100MHz AT-CUT 3RD OVERTONE CRYSTAL OSCILLATOR

At the same time, an AT-cut 3rd overtone 100MHz quartz crystal resonator was used to achieve a 100MHz low phase noise voltage-controlled crystal oscillator. The followed purpose is using this oscillator to realize direct temperature compensation, when the frequency is at 100MHz (the approach associated with frequency multiplication or adding series inductance is abandoned). The main parameters used in this experiment are illustrated in table II.

<table>
<thead>
<tr>
<th>Description of design</th>
<th>Equivalent Dynamic capacitance</th>
<th>Equivalent Resistor</th>
<th>Unload Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT 3rd overtone</td>
<td>1.2F</td>
<td>10.0Ω</td>
<td>132K</td>
</tr>
</tbody>
</table>

Pierce oscillating circuit is employed in this measurement. Based on the calculation from [10], the load Q of this crystal resonator in the given oscillating circuit is about 28000. In the same way, on the basis of Leeson formula, the predicted phase noise curve of AT-cut 3rd overtone 100MHz quartz crystal oscillator can be seen below in figure 4.

Figure 5 shows the simulation curve of the phase noise of the oscillator, using ADS software of Agilent, and the nonlinear transistor model is adopted.

TABLE II. THE PARAMETERS OF CRYSTAL RESONATOR

Fig. 4. Predicted phase noise graph of AT-cut 3rd overtone 100MHz quartz crystal oscillator

Fig. 5. Simulated phase noise plot of AT-cut 3rd overtone 100MHz quartz crystal oscillator
Under the instruction of the prediction and simulation, AT-cut 3rd overtone 100MHz quartz crystal oscillator prototype is measured. The actual phase noise curve of the oscillator, obtained by E5052 signal source analyzer, is shown in figure 6.

By comparison of figure 4-6, one can find that the measured phase noise characteristic is also close to the prediction and simulation results.

IV. PREDICTION, SIMULATION AND VERIFICATION OF THE PHASE NOISE OF 120MHz SC-CUT 5TH OVERTONE QUARTZ CRYSTAL OSCILLATOR

120MHz SC-cut 5th overtone quartz crystal oscillator is also adopted to do the similarly experiment. The major parameters of this crystal resonator are listed in table III.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equivalent Dynamic Inductance L_q</td>
<td>8.8 mH</td>
</tr>
<tr>
<td>Equivalent Resistor R_q</td>
<td>63Ω</td>
</tr>
<tr>
<td>Equivalent Dynamic capacitance C_q</td>
<td>2 × 10⁻¹⁶ F</td>
</tr>
<tr>
<td>Shunt capacitance C_0</td>
<td>3.2 pF</td>
</tr>
</tbody>
</table>

The unload Q of this crystal resonator is approximately 105K. The oscillating circuit adopted in this measurement is Pierce oscillating circuit. On the basis of calculation from [11], the predicted phase noise plot of 120MHz SC-cut 5th overtone quartz crystal oscillator is illustrated in figure 7.

Then, we obtained the simulated phase noise curve by using ADS software of Agilent and the nonlinear transistor model.

On the guidance of prediction and simulation, actual 120MHz SC-cut 5th overtone quartz crystal oscillator prototype is measured. The phase noise curve obtained by E5052B signal source analyzer of Agilent is shown in figure 9.
By making a comparison among figure 7-9, it shows that the measured phase noise characteristic is close to the prediction and simulation result.

V. COMPARISON AND ANALYSIS

The comparison among the prediction, simulation and measurement of the phase noise of three low phase noise crystal oscillators are illustrated in table IV.

One can see from table IV that the predicted phase noise of the oscillator by using Leeson model and the simulated phase noise, obtained by ADS software with nonlinear transistor model, are both near the actual measured phase noise at 100Hz and far away offset the carrier frequency. The offset frequency near the carrier is rather obvious. This might be caused by the nonlinear amplitude-frequency (AF) characteristic of the resonator or the temperature fluctuation of measure environment. Further study on this phenomenon is continued.

The prediction and simulation methods given by this paper might be beneficial to simplify the design progress of the low phase noise crystal oscillator.

<table>
<thead>
<tr>
<th>Oscillators</th>
<th>Type</th>
<th>specification and parameters of the resonators</th>
<th>prediction values of the oscillator’s phase noise</th>
<th>simulation results of the oscillator’s phase noise</th>
<th>measured results of the oscillator’s phase noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>100MHz low phase noise crystal oscillator</td>
<td>LT10.7MHz low phase noise crystal oscillator</td>
<td>C1=0.579pF, Q0=1.24K, C0=2.5pF</td>
<td>-89dBc/Hz/10Hz</td>
<td>-88dBc/Hz/10Hz</td>
<td>-85dBc/Hz/10Hz</td>
</tr>
<tr>
<td>120MHz low phase noise crystal oscillator</td>
<td>100MHz low phase noise crystal oscillator</td>
<td>Quartz, AT, 3th overtone</td>
<td>-99dBc/Hz/10Hz</td>
<td>-100dBc/Hz/10Hz</td>
<td>-92dBc/Hz/10Hz</td>
</tr>
<tr>
<td>120MHz low phase noise crystal oscillator</td>
<td>120MHz low phase noise crystal oscillator</td>
<td>Quartz, SC-cut 5th overtone</td>
<td>-106dBc/Hz/10Hz</td>
<td>-112dBc/Hz/10Hz</td>
<td>-104dBc/Hz/10Hz</td>
</tr>
</tbody>
</table>

REFERENCES

The Border Effect in Frequency Signal Processing and the Phase Measurement with Arbitrary Frequency Relationship

Wei Zhou, Lina Bai, Zhiqi Li, Faxi Chen, Xiaotian Cao, Yadong Duan, Xuyang Zhou, Longfei Xu
Dept. of Measurement and Instrument, Xidian University, Xi’an, Shannxi, China

Abstract—Based on the achievement of phase coincidence detection, the processing and elimination of quantization error is possible in digital frequency measurement. Because of the limited resolution of phase detection, a fuzzy area is to be formed. The higher precision can be obtained using the border stability of fuzzy areas. The decisive factor of the actual precision is the resolution stability of the border of fuzzy area, called the border effect. Therefore, utilizing the circuit with ns resolution one can obtain ps level or higher precision. On the basis of improving measuring resolution significantly and the phase variation regularity of periodic signals, through the frequency measurement with phase continuous, the measured frequency signal is compared with its theoretical nominal frequency by multiple periods and with measuring gate time one by one. Eventually, the measurement of phase variation of any measured signal in a wide frequency range can be realized. Compared with other measurement techniques and instruments it is a great advancement. Because of the phase information is obtained for the different even very complicated frequency signals, higher frequency measurement precision can be realized without any frequency transformation. It is possible to realize higher measurement and control precision.

Keywords—the border effect; the fuzzy area; the virtual reconstruction

1. BORDER EFFECT IN MEASUREMENT

Although the increasing measurement precision is mainly manifested in the measurement resolution improvement[2][3], the limitation of the actual measurement conditions, such as the resolution of the device itself, limited response time, the quantization error in digital processing, the identification of limited time to deal with for periodic signal and noise would have leaded inevitably the problems[4][5] with low resolution for all the measurement process.

Due to the generally existence and probability of measuring resolution, as well as the value of resolution stability is higher than itself, therefore, we can take advantages of the resolution stability to improve the measurement resolution. The measurement resolution of the component leads to the existence of the measurement fuzzy area near the true value. As shown in Fig. 1 and Fig.2:[6][7][8]

As shown in Fig.1, two compared signals. f1 and f2, are close and have a slight deviation in frequency, and V(t) is the phase difference between them[9][10]. The phase coincidence detection will produce two overlap pulse signals A and B[10][11], as shown in Fig.2 (a), in each least common multiple cycle Tnn. Pulse A may be one of any pulses in a, b and c in Fig.1; and pulse B is also the same. Enlarging the coincidence pulse A or B, we can observe the “real fuzzy area” in the measurement process, as shown in Fig.2(b).

![Overlap pulse from coincidence](image1)

(a) Overlap pulse from coincidence detection

Fig.2 the Result of phase coincidence detection

While perfect measurement result can be achieved at the measured value, however the fact that the position of the measured value is not always easily accurately obtained, makes it difficult to achieve the ideal accuracy. The detection resolution has more stability at the fuzzy borders than the other position in the fuzzy area. If measurement results periodic repeatedly appear during the process, through two times border of fuzzy area (point a, and point b as shown in Fig.3) for measurement, the error of the measurement value to fuzzy area border can be eliminated, (namely e1 and e2). Thus, the actual measured of L0 and L = [14][15], measuring at the border of fuzzy area, is almost equal, and eventually measurement error is the smallest, and much higher measurement precision can be obtained. The resolution can be improved two to three orders of magnitude using border stability of resolution in measurement accuracy than the resolution itself. If the border effect is used in the metering circuit with ns scale resolution, the actual measurement accuracy can reach the ps scale.[7][11]

![Fuzzy area from the detection overlap pulse](image2)

(b) Fuzzy area from the detection overlap pulse

Fig.3 Border effect in measurement process
II. FREQUENCY MEASUREMENT USING NEW PHASE PROCESSING

Phase is one of the most important parameters in all the periodic phenomena.\cite{6,8} Phase processing is widely applied in frequency measurement and control fields. The conventional phase processing can only be carried out between signals with the same nominal frequency and it is difficult to directly realize the phase measurement for signals with wide frequency range, even cannot be achieved in some cases.\cite{6,7,8} And a number of frequency transformation circuits have to be incurred and the development and research in relative fields are severely restricted by the condition \cite{7,8}. Therefore, a new approach realizing directly phase measurement with high resolution is proposed and it has been confirmed for wide frequency range.

A. the PRINCIPLE of NEW PHASE PROCESSING

Based on the border effect with high measurement resolution and the law of periodic variation of phase difference between periodic signals,\cite{1,2} we confirmed and completed an phase measurement method which can directly obtain phase difference without any frequency transformation in any frequency range, and which is a completely different from the traditional method of phase compared approaches. The waveform principle and experiment platform are shown in Fig.4:

\begin{equation}
\tau_{\phi} = \tau - \phi
\end{equation}

\begin{align}
\tau_{\phi} = N_{x}T_{x} - \tau
\end{align}

Equation (3) is true under the condition of high precision of frequency standards comparison, namely $N_{x} = N_{sn}$.

Along with the extension of compared time, the totally accumulating phase variation $t_{t}(\theta)$ in continuous compared time can be described as (4):

\begin{equation}
t_{t}(\theta) = t_{1}(\phi) + t_{2}(\phi) + t_{3}(\phi) + \ldots + t_{i}(\phi)
\end{equation}

According to the theories mentioned above, we have obtained phase variation in arbitrary frequency range by the new phase processing method.

B. the VERIFICATION FOR PHASE PROCESSING

To confirm the practicability of the method proposed in this paper, relative experiments have been conducted. The concrete process is described as Fig.4(b). After signal processed, we apply two compared signals $f_{r}$ and $f_{x}$ to the input of the phase coincidence detector, depend on the stability of the border of fuzzy area or the stability of the coincidence detection resolution, namely the border effect, obtain the high precision phase coincidence pulse to trigger the reference gate to form the actual continuous measurement gate, and finally realize continuous measurement gate. Counting the cycles of $f_{x}$, we can get the $f_{sn}^{'x}$, which is the same frequency nominal value with the measured signal and the part of accumulating phase difference $t(\phi)$ can be obtained by the compared between $f_{x}$ and $f_{sn}^{'x}$ in the corresponding measurement gate $\tau$. Finally, a series of $t(\phi)$ can be obtained along with the extension of measured time.

C. ANALYSIS for EXPERIMENT DATA

In the process, we use the frequency generator HP8662A generates (12.8M+10)Hz signal as the measured signal $f_{x}$ and its inner frequency standard 10MHz as reference signal to input the measurement system to realize the self-calibration experiment. We obtain frequency stability of self-calibration experiment result by analyzing and processing the data, shown in the Fig.5.

![Fig.5 Frequency stability of crystal self-calibration](image)
The self-calibration can reach the resolution at $9 \times 10^{-12}/s, 8 \times 10^{-13}/10s, 8 \times 10^{-14}/100s, 8 \times 10^{-15}/1000s, 8 \times 10^{-16}/10,000s,$ and meet the presumptive requirement. Due to the compared time is not enough, the curve descend sharply at 30,000s; according to the former analysis, the system can reach the resolution of $10^{-17}/$day if the measurement time can last over 10days.

This result reflects that the phase information of measurement means can be applied in complex frequency control, on the other hand, it is also revealed that the method will play a major role in the development of precision measurement and control fields.

**CONCLUSION**

Phase information plays a quite important role in many science and engineering fields. Due to the width and complexity of periodic signals, direct detection of the phase variations for different frequency is impossible, or too complex according to traditional methods. The situation leads to complexity in many aspects, such as frequency control, signal processing and measuring equipment, broadband phase processing, quantum frequency standard and comparison, electronics, phase-locked loop, navigation and positioning, periodic physical phenomena exploration, communication and so on[16][17][18]. We also found that the phase processing method will play a vital role for many complex measurement problems. Therefore, the measurement method without frequency conversion directly obtains the phase information will be an unprecedented breakthrough in some areas of frequency and phase. It will produce revolutionary advances in many technological fields It should be emphasized that the far-reaching meaning of the method mentioned in the paper is that a novelty approach, namely reconstruction for the measured signal, is proposed to phase measurement with arbitrary wide frequency range.

**ACKNOWLEDGMENT**

This work is partially supported by the National Natural Science Foundation of China under Grant No. 61201288, the Open Fund of National Key Laboratory of Aerospace Dynamics under Grant No. 2013ADL-DW0402, Xi’an Science and Technology Plan Projects under Grant No. CXY1351(6), Shaanxi Natural Science Foundation Research Plan Projects under Grant No. 2014JM2—6128.

**REFERENCE**


A 250nm CMOS Low Phase Noise Differential VCO Circuit Without Varactors

Anatoly V. Kosykh, Konstantin V. Murasov, Aleksandr N. Lepetaev, Sergey A. Zavyalov
Omsk State Technical University
Omsk, Russia.
E-mail: avkosykh@omgtu.ru

Abstract— In the article the realization of the new circuit of differential VCO without varactors in 250 nm CMOS technology has been described. The tuning of VCO is realized as a circuit changing load of oscillator. The control curve for differential VCO without varactors is presented. The results of comparing of phase noise for circuits differential oscillators based on N-channel and P-channel transistors has been described. (Abstract)

Keywords: Quartz resonator; phase noise; CMOS; differential oscillator; excitation circuit; temperature compensation

I. INTRODUCTION

Quartz crystals are widely used in references of a frequency source required for various types of radio communication systems, telephony, data processing, etc. The most significant factor in the instability oscillator frequency using a quartz resonator is the temperature dependence of the resonance frequency of quartz. There are two basic techniques to minimize thermal instability of crystal oscillators: oven control or temperature compensation. Use of oven control is related with sufficiently high energy consumption and an increase in dimensions, which excludes the use of this method for integrated circuit frequency reference sources. Method of temperature compensation based on the control signal for the oscillator circuit depending on temperature. This method can significantly reduce the deviation of the oscillator frequency caused by the change in temperature.

II. THE STATEMENT OF THE PROBLEM

In most types of technological processes based on CMOS, the implementation of varactors having high Q, low resistance and the necessary restructuring of capacitance is not possible. Using as an integrated varactors gate capacitance of the MOS transistor [1] is associated with a large initial capacitance gate and the magnitude of relative tuning of not more than 50%. To provide the operation mode MOSFET gate as a varactor it is necessary that the gate potential would be smaller than potentials of electrically connected regions by the drain and the source of MOSFET. This requirement is realized by the introduction of the separation capacitance and supplying a bias voltage to the gate of the transistor. These actions lead to the reduction in the quality factor of MOS varactor and a relatively large size of the tuning system on a chip. Moreover, in a number of schemes oscillators because of the large initial capacitance, MOS varactors are not applied at all, because circuit will be overloaded.

In the case of necessity of increasing a bank of capacitances may be used, switched by MOSFET switches [2], or by the combination of a bank of capacitances and a MOS varactor [3].

III. THE DIFFERENTIAL QUARTZ OSCILLATOR

A fundamentally different way to control the frequency oscillator may be implemented by changing the load of an oscillator. On fig. 1 is a schematic diagram of a differential crystal oscillator with frequency oscillation of 10 MHz has been presented.

Fig. 1       Differential xtal oscillator circuit

In the article [4] it is shown the possibility of reducing the phase noise by replacing transistors of differential pair with a negative resistance on the transistors of opposite conductivity types of transistors, giving the advantage by 6 dB in the near field (fig.2).

Fig. 2        Phase noise oscillator circuit with P-MOS and N-MOS
In the n-MOS transistors the channel is induced in the surface layer of semi-conductor, while the transistors of opposite conductivity (p-MOS) channel are induced by immersion in a thickness of the semiconductor. As a result, the charge carriers in the n-MOS transistor, which are closer to the surface of the semiconductor, have a greater probability of falling in the trap of the oxide layer, which are one of the main sources of flicker noise.

The difference in mobility between the parameters of the holes and electrons results in a significantly lower probability of tunneling silicon oxide film holes than electrons that reduces the level of flicker noise in 5 ... 6 times. Higher density distribution of donor-type traps in comparison with the density of acceptor-type traps explains further the difference in the level of flicker noise transistors with different types of conduction channels.

On transistors P3, P5 (fig. 1) differential pair with a negative resistance is collected. It should be noted that the use of p-channel transistors as differential pairs allows to reduce the phase noise at the frequency offset 1 kHz to 6 dB. As a variable load circuit C1, N1 and C3, N4 have been used.

IV. VCO ANALYSIS

To analyze the properties of the oscillator we will use the representation of the tuning circuit (fig. 3) with variable resistor.

After analysis of the VCO circuit with a circuit configuration shown in Fig. 3 the dependence of the frequency of the VCO of the resistance have been obtained (fig. 4).

On figure 4. 2 areas can be identified. When the resistance increases from hundreds of ohms to 1 kOhm the increase in the frequency of the VCO with increasing resistance has been observed. With further increase in resistance the VCO frequency decreases.

The amplitude of voltage of variable resistor (fig. 3) is not more than 100 mV. Passing to the analysis of the circuit frequency control based on the MOS transistor, it is necessary to calculate the size of the transistor for changing the channel resistance within a certain range when changing the control voltage. By the results of the modeling of this circuit using method modeling oscillators [5], we define the control characteristic. The changing of the control voltage from 900 mV to 1.8 V leads to a continuously tunable oscillator output frequency of 3.5 kHz, which is more than enough to compensate for the temperature dependence of the oscillator frequency. VCO frequency dependence on voltage control (fig. 5) have 1/x^2 character, and frequency shift. This can be explained by the quadratic dependence of the channel resistance of the FET and the presence of parasitic capacitances.

Nonlinear characteristics of the frequency control of the generator in 3.6 kHz frequency range is 33%. After tuning circuit linearization, nonlinearity of the curve control decreases to 7.9 %. However, the tuning range was reduced to 1.6 kHz. With a high accuracy this characteristic can be approximated by a polynomial of the second degree, making this non-linearity compensated by the introduction of relevant amendments to the synthesizer of compensating function.

Unlike common VCO circuits, based on the Pierce oscillator circuit, the differential crystal oscillator circuit has a good temperature stability of its own, not exceeding 300 Hz in a temperature range of -60 ... 125 C at a commensurate level of phase noise, less than -155 dBc/Hz on offset frequency 10 kHz.
V. PRACTICAL IMPLEMENTATION

VCO has been implemented based on Si-Ge CMOS 250nm process. On fig. 6 the fragment of the topology of the crystal containing the designed VCO is presented. The active area of VCO is 0.0336 mm$^2$ (= 210 × 160 µm$^2$).

Fig. 6 VCO layout

REFERENCES


Novel gyroscopic mounting for crystal oscillator (payload) applied in high dynamic host vehicle (platform) to improve its output stability

Maryam Abedi, Tian Jin
School of Electronic and Information, Beihang University
Beijing 100191, China
E-mail: abedi_maryam@buaa.edu.cn

Abstract—when crystal oscillator applied in high dynamic host vehicle, dynamic loads impact its short and medium-term stability. In order to suppression and/or reduction these impacts, gyroscopic mounting is proposed to install oscillator on it.

One of the main parameters which affect dynamic loads impacts on oscillator output is crystal g-sensitivity vector. Therefore, at first, statistical study has been accomplished on this parameter for SC-cut crystals which used in high dynamic oscillators. This study reveals that the angle between this vector and crystal surface \( \phi \) nearly follows Rayleigh distribution with \( 1\sigma =19^\circ \). Thus, this vector tends to be close to crystal surface not perpendicular to it. Furthermore the analysis results of using gyro-mounting shows the acceptable results for quartz crystals with \(|\phi|<2\sigma=38^\circ\) (85%).

To prove gyro efficiency, its effect on each dynamic load is analyzed separately. The results of related analyses reveal that gyro shows its best for attitude change of host vehicle, such that it totally suppresses relevant disturbances. In the case of steady state load, sinusoidal and random vibrations, gyro best effects appear for \(|\phi|<30^\circ\) and \(\beta\) angles which are not too close to 90°.

Totally when high dynamic crystal oscillator is installed on gyro-mounting, the probability of observing instability on its output is reduced.

Keywords—crystal oscillator, SC-cut, high dynamic, short-term instability, frequency jitter, phase noise, random vibration, sinusoidal vibrations, steady state acceleration, gyroscope.

I. INTRODUCTION

![Fig. 1. Angular orientation of dynamic load (A) and g-sensitivity vector (Γ) of quartz crystal blank; (left) typical state, (right) critical state (β=0).](image)

When crystal oscillator is used in high dynamic host vehicle which could be spaceborne, airborne or seaborne, e.g. launch vehicle, aircraft, helicopter, submarine, etc., its output is influenced by a large number of environmental effects. One of which is mechanical environment in which crystal oscillator exposed to a variety of dynamic loads modulated on its output and degrade its short and medium-term stability [1]. In some applications, e.g. GNSS receiver [2], Short term instability of crystal oscillator degrades the performance of total system. Dynamic loads-induced instabilities appear as disturbances on signal frequency and phase (clock drift and clock bias/T).

A. Short-term instability (clock drift)

\[
\Delta f (\tau=1s) = f_0 \Gamma^A \cos \alpha \cos \beta \text{ (Hz)}; \\
\Delta \varphi (B_n=1Hz)= 2\pi \int \Delta f dt = 2\pi f_0 \Gamma^A \int A dt \text{ (rad)}; \\
\epsilon (f) = 20 \log \left| \frac{\Delta \varphi}{2} \right| \text{ (dBc/Hz)};
\]

Where: \(\Delta f=\) frequency disturbance; \(\Delta \varphi=\) phase disturbance; \(\epsilon(f)=\) phase noise; \(\Gamma^A=\) g-sensitivity vector of crystal blank (1/g); \(\alpha=\) Angle between \(\Gamma^A\) and \(A\); \(\beta=\) Angle between pages pass through \(A\) and \(\Gamma\); \(f_0=\) oscillator frequency (Hz) and \(T=\) mission period of host vehicle.

II. DYNAMIC LOADS

![Fig. 2. High dynamic host vehicle (M1, C1, K1), dynamic load \(\ddot{y}\), crystal oscillator installed on it (m2,c2,k2), dynamic load \(\ddot{y}\).](image)

Dynamic loads applied by high dynamic host vehicle on crystal oscillator (Fig.2) could be divided into [1, 4-10]:

1. Medium-term instability sources:

   **Attitude changes** of host vehicle on its trajectory.
2. Short-term instability sources:

   **Steady State Acceleration:** Time variant trust of host vehicle on longitudinal and lateral directions.

   **Sinusoidal Vibrations:** A series of low frequency sinusoidal vibrations.

   **Random Vibration:** A combination of bandlimited sinusoidal vibrations with random amplitude, frequency and phase.

III. G-SENSITIVITY VECTOR

Each crystal blank has private g-sensitivity vector (Fig.1). Apart from its magnitude ($\Gamma$), its angular orientation ($\phi$, $\theta$) influences the impact of dynamic load on crystal oscillator output signal (Eq.1-4).

Since in our study, angle $\phi$ (Fig.1) plays important role on induced instability on oscillator output, therefore Statistical study [11- 17] has been done on this angle for different SC-cut crystal blanks which used in high dynamic oscillators. According to this study, Distribution of this angle is as shown in Fig.3. As seen it nearly follows Rayleigh distribution with $1\sigma=19^\circ$ (36.36%), $2\sigma=38^\circ$ (84.85%), $3\sigma=57^\circ$ (98.48%) and $\phi_{min}=2.6^\circ$, $\phi_{max}=62^\circ$.

![Fig. 3. Distribution of angle $\phi$ for different SC-cut crystals](image)

In this paper, $|\phi|<30^\circ$ with distribution 64% supposed as the most probable $\phi$ values for chosen crystal oscillator (Fig.3).

IV. REDUCTION OF INSTABILITY CREATED BY DYNAMIC LOADS ON CRYSTAL OSCILLATOR OUTPUT

According to the equations 1-4, frequency and phase disturbances are functions of $\alpha$ and $\beta$ cosine (Fig.1), thus when A moves away from $\Gamma$, improvements give arise on oscillator output signal.

The first idea to suppress these disturbances is to fix the applied dynamic load perpendicular to $\Gamma$ in any given moment. Since the angular orientation of $\Gamma$ is different among a series of carefully designed crystals with the same cut, same vibration mode and same overtone resonant frequency, therefore the implementation of this idea is not feasible.

Therefore the main idea to reduce these disturbances is to hold the dynamic applied load perpendicular to crystal surface in any given moment. By this way, for specified range of angle $\beta$ (Fig.1), where $|\beta|<\cos^{-1} (\sin \phi /\cos \alpha)$, $\alpha$ is increased as much as possible. Then dynamic loads-induced disturbances are reduced as low as possible and consequently oscillator output stability will be improved.

V. GYRO-MOUNTING

In order to implement mentioned idea, a mounting is introduced for crystal oscillator to install it on electronic board and at the same time give it the freedom to rotate freely around roll, pitch and yaw. By this way, in each moment resultant applied load will be perpendicular to oscillator surface. Since the structure of this mounting is similar to gyroscope, it so-called gyro-mounting (Fig.4).

This mounting must be as fast as possible to respond to dynamic loads; therefore it should be manufactured of low-density materials as well as the softness of surfaces should be as high as possible.

Gyro-mounting is a passive component and doesn’t need to any electric equipment and power supply. Moreover it is neither expensive nor complicated.

After installation of crystal oscillator on it, dynamic loads-induced disturbances will be changed as:

\[ \Delta f_{gyro} (\tau=1s) = f_0 \frac{\Gamma A \sin \phi}{(Hz)} \]  
\[ \Delta \phi_{gyro} (Bn=1Hz) = 2\pi \int \Delta f_{gyro} dt \ (rad) \]  
\[ \xi (f) = 20 \log_{10} \frac{1}{2} \Delta \phi_{gyro} (dBc/Hz) \]

Where: $\Delta f_{gyro}$: frequency disturbance; $\Delta \phi_{gyro}$: phase disturbance; $\xi (f)$: phase noise.

It is worth remembering that max gyro effect appears in critical state (Fig.1) when the angle between A and $\Gamma$ is minimum value. In this case, gyro prevents of maximum probable instabilities in system.

I. IMPACTS OF DYNAMIC LOADS ON STABILITY OF CRYSTAL OSCILLATOR OUTPUT AND PROOF OF GYRO EFFICIENCY

In this paper to do numerical analysis, high dynamic oscillator assumed as GPS disciplined crystal oscillator installed on launch vehicle ARIAN. Since this kind of oscillator is used as primary frequency standard for calibration and metrology in laboratories. (OCXO, thickness shear mode, SC-cut, 3rd overtone, $f_0= 10.23 \ MHz$, $\Gamma=10^{-9} /g$, $\sigma_\alpha$ (Allan deviation) = $10^{-14}$).

A. Attitude change of host vehicle

When during mission period (T) of host vehicle, its attitude changes $\pi \tau$, gravity acceleration (g) plays role of dynamic load and makes disturbances on oscillator output. As well as in the
case of host vehicles which have ascent phase from ground surface (e.g. launch vehicles), as altitude increased, magnitude of g decreased therefore it appears as dynamic load. Thus when attitude changes and either g magnitude or the angle between Γ and g is time variant (Fig.5, left), this dynamic load i.e. g acceleration, makes disturbances on oscillator output as clock drift and clock bias in mission period T.

CLOCK BIAS

\[ \Delta f/f_0 = \Gamma \frac{g_{av}}{g} \cos (\phi + \delta) = 2n \Gamma \frac{g_{av}}{2n} \text{ppb/T} \]  
(8)

Note: Rate of change of g by altitude is not significant; therefore in calculations it has been substituted by \( g_{av} \), on trajectory.

\[ g_0/g = (r/r+h)^2; g_{av} = 1/h \int g_h \, dh \]  
(9)

FREQUENCY DEVIATION

\[ |\Delta f| = f_0 \Gamma \cos \theta = f_0 \cos (\phi + \delta) \cos \theta; \]  
(10)

Where: 0 < \( \phi < \pi \), \( r \) = earth radius, \( h \) = altitude above earth’s surface, \( g_h \) = acceleration gravity on altitude h.

Fig. 5. (left) crystal subjected to attitude change (right) frequency deviation for \( h = 0-1000 \) (km), \( n = 2, f_0 = 10.23 \) (MHz), \( \Gamma = 10^{-9} / g \), \( \beta = 0 \), 0 < \( \phi < 360^\circ \).

GYRO EFFECT

Due to using gyro-mounting the angle between g and Γ remains constant on host vehicle trajectory. If the change of g is neglected then it can be concluded that all attitude change induced disturbances, will be suppressed by gyro (Fig.5, right).

B. Steady state acceleration

Fig. 6. longitudinal Steady State Acceleration for launch vehicle Arian5 (max lateral load=0.2 g) [4].

For launch vehicle araian5 [4], Longitudinal steady state acceleration changes between 0 < \( a_{long} < 4.2 \) g on its trajectory as shown in Fig.6, and lateral acceleration changes between 0 < \( a_{lat} < 0.2 \) g. Impact of this load on stability of crystal oscillator output directly depends on how to install it on host vehicle (Fig.7). To calculate maximum gyro efficiency, analyses have been done by assumption of vertically installation.

**TABLE I. ANALYSIS OF GYRO EFFICIENCY FOR CRYSTALS WITH \( \phi = 2.6^\circ \).**

<table>
<thead>
<tr>
<th>( \phi )</th>
<th>Fixed oscillator</th>
<th>Using gyro</th>
<th>Gyro effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-30°</td>
<td>0.0429</td>
<td>0.0019</td>
<td>-0.0410</td>
</tr>
<tr>
<td>0-90°</td>
<td>3.57×10^{-5}</td>
<td>9.28×10^{-6}</td>
<td>+5.7×10^{-9}</td>
</tr>
<tr>
<td>( \beta )</td>
<td>Gyro has great effect but still frequency deviation &gt; Allan deviation (10^{-4} Hz).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>It only occurs for ( \beta ) angles too close to 90°.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>Great gyro effect for ( \beta &lt; 70° ) but its result isn’t perfect because phase noise is still &gt; Allan deviation (-70.05).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>This drawback only occurs for ( \beta ) values too close to 90°.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II. ANALYSIS OF GYRO EFFICIENCY FOR CRYSTALS WITH \( \phi = 30^\circ \).**

<table>
<thead>
<tr>
<th>( \phi )</th>
<th>Fixed oscillator</th>
<th>Using gyro</th>
<th>Gyro effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-30°</td>
<td>0.0426</td>
<td>0.0215</td>
<td>-0.0157</td>
</tr>
<tr>
<td>0-90°</td>
<td>3.69×10^{-5}</td>
<td>0.0100</td>
<td>9.69×10^{-14}</td>
</tr>
<tr>
<td>( \beta )</td>
<td>Gyro has good effect but still frequency deviation &gt; Allan deviation (10^{-4} Hz).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>This draw back only occurs for ( \beta ) angles too close to 90°.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>Best effects occur for ( \beta &lt; 45° ) but its result isn’t perfect because phase noise is still &gt; Allan deviation (-70.05).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>This big drawback only occurs for ( \beta ) close to 90°.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp;</td>
<td>Drawsbacks with small values strats from ( \beta &lt; 60° ).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 7. Steady state load applied on crystal when installed on host vehicle (left) horizontally; (right) vertically.

I. Impact on crystal oscillator short-term stability

This load appears as frequency deviation and phase noise on oscillator output and impacts its short term stability (\( \tau = 1s \)).

According to Fig.8, gyro specially shows its great effects for the most probable g-sensitivity angles i.e. 2.6°<\( \phi < 30^\circ \). Therefore gyro efficiency will be between values have been shown in Fig.9 and summarized in table I and II.

It is considerable that besides reduction of disturbances, gyro shifts the max values from the most probable g-sensitivity angles i.e. 2.6°<\( \phi < 30^\circ \) to low probable (\( \phi > 30^\circ \)).

C. Sinusoidal Vibrations (2-100 Hz)

Sinusoidal vibrations applied on payload installed inside fairing of ARIAN5 [4] is as shown in Fig.10. The Magnitude and angular orientation of these vibrations are as illustrated in Fig.11. The conic area points out that this load could be applied
on crystal surface with any angle \(0<\beta<360^\circ\), and it directly depends on how to install oscillator on host vehicle.

\[ \text{Fig. 10. Arian 5 sinusoidal vibrations, (blue) longitudinal, (red) lateral.} \]

\[ \text{Fig. 11. Magnitude and angular orientation of sinusoidal vibrations as } A=Ax_0\sin(2\pi f t),0<\beta<2\pi, (\text{left}) \varphi=38.7^\circ; (\text{middle}) \varphi=31^\circ; (\text{right}) \varphi=36.87^\circ. \]

1) Impact on crystal oscillator short-term stability
Sinusoidal vibrations appear as frequency jitter and phase noise on oscillator output and impact its short term stability.

According to Fig. 12, gyro-mounting shows its great effects especially for the most probable angles \(2.6^\circ<\varphi<30^\circ\). Therefore gyro efficiency will be between values have been shown in Fig. 12 and summarized in table III and IV.

| TABLE III. | ANALYSIS OF GYRO EFFICIENCY FOR CRYSTALS WITH \( \varphi=2.6^\circ \) |
| \( \Phi=2.6^\circ \) | \( 0<\beta<90 \) | Fixed oscillator | Using gyro | Allan deviation | Gyro effect |
| \( \Delta f_{\text{max}}(Hz) \) | 0.0086 | 5.94×10^{-4} | 0.0001 | -0.008 |
| \( \Delta f_{\text{min}}(Hz) \) | 1.14×10^{-4} | 4.64×10^{-5} | 1×10^{-4} | 3.5×10^{-5} |
| \( \xi(f) \) | -53.31 | -76.56 | -92.04 | 23.25 |
| \( \xi(f)_{\text{min}}(dBc/Hz) \) | -124.92 | -112.7 | -126.02 | 12.22 |

2) Impact on crystal oscillator short-term stability
Sinusoidal vibrations appear as frequency jitter and phase noise on oscillator output and impact its short term stability.

According to Fig. 12, gyro-mounting shows its great effects especially for the most probable angles \(2.6^\circ<\varphi<30^\circ\). Therefore gyro efficiency will be between values have been shown in Fig. 12 and summarized in table III and IV.

| TABLE IV. | ANALYSIS OF GYRO EFFICIENCY FOR CRYSTALS WITH \( \varphi=30^\circ \) |
| \( \Phi=30^\circ \) | \( 0<\beta<90 \) | Fixed oscillator | Using gyro | Allan deviation | Gyro effect |
| \( \Delta f_{\text{max}}(Hz) \) | 0.0122 | 0.0065 | 0.0001 | -0.0057 |
| \( \Delta f_{\text{min}}(Hz) \) | 1.64×10^{-3} | 0.0051 | 1×10^{-4} | 3.5×10^{-5} |
| \( \xi(f) \) | -50.31 | -75.56 | -92.04 | 23.25 |
| \( \xi(f)_{\text{min}}(dBc/Hz) \) | -121.71 | -91.84 | -126.02 | 29.87 |

- This big value drawback occurs only for \( \beta \) close to 90°.
- Drawbacks with small values start from \( \beta>60^\circ \).
D. Random Vibration \((20<\nu<2000\text{ Hz})\)

Random Vibration is a bandlimited noise with Gaussian distribution which is created by different sources. Generally random vibration defined by PSD \((g^2/\text{Hz})\) or \((\text{dBc/Hz})\) with respect to its source. Regardless to these sources, random vibration causes short term instability as frequency jitter and phase noise on oscillator output.

As an example to show gyro effect, mechanical random vibration is analyzed according to PSD \((g^2/\text{Hz})\) of Arian4 host vehicle; \((1^{\text{st}}\text{ and } 3^{\text{rd}}\text{ pics.})\) using gyro; for critical state \(\beta=0\) and for different values of \(0<\varphi<62^\circ\), \(\Gamma=10^9\), \(f_0=10.23\) (MHz), \(2<f,<100\) (Hz), \(\sigma_2=10^{-11}\).

Since Random Vibration is a combination of all the frequencies at the same time, at first it is necessary to configure this load in time domain. According to the parseval’s law, \(g_{\text{rms}}\) is equal to \(\sigma\) of Random Vibration. Therefor Random Vibration is defined in time domain by calculation of \(g_{\text{rms}}\) from PSD (Fig.14, right).

**1) Random Vibration in time domain**

- **Calculation of \(g_{\text{rms}}\)**
  
  \[ g_{\text{rms}} = \sqrt{\text{PSD} (g^2/\text{Hz}) \times f \text{ (Hz)}} \]  \hspace{1cm} (11)

  - From point A to B: \(20<\nu<150\); \(\text{PSD/PSD}_1 = (f/f_1)^{3\text{/2}}\rightarrow \text{PSD}=7.7778\times10^{-6} f^2 \text{ (g}^2/\text{Hz)}\) \hspace{1cm} (12)
  - From point B to C: \(150<\nu<700\); \(\text{PSD} = 0.04 \text{ (g}^2/\text{Hz)}\) \hspace{1cm} (13)
  - From point C to D: \(700<\nu<2000\); \(\text{PSD/PSD}_2 = (2/f)^4\rightarrow \text{PSD}=28/f \text{ (g}^2/\text{Hz)}\) \hspace{1cm} (14)

  \[ g_{\text{rms}} = 7.3 \text{ g} \rightarrow 1\sigma = 7.3 \text{ g}; 2\sigma = 14.6 \text{ g}; 3\sigma = 21.9 \text{ g}. \]  \hspace{1cm} (15)

**b) Impact on crystal oscillator short-term stability**

To Analysis gyro effect on random vibration it is necessary to determine 3 unknown parameters, angles \(\xi, \beta\) and \(\varphi\) (Fig.15).
If \( \xi = \varphi \) and \( \beta = 0 \), thus \( \Gamma \) and \( \Lambda \) coincide each other, therefore gyro shows its best positive effects especially for \( \varphi < 30^\circ \), (Fig.17 and Fig.16 left). If \( \xi < \varphi = \pi/2 \), then \( \Gamma \) and \( \Lambda \) are perpendicular to each other, in this case gyro shows the worst drawback, (Fig.16 right). As summarized in table V, analysis of these 2 cases says that totally when crystal oscillator installed on gyro, the probability of observing random vibration-induced instability on oscillator output is reduced.

### TABLE V. MAX GYRO EFFECTS AND DRAWBACKS

<table>
<thead>
<tr>
<th>Random vibration induced instability</th>
<th>( 2^\circ&lt;\varphi&lt;30^\circ )</th>
<th>Fixed oscillator</th>
<th>Using gyro</th>
<th>Allan deviation</th>
<th>Gyro effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max gyro effect on ( \Delta f(Hz) )</td>
<td>0.2243</td>
<td>0.0078</td>
<td>0.0001</td>
<td>-0.2165</td>
<td></td>
</tr>
<tr>
<td>Max gyro drawback on ( \Delta f(Hz) )</td>
<td>0.0039</td>
<td>0.1121</td>
<td>0.0001</td>
<td>0.1082</td>
<td></td>
</tr>
<tr>
<td>Max gyro effect on ( \gamma f(dBc/Hz) )</td>
<td>-85.03</td>
<td>-14.17</td>
<td>-152.04</td>
<td>-29.14</td>
<td></td>
</tr>
<tr>
<td>Max gyro drawback on ( \gamma f(dBc/Hz) )</td>
<td>-120.19</td>
<td>-91.31</td>
<td>-152.04</td>
<td>28.88</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE VI. MAX GYRO EFFECT ON CRYSTALS SHORT-TERM INSTABILITY.

<table>
<thead>
<tr>
<th>Attitude change (Fig.5)</th>
<th>Gyro supresses totally this load induced disturbances,clock bias and ( \Delta f ).</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Phi=2^\circ, \beta=0 )</td>
<td>Steady State Acceleration (Fig.8)</td>
</tr>
<tr>
<td>Fixed oscillator</td>
<td>Using gyro</td>
</tr>
<tr>
<td>( \Delta f_{max}(Hz) )</td>
<td>0.0429</td>
</tr>
<tr>
<td>( \gamma f_{max}(dBc/Hz) )</td>
<td>-17.34</td>
</tr>
<tr>
<td>Gyro reduces disturbances greatly.</td>
<td></td>
</tr>
</tbody>
</table>

### Sinusoidal Vibrations (Fig.12)

<table>
<thead>
<tr>
<th>Random Vibration (Fig.17)</th>
<th>Gyro reduces disturbances close to safety margin of Allan deviation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta f_{max}(Hz) )</td>
<td>0.0064</td>
</tr>
<tr>
<td>( \gamma f_{max}(dBc/Hz) )</td>
<td>-89.87</td>
</tr>
<tr>
<td>Gyro reduces disturbances near to safety margin of Allan deviation.</td>
<td></td>
</tr>
</tbody>
</table>

### Sinusoidal Vibrations (Fig.13)

<table>
<thead>
<tr>
<th>Random Vibration (Fig.17)</th>
<th>Gyro reduces disturbances close to safety margin of Allan deviation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta f_{max}(Hz) )</td>
<td>0.2243</td>
</tr>
<tr>
<td>( \gamma f_{max}(dBc/Hz) )</td>
<td>-85.03</td>
</tr>
<tr>
<td>Gyro reduces disturbances near to safety margin of Allan deviation.</td>
<td></td>
</tr>
</tbody>
</table>

### REFERENCES

2. Maryam Abedi, Tian Jin, “Improvement in tracking loop threshold of high dynamic GNSS receiver by installation of crystal oscillator on gyroscopic mounting”,in press.
Abstract—The New Horizons mission has provided an unprecedented opportunity to observe the evolution of two ultra-stable oscillators in deep space for over nine years. The unusual architecture of the New Horizons communications system allows for direct observation of one of the two USOs during every tracking pass. The communications system is based on a transceiver as opposed to the normal transponder. The downlink is independent of the uplink in this architecture and so the USO is solely responsible for the downlink carrier frequency. Both USOs have greatly stabilized during the nine years of flight and have recently changed the direction of the frequency drift. The history since launch in January 2006 and the recent drift history are presented and discussed.

Keywords—ultra-stable oscillators, deep space tracking

I. INTRODUCTION

The New Horizons (NH) mission to the Pluto and subsequent exploration of the Kuiper Belt structure in our outer solar system is expected to achieve a major mission goal by its impending flyby of the Pluto-Charon double planet in July 2015. One of the primary science objectives for NH during the Pluto-Charon flyby is the chemical characterization of any possible atmosphere at Pluto, including assessment of atmospheric density and dynamic interaction with Pluto’s surface. The measurement of Pluto’s possible atmosphere will be accomplished by an NH on-board radio science experiment, known as REX. To achieve its science objectives, REX requires the frequency stability of an ultra-stable oscillator (USO) with Allan deviation of no more than 1x10^{-13} over time intervals from 1 to 1000 s.

Planetary radio science is commonly performed using the downlink. REX is non-traditional in that signal interrogation will be performed on-board NH using the uplink provided by the NASA Deep Space Network ground stations. This is because, at over 30 AU from Earth, the NH downlink S/N would be too low to achieve the desired science return [1]. The two Johns Hopkins University Applied Physics Laboratory (JHU/APL) USOs A & B onboard NH are allowed to continuously operate during the spacecraft’s flight to Pluto-Charon to assure the maximum state of readiness for REX, while improving their frequency stability and reducing drift (aging). Fig. 1 shows the placement of USOs A & B on the spacecraft. As well as supporting REX, the USOs provide the signal reference for the spacecraft’s transceiver and non-coherent navigation system.

Fig. 1. The New Horizons spacecraft carries two JHU/APL ultra-stable oscillators (USO) to support the REX, radio science experiment of Pluto’s possible atmosphere. The USO are located in the figure as well as other major telecommunications sub-systems of the spacecraft.

The NH non-coherent navigation system provides the opportunity for precise determination of USO’s A & B in-flight frequency performance during the entirety of every downlink pass [2]. This represents an unprecedented amount of data for USOs in deep space. In this paper, we will report on the frequency stability of each USO over the nine years since launch. We will discuss the treatment of general and special relativity, including the Shapiro delay. The frequency performance of each NH USO demonstrates distinct interactions with various space environments, such as with Jupiter’s radiation belt in Feb. 2007 and the impact of the normal 5 RPM spin rate on the two USOs. The frequency drifts of USOs A & B are currently -1.3x10^{-13}/day and +1x10^{-13}/day, respectively. It is possible both USOs may be on the cusp of a
re-curve in their drift behavior. We will discuss our confidence in this possibility, and any concern to the performance of REX.

II. NEW HORIZONS SUPPORT OF TRACKING OPERATIONS

A. Doppler

Without additional hardware and software, the New Horizons transceiver could not support accurate Doppler velocity measurements. The reference oscillator is not stable enough to make one-way Doppler measurements accurate enough for navigation. The New Horizons communications system includes a small amount of hardware that creates telemetry that can be used to accurately monitor the ratio of the uplink and downlink frequencies at the spacecraft. Software in the Mission Operations Center uses this telemetry to convert the raw downlink carrier phase measurements into those that would have been produced by a transponder. These converted files are then used by the navigation team without further consideration of the communications system architecture.

B. Range

Ranging is supported by programming the uplink frequency so that the ratio of the downlink and uplink frequencies at the spacecraft is very close to the nominal X-band turnaround ratio of 880/749. The uplink frequency at the spacecraft is typically within 1 Hz of the value that would exactly produce this ratio. An error at this level is not sufficient for Doppler measurements, but is sufficient to support accurate ranging.

III. REFERENCE OSCILLATOR DETERMINATION

Two methods are used to compute the reference oscillator frequency from the observed downlink carrier phase. These are referred to as a one-way method and a two-way method depending on whether or not the uplink station frequency is involved in the calculation.

A. One-way Method

One method combines this observed carrier phase with the ephemeris of the spacecraft ephemeris and the knowledge of the motion of the ground station during the measurement. Additionally, the rotation state of the spacecraft (normally either 3-axis stabilized or rotating at 5 RPM) and the downlink polarization (LCP or RCP). This method can be applied to every downlink pass.

B. Two-way Method

A second method is applied when a tracking pass involves both an uplink and a downlink. For the majority of the last two years, most of this tracking has been “3-way” tracking in which the uplink ground station is different from the downlink station due to the long round-trip light times. The current round-trip time is approximately 9 hours. This method has the advantage that it places less reliance on the ephemeris. Because of the nature of the non-coherent navigation method, the range rate can be solved for given the telemetry used to convert the observed downlink carrier phase to its transponder equivalent. The ephemeris is still used, but errors have less of an effect than for the one-way calculation described above.

For this two-way method, the observed carrier phase and the telemetry are combined with the information used for the one-way calculation plus the history of the uplink frequency, the motion of the uplink station, and the round-trip light time.

Relativistic effects are accounted for to express the oscillator frequency in the frame of reference of the spacecraft and the ground potential of the ground station at launch. Other relativistic effects such as the Shapiro delay are accounted for.

IV. OBSERVATIONS

Fig. 2 shows the history of the primary USO, since launch on January 20, 2006. The figure shows the deviation of the USO frequency from its nominal value of 30 MHz. The deviation in Figs. 2 through 5 are expressed in terms of parts-per billion relative to this nominal frequency. The gaps in Figs. 2 through 5 are those periods during which the spacecraft was in hibernation and no tracking was performed. Tracking was normally performed briefly around the beginning of each year and then more extensively during the annual check-out each summer.

The nine years of light have afforded the USO time to not only settle to a drift rate that is two orders of magnitude smaller than the drift rate that existed immediately after launch, but it has allowed the emergence of another process by which the drift rate has reversed direction. The history of USO A since Jan. 1, 2014 is shown in Fig. 3. The apparent change in the direction of the frequency drift is seen in the data that followed the beginning of 2015. When the geometry is such that the line-of-sight between Earth and the spacecraft is near the Sun, the noise in the measurements is exaggerated, as seen by the indicated portions of the frequency history curve.

![Fig. 2. USO A drift history since launch in January 2006.](image)
The history of the drift of USO B is shown in Figs. 4 and 5. The history of USO B is roughly the inverse of the drift observed for USO A. USO B began the mission with a positive offset and drifted to a higher frequency, while USO A started negative and drifted to a lower frequency. As with USO A, the drift rate of USO B turned around and is drifting to a lower frequency during the most recent observations.

An additional interesting feature of the frequency history shown in Fig. 5 is the dependence of the frequency of USO B on the spin state of the spacecraft. Two intervals are highlighted in Fig. 5. In each case, the oscillator frequency dropped by around 3.5 x 10^{-11} when the spacecraft changes from the normal 5 RPM spin rate to being 3-axis stabilized. This behavior has been observed since launch.

In [3], we described the position of USO B as 21" from the central spin axis. The position of USO A is 5" from the central spin axis. The stabilizing rotation of the spacecraft at 5 RPM imparts ~0.015 g’s to USO B resulting in a frequency shift in the order of 2 to 3x10^{-11} for an assumed resonator acceleration sensitivity of ~2x10^{-9} per g. The acceleration at USO A is about 77% less. When higher spin rates were employed early in the mission, the frequency offset in the USO’s was observed to be proportional to the square of the spin rate, which should be expected for a stress on the resonator imposed by the centripetal acceleration.

V. ANALYSIS OF AGING CHARACTER

Beyond the curiosity we hold toward the observed aging performance of the two USOs on-board New Horizons is also the practical question of what frequency drift rate should be expected at the mission critical phase of collecting radio-science data from REX. For the majority of the cruise both USOs have displayed a gradually decreasing, yet assumed monotonic aging character. The apparent coincidence of both USOs undergoing reversal of frequency drift within several months of the Pluto-Charon encounter gives cause to examine whether either of these newly evolving processes could compromise the required frequency stability of REX.

In [4], the mechanism for reversal in quartz resonator aging is generally attributed to competing processes that possess
differing time activation profiles. In the majority case, most aging in quartz resonators is associated with the adsorption or desorption of unwanted materials, transferring mass from the surface of the quartz. For a 5 MHz, SC cut resonator of the type in the New Horizons USOs, a change in equivalent mass of one atomic layer of quartz would contribute about $1 \times 10^{-6}$ change in the resonator’s frequency. The observed total change in both USOs over nine years has been approximately $25 \times 10^{-9}$, fundamentally yielding the sense of scale which we are considering in the current state of either USOs aging character. In other words, at the extremely low drift rates in the order of $1 \times 10^{-12}$ per day, the equivalent change of mass on the USO resonators is at the decades of quartz atomic equivalents.

This leads us to believe that, if a reversal in aging is occurring, the processes must be more subtle than mass transfer alone, leaving mechanisms such as diffusion or migration of electrode metallization at the quartz interface layer, and stress changes in the electrode thin film. Mechanical creep in the glass housing of the resonator and the brazing joints of the quartz mounting structure may also become apparent at the observed aging rates. Likely less involved is oxidation from housing infiltration, since the units have been in the deep vacuum of space since launch. Also, we are less likely to consider stressing from direct mounting structure changes, other than during spin stabilization periods, since the units are experiencing essentially no acceleration from gravity.

Nonetheless, the extensive frequency data shown in Figs. 3 and 5 for USOs A and B, respectively, indicates the appearance of a reversal in the aging trend. Fig. 6 shows the most recent segment of data for USO A since its emergence from solar occultation in Dec. of 2014. The vertical frequency scale is referred from the data base unit of $1 \times 10^{-9}$ and mean normalized, so that the full magnitude of the frequency change is about $\pm 7.5 \times 10^{-12}$. Over the 110-day period, a slight rise in frequency in the order of $5 \times 10^{-12}$ seems to appear at day 40 (Jan 16th, 2015).

Being aware that we may also be observing noise mechanisms over ever greater time intervals, we recognize the prudence in analyzing for these behaviors. Most significantly, both USO’s flicker floors are on the order of $1 \times 10^{-13}$ starting at a few seconds and expected to present this noise level for extended periods of time, perhaps on the order of several hours. Fig. 7 is the Hadamard deviation of the data shown in Fig. 6. A green line segment is placed at the greatest time intervals representing dependence in tau of +1, sometimes described as flicker walk of frequency. Again the vertical scale of Fig. 7 is referred to the data base unit of $1 \times 10^{-9}$. Therefore encouragingly, we can estimate the apparent flicker of frequency noise of USO A in the received downlink as presenting in the low $10^{-13}$ out to several hours, as should be expected.

Fig. 7. Hadamard deviation of USO A since emergence from the latest solar occultation. The green line segment is proportional to a +1 tau dependence described as flicker walk of frequency. According to [5], a change of drift should present itself as a +2 dependency in the Hadamard.

Since it is know that the Hadamard is not sensitive to frequency drift, the emergence of a higher order noise mechanism at time intervals greater than 10,000 s must be attributed to uncertain dynamics in this region. Or, as derived in [5], the higher order uplift in the Hadamard may be the beginning of a very slow change in the drift behavior of USO A, which should be indicated from a tau dependence of +2.

Another way to analyze the change behavior of frequency data is to use the dynamic variance approach. The dynamic approach, as derived in [6], divides the frequency data into windows of time that cover the entirety of the data set. Then, the variance of choice is performed for each window over an extension of the base tau to as large as can be resolved for the variance of choice. Fig 8 is the plot of the dynamic Hadamard variance of the data shown in Fig. 6 for USO A. The result in Fig. 8 is a complicated surface with deep variations spanning more than a magnitude across the individual time windows. The basic interpretation of Fig. 8 is that the surface is mostly flat over the axis labeled “Log AF”, or time intervals from 370 to 62000 s, meaning that the noise process is representative of flicker FM. The greater variation, which appears as deep
notches relative to the “Windows” axis, indicates a rippling in the magnitude of flicker intensity across the data. From this analysis of Fig. 8, the description of the frequency character of USO A is more consistent with the emergence (or ability to observe) a higher order noise process, such as flicker walk of frequency, than a change of drift which would be expected to have a smooth up-lifting nature to the surface as time intervals increase along the Log AF axis in Fig. 8.

Fig. 8. Dynamic Hadamard variance for USO A since emergence from the latest solar occultation. The axis labeled “Window #” consists of slices of time of slightly larger than a day. The axis labeled “Log AF” should be interpreted as the typical tau axis of variance plots scaled to the base of the time interval range, in this case 3700 s. The greater noise variation, which appears as deep notches relative to the “Windows” axis, indicates a rippling in the magnitude of flicker intensity across the data.

Our final approach toward determining whether the apparent reversal in the aging trend of USO A is a change in drift direction examines just the rising tail of the data of Fig. 6, starting at about day 36 (about Jan. 12th 2015 of Fig. 3). In Fig. 9, The Allan deviation of the USO A data from day 36 to day 110 is shown along with a green line indicating random walk with dependence in tau of +0.5. Surprisingly, this data which gives the highest impression that the drift is changing direction in USO A fits very well to a random walk process. Taken together with the previous analyses, the Allan deviation result of Fig. 9 argues convincingly that the drift mechanism observed in USO A since launch has effectively diminished, such that only higher order noise processes are observed.

VI. CONCLUSION AND DISCUSSION

Both USOs of the NH spacecraft appear to show a reversal of long term aging after nine years of monotonically improving behavior. The apparent change in aging drift is occurring just before the Pluto-Charon encounter. Starting from the last solar occultation of Dec 2014, the frequency of USO A will remain continuously monitored through the Pluto-Charon encounter, expected in July of 2015. Consistent with [4], the frequency data obtained through this period should be sufficient to conclusively establish whether an extremely slow aging mechanism, with a reverse characteristic than that observed over the last nine years, exists in the resonator. More likely, if an aging reversal is occurring in both USO’s, it is newly induced by some speculative process.

Nonetheless, dynamic variance analysis indicates the possibility the USOs are simply revealing higher order noise processes, such as white frequency (random) and flicker frequency walk, not previously observable until the drift mechanisms have practically subsided.

At this time while no strong conclusion for aging reversal in the New Horizons USOs can be formed, there is no reason to expect any remaining drift mechanism in the resonators could emerge strong enough to prevent REX from achieving its goals.

ACKNOWLEDGMENT

The authors thank the New Horizons Project for the support of this work.

REFERENCES

A CMOS LC-Based Frequency Reference with ±40ppm Stability from -40°C to 105°C

Timing Products Division, Si-Ware Systems, Cairo, Egypt
E-mail: nabil.sinoussi@si-ware.com

Abstract—This work presents a highly stable monolithic integrated CMOS LC-based frequency reference. The frequency reference is based on a Self-Compensated Oscillator (SCO) architecture where the LC tank operates at a specific phase \( \phi_{\text{NULL}} \) where frequency sensitivity versus temperature is minimum. A new compensation technique is applied over \( \phi_{\text{NULL}} \) to further optimize frequency stability and extend the temperature range. The new technique is based on an analog approach and induces a minimum impact on oscillator phase noise, current consumption and die area. Utilizing this technique, the temperature range has been extended to (-40-105°C) with a ±40ppm frequency stability. Achieved performance makes it possible for the SCO to be introduced to automotive applications where crystals suffer vibration induced stability issues.

Keywords—Self-Compensated Oscillator; Compensation Block; Automotive.

I. INTRODUCTION

Evolving applications like wearable devices and internet of things are continuously stimulating electronic systems towards further integration and miniaturization. As the heart of any electronic system, reference clock generators have to cope with these stimulants. However, during the past decade, the volume of quartz crystals, the foundation of most of the clock generators in any electronic application, seemed saturating [1]. While the developing MEMS technology has proven its capability to fit inside quite challenging form factors with adequate performance, the corresponding well-established quartz technology is not as miniature [2]. Nevertheless, a pure CMOS clock generator would be superior from the aspect of miniaturization in the sense that it can be integrated inside standard CMOS System-on-Chip’s (SoCs). Few promising pure CMOS LC-based clock generators have been demonstrated and even commercialized as stand-alone solutions with a great potential for full integration in the future [3], [4] and [5].

Moreover, Crystal Oscillators (XOs) are sensitive to vibration [6]. This imposes difficulties in the assembly of electronic systems in some applications such as automotive applications. On the other hand, CMOS oscillators are expected to have less sensitivity to vibration [7]. However, CMOS oscillators cannot yet fill this gap due to the stringent stability and reliability requirements imposed by automotive applications. One of the most challenging barriers is the operating temperature range of the top grades e.g. grade 2 automotive electronics is required to operate in the range (-40-105°C) [8]. Such a wide temperature range represents a challenge even for XOs to be achieved in a cost efficient method [9]. Actually, the best performances reported so far for pure CMOS oscillators do not exceed the industrial temperature range (-40-85°C) [3], [4], [5] and [10]. This motivated the work herein as a continuation to the earlier research.

Pure CMOS LC-based SCO has been introduced in [11]. The SCO has demonstrated ±50ppm frequency stability across the temperature range (-20-70°C) [3] and ±100ppm across (-40-85°C) [10]. However, the frequency deviation across temperature (Δ\(f_{TC}\)) of the SCO totally relied upon the LC tank engineering without any further control knobs. This limited the SCO capability to expand its temperature range even further.

Nevertheless, the unique phase-based architecture of the SCO has allowed a novel phase compensation technique, presented in this work, which further expands the SCO temperature range. The intrinsic Δ\(f_{TC}\) of the SCO is orders of magnitude less than that of conventional LC oscillators [3]. Thus, the SCO offers a superior starting point for the compensation technique which simplifies the architecture to a great extent yielding negligible overhead in the current, area and system complexity. The new technique enables the SCO to achieve a distinct Δ\(f_{TC}\) of ±40ppm across the temperature range (-40-105°C).

This paper is organized as follows. Section II briefly highlights the SCO concept. Section III illustrates the novel phase compensation technique. Section IV summarizes the measurement results. Finally, Section V concludes the main points of the paper.

II. LC-BASED SELF-COMPENSATED OSCILLATORS

Fig. 1 depicts that any standard CMOS LC tank has a specific operating phase at which the temperature dependence of the tank impedance decreases dramatically. This phenomenon is denoted as the Temperature Null (TNULL) phenomenon and that specific phase at which the temperature dependence is minimized is denoted as \(\phi_{\text{NULL}}\) [11]. Conventional LC oscillators operate around the zero phase of the LC tank, thus, produce Δ\(f_{TC}\) of thousands of ppms across a given temperature range. On the other hand, the SCO utilizes the TNULL phenomenon by forcing the LC tank to operate around \(\phi_{\text{NULL}}\) yielding Δ\(f_{TC}\) of only tens of ppms [11].

The SCO is capable to operate at \(\phi_{\text{NULL}}\) by using intelligent circuit techniques. In [3], two techniques are illustrated in details. One of which is the Quadrature SCO (IQ-SCO) that has
been chosen to demonstrate the phase compensation technique in this work. For the sake of convenience, the IQ-SCO is briefly explained in this section.

Fig. 2 shows the IQ-SCO which is based on the conventional quadrature LC oscillator architecture. It is known that the two tanks of such an oscillator operate at a non-zero phase ($\phi$) such that: $\tan(\phi) = Gmc/Gmo$, where $Gmc$ is the coupling transconductance and $Gmo$ is the main oscillator transconductance [12]. The IQ-SCO utilizes a programmable infrastructure to control the value of $Gmc$, and in turn the value of $\phi$ precisely. $Gmc$ is adjusted such that the LC tank operates at its TNULL phase i.e. $\phi = -\phi_{\text{NULL}}$ and the oscillator is thus self-compensated. The programmable infrastructure can control the phase of the LC tank with an accuracy of $\pm0.05^\circ$. An Automatic Amplitude Control loop (AAC) determines the oscillation amplitude such that the transconductances operate in a rather linear regime; hence, reduce the higher order harmonic effects. More details about the IQ-SCO architecture is found in [3].

III. PHASE COMPENSATION TECHNIQUE

The inherent performance of an SCO cannot be better than $\pm100$ppm across (-40-85°C). In this section, a novel compensation technique for the SCO is proposed to further improve the frequency stability versus wider extended temperature ranges.

Fig. 3 illustrates a block diagram for the proposed compensation concept. The SCO is initially trimmed to operate at $\phi_{\text{NULL}}$. The new phase compensation concept idea is to induce temperature dependence to the phase of the impedance of the LC tank $\Delta\phi_{\text{NULL}}$ around the trimmed value of $\phi_{\text{NULL}}$. The temperature-dependent variation of $\phi_{\text{NULL}}$ in return, causes a deviation of the SCO frequency across temperature. The induced frequency deviation $\Delta\phi_{\text{NULL}}$ is utilized to control the inherent frequency deviation of the SCO. Initially, the temperature of the SCO is detected by an analog temperature sensor. Afterwards, the Compensation Block generates the Control Signal, ($S(T)$), that is dependent on the sensor’s output. The control signal is then used to vary $\phi_{\text{NULL}}$. The compensation block is programmable where $S(T)$ is generated with different magnitudes and different temperature profiles. The proper magnitude and temperature profile are selected so as to compensate the inherent deviation of the frequency of the SCO at $\phi_{\text{NULL}}$ and produce a better stability across a wider temperature range. Fig. 4 illustrates an inherent parabolic frequency deviation across temperature of an SCO. The Compensation Block produces an $S(T)$ that induces the inverse parabolic behavior. The net compensated $\Delta f_{\text{TC}}$ is fully dependent on the characteristics of $S(T)$ where the residue illustrated in figure illustrates the difference between the inherent SCO behavior and the compensating function.

One of the most important merits of the proposed compensation technique is that $S(T)$ is of an analog nature. The Compensation Block utilizes Band Gap (BG), Proportional to Absolute Temperature (PTAT) and Complementary to Absolute Temperature (CTAT) voltages and currents to produce the different compensating temperature profiles. Thus,
The inherent parabolic behavior of the SCO is compensated by the inverse parabolic compensating function.

Fig. 5. The implementation of the Compensation Block within an IQ based SCO where $S(T)$ controls the coupling transconductance stages.

the compensation technique does not induce any phase noise deterioration or significant current consumption increase that will be associated with the utilization of a digital temperature sensor and a digital compensation technique.

Fig. 5 illustrates the proposed implementation of the compensation technique. The implementation is based on the IQ-SCO. The generated control signal, $S(T)$, is applied to the coupling transconductance cells; $G_{mc}$. The variation in the phase of the LC tank impedance around $\phi_{\text{NULL}}$, is defined by $\Delta\phi_{\text{NULL}} = \tan^{-1}(\Delta G_{mc}/G_{mc})$, where $\Delta G_{mc}$ represents the variation in the coupling stages transconductance induced by $S(T)$.

IV. MEASUREMENT RESULTS

The compensation technique circuitry has been integrated with an IQ based SCO architecture and the chip has been fabricated on a 0.18$\mu$m CMOS process with a single-poly and 6 aluminum metal layers. The SCO parts have been assembled within 5.0X3.2 ceramic packages. All parts have been trimmed to operate at $\phi_{\text{NULL}}$, where operation at $\phi_{\text{NULL}}$ equalizes the frequency of the SCO at the temperature range extremes; in this case $-40$ and $105^\circ$C. Temperature deviation at room temperature has been measured. Based on the measured frequency deviation relative to the frequency at the temperature extremes, the proper magnitude and temperature profiles of $S(T)$ have been selected for each part. The net compensated frequency stability of parts has been characterized by inserting the parts in a temperature chamber and sweeping the temperature from $-40^\circ$C to $110^\circ$C in steps of $10^\circ$C. Fig. 6 illustrates the compensated stability results of 20 parts; 10 parts trimmed to operate at $12.288\text{MHz}$ and 10 parts trimmed to operate at $62.500\text{MHz}$. All parts across a wide temperature range of ($40-105^\circ$C) are illustrating an excellent performance of $\pm40\text{ppm}$. Results illustrate clearly the merits and power of operating the SCO at $\phi_{\text{NULL}}$ and applying a compensating technique for optimizing the frequency stability of the SCO. This performance is unprecedented and has not been reported previously for any LC-based reference oscillator.

V. CONCLUSION

An unprecedented performance of $\pm40\text{ppm}$ frequency stability across a ($40-105^\circ$C) of an LC based reference oscillator has been reported. The LC based oscillator is an SCO operating at $\phi_{\text{NULL}}$, with a novel compensation technique applied to further improve frequency stability across a specific temperature range or extend the temperature range. The technique is based on an analog approach with minimum impact on phase noise or current consumption or even architecture complexity. Results illustrate clearly the potential of the compensated SCO of being utilized as a reference clock.
for automotive applications that require a temperature range of (-40-105°C). Furthermore, the power of the new proposed technique makes it possible to extend the temperature range to (-40-125°C) by applying more trimming insertion temperatures or different compensating temperature profiles across subsets of the temperature range.

REFERENCES


Effects of Pressure and Bias Voltage on the Phase Noise of CMOS-MEMS Oscillators

Wan-Cheng Chiu\textsuperscript{1}, Ming-Huang Li\textsuperscript{1}, Chao-Yu Chen\textsuperscript{1}, and Sheng-Shian Li\textsuperscript{1,2}

\textsuperscript{1}Inst. of NanoEngineering and MicroSystems, \textsuperscript{2}Dept. of Power Mechanical Engineering
National Tsing Hua University, Hsinchu, Taiwan
E-mail: wancheng5990@gmail.com

Abstract—In this work, we present a comprehensive study on the effects of environmental pressure ($P$) and resonator dc-bias voltage for the phase noise of a monolithic CMOS-MEMS oscillator. In order to access the practical utility of CMOS-MEMS oscillators for versatile applications, a double-ended tuning fork (DETF) MEMS resonator oscillator is used as a case study. In the ambient pressure, the oscillation ensues at a minimum $V_p = 30V$ and shows a phase noise (PN) of $-86 \text{ dBc/Hz}$ at 1-kHz offset and $-99 \text{ dBc/Hz}$ at 1-MHz offset. On the other hand, a low-$V_p$ CMOS-MEMS oscillator with IC compatible voltage (i.e., $V_p = 3V$, leading to an equivalent motional impedance $R_m$ of 100 M$\Omega$) is also demonstrated in a vacuum chamber ($P < 1 \text{ mTorr}$) with a PN of $-94 \text{ dBc/Hz}$ at 1-kHz offset and $-98 \text{ dBc/Hz}$ at 1-MHz offset, respectively.

Keywords—MEMS; resonators; sustaining circuits; oscillators; phase noise; CMOS; monolithic integration.

I. INTRODUCTION

During the past decades, microelectromechanical systems (MEMS) oscillators attract substantial interests in emerging applications, such as timing reference and environmental sensing owing to their small size and high level integration with integrated circuits (IC) \cite{1}. In particular, the recent advancement in CMOS-MEMS technologies demonstrates advantages over other solutions, including monolithic MEMS-circuit integration, high fabrication yield, and harnessed process variations, so as to provide a cost-effective solution for single-chip oscillator designs \cite{2-4}.

To understand the ultimate performance that CMOS-MEMS oscillators can achieve, the phase noise (PN) of such monolithic oscillators are broadly studied. The prior works in \cite{4-5} have studied CMOS-MEMS oscillators operated in air ambient with lower $Q$-factors. The low-$Q$ and severe nonlinearity in their MEMS (and NEMS) resonator devices sets the baseline for the closed-to-carrier phase noise. On the other hand, we have demonstrated the ultimate phase noise performance for a CMOS-MEMS \textit{nonlinear} oscillator under given $Q$-factor and circuit topology very recently \cite{6} using a double-ended tuning fork (DETF) resonator.

To further understand the phase noise behavior for CMOS-MEMS oscillators under different resonator dc-bias ($V_p$) and environmental pressure ($P$), the DETF CMOS-MEMS oscillator in \cite{2} is again used to complete this study. In this work, the oscillator performance under various \textit{extreme conditions} is reported. The comparison of the measurement results with theoretical and simulation models are also reported.

II. CMOS-MEMS OSCILLATOR AND PHASE NOISE

A. CMOS-MEMS Resonator

Fig. 1 schematically shows a CMOS-MEMS oscillator based on a DETF MEMS resonator. The MEMS resonator is 136 $\mu$m long and 5 $\mu$m wide with a resonance frequency of 1.2 MHz. Due to the fabrication process limit, the equivalent in-plane capacitive transducer gap of 1 $\mu$m is used for the resonator \cite{2}. Based on the capacitive resonator model \cite{1}, for a given resonator dimension and gap spacing, $R_m$ is inversely proportional to $Q$-factor ($Q$) and dc-bias ($V_p$), i.e.,

$$R_m \propto \frac{d^4}{V_p^2 \times Q}. \quad (1)$$

Note that for a flexural-mode MEMS resonator operated in MHz range, $Q$-factor is highly pressure-dependent \cite{1}. Therefore, any change in environmental pressure ($P$) and dc-bias of the resonator directly impacts $R_m$ of a given resonator.

Fig. 2 shows the expected $R_m$ for the DETF CMOS-MEMS resonator based on the measured data in vacuum ($R_m = 320k\Omega$, $P = 2,100$, $V_p = 60V$). According to (1), $R_m$ for the resonator under different $V_p$ and $Q$ can be estimated.
oscillator topology is chosen for avoiding Q-loading effect. Moreover, TIA-based topology offers great flexibility on PN tuning and optimization [6]. Based on the noise sources illustrated in Fig. 1, the linear time invariant (LTI) PN model can be derived as

\[ S_{PN} = \frac{1}{2V_{OUT}} \left[ 4kT R_m + \frac{7}{\Delta f} R_m^2 \right] \left[1 + \left( \frac{f_o}{2Q_m f_m} \right)^2 \right] \] (2)

where \( R_m \) is the motional impedance of the resonator, \( i_m/\Delta f \) is the TIA input-referred current noise, \( f_o \) is the carrier frequency, \( f_m \) is the offset frequency, and \( V_{OUT} \) is the output voltage.

Obviously, for a high-\( R_m \) resonator in this work, the \((i_m/\Delta f)^2 R_m^2\) noise component is larger than \(4kT R_m\). Therefore, the phase noise spectrum is scaled by \( R_m^2 \) if circuit noise remains constant. Fortunately, the input-referred noise of the TIA circuit designed in this work can be adjusted by means of a tunable feedback resistor [2]. Therefore a reasonable PN can be achieved even under extremely high \( R_m \), such as \( R_m > 20 \text{M} \Omega \). Please note that the nonlinear phase noise effects are not incorporated in the model of (2) for simplicity. If the resonator is biased with a high \( V_P \), for example, \( V_P = 75V \), the closed-to-carrier phase noise will be degraded due to nonlinear amplitude-to-phase noise conversion [6].

B. Oscillator Phase Noise

In this work, a transimpedance amplifier (TIA)-based oscillator topology is chosen for avoiding Q-loading effect. Moreover, TIA-based topology offers great flexibility on PN tuning and optimization [6]. Based on the noise sources illustrated in Fig. 1, the linear time invariant (LTI) PN model can be derived as

\[ S_{PN} = \frac{1}{2V_{OUT}} \left[ 4kT R_m + \frac{7}{\Delta f} R_m^2 \right] \left[1 + \left( \frac{f_o}{2Q_m f_m} \right)^2 \right] \] (2)

where \( R_m \) is the motional impedance of the resonator, \( i_m/\Delta f \) is the TIA input-referred current noise, \( f_o \) is the carrier frequency, \( f_m \) is the offset frequency, and \( V_{OUT} \) is the output voltage.

Obviously, for a high-\( R_m \) resonator in this work, the \((i_m/\Delta f)^2 R_m^2\) noise component is larger than \(4kT R_m\). Therefore, the phase noise spectrum is scaled by \( R_m^2 \) if circuit noise remains constant. Fortunately, the input-referred noise of the TIA circuit designed in this work can be adjusted by means of a tunable feedback resistor [2]. Therefore a reasonable PN can be achieved even under extremely high \( R_m \), such as \( R_m > 20 \text{M} \Omega \). Please note that the nonlinear phase noise effects are not incorporated in the model of (2) for simplicity. If the resonator is biased with a high \( V_P \), for example, \( V_P = 75V \), the closed-to-carrier phase noise will be degraded due to nonlinear amplitude-to-phase noise conversion [6].

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

Fig. 3 shows an optical photo of the monolithic CMOS-MEMS oscillator. The CMOS-MEMS circuit is designed in an open-loop configuration for resonator characterization. The closed-loop oscillator can be formed by a single bond wire.

A. Measurement Results

Fig. 4 shows the open-loop measurement result of a DETF CMOS-MEMS resonator. Its Q-factor is around 1,700 in vacuum and 150 in air, respectively. The squeezer film damping is expected to be the main loss mechanism that limits Q-factor of the DETF resonator in the ambient pressure. An off-resonance rejection of 20 dB for both vacuum and air conditions ensures single-mode oscillator operation.

Fig. 5(a) shows the measurement result of a monolithic CMOS-MEMS oscillator operated in air. The minimum \( V_P \) of 30V is used to perform successful oscillation in air, which corresponds to \( R_m \) of 20 M\( \Omega \). The phase noise of -86 dBc/Hz at 1-kHz offset and -99 dBc/Hz at 1-MHz offset are demonstrated. As \( V_P \) increases from 30V to 90V, \( R_m \) decreases from 20 M\( \Omega \) to 2 M\( \Omega \), by which a 15-dB reduction in far-from-carrier phase noise is observed. The phase noise improvement not only comes from \( R_m \) reduction but also accounts for the 2X carrier power increase. Such a phase noise for \( V_P > 65V \) is better than...
that of the prior arts as indicated in Fig. 5(a) [4][7]; however, it necessitates high bias voltage. This will become a possible implementation in future once HV-CMOS or BCD technology platforms are applicable.

Fig. 5(b) further captures the oscillator PN in vacuum, where the minimum bias used is only 3V. Under such low $V_P$, the equivalent $R_m > 100\, \text{M}\Omega$ is expected (cf. Fig. 2). However, with proper TIA biasing, an ultra-low input-referred current noise density around 100 fA/\sqrt{Hz} at 1-MHz can be obtained from simulation. As a result, such low input-referred noise compensates the resonator loss, thus resulting in a best-case phase noise floor of -98 dBc/Hz. Through multiple samples, the average phase noise floor for 3V CMOS-MEMS oscillators is around -87 dBc/Hz. Fig. 5(b) further shows an oscillator biased at $V_P = 33$V and a phase noise floor of -115 dBc/Hz is demonstrated. Compared with Fig. 5(a), if the oscillator is operated in vacuum, the required $V_P$ can be reduced by 3X for similar far-from-carrier PN performance (i.e., 90V to 33V).

B. Discussions

Fig. 6 plots the phase noise of the CMOS-MEMS oscillator under various biasing conditions in vacuum. As expected, the low-$V_P$ oscillator shows the worst phase noise due to its large motional impedance. On the other hand, as the dc-bias voltage increases from 6V to 25V, both the closed-to- and far-from-carrier phase noises are reduced. In this case, the phase noise improvement attributes to the motional impedance reduction. Please note that the slope of the closed-to-carrier phase noise for $V_P = 25$V is steeper than the case of $V_P = 6$ V. This is caused by the increase of amplitude-to-phase noise conversion due to bias-dependent nonlinearities [6]. As a result, the worst closed-to-carrier phase noise for high-$V_P$ case, $V_P = 70$V, can be explained by the MEMS nonlinearity. Based on the measurement results, the best-case phase noise can be attained by medium-$V_P$ (at the range of 25V to 45V in this case study) for a best trade-off between low-$R_m$ and nonlinearity.

Fig. 7 presents the comparison between circuit simulation, theoretical prediction, and measurement result for a DETF CMOS-MEMS oscillator in air with $V_P = 90$V. The extracted linear RLC resonator model is used for simulation; therefore the resonator nonlinearity will not be considered in simulations. With a foundry CMOS model, the Cadence Spectre [8] is chosen for PN simulation using PSS and PNOISE analysis. The PN calculation using (2) is based on the simulated circuit noise ($i_n^2/\Delta f$) and measured oscillator output powers ($V_{OUT} = 960\, \text{mV}_{pp}$). As shown in Fig. 7, the flicker noise up-conversion might be over-estimated in MEMS oscillators in Cadence simulations, since the measured PN for a low-$Q$ oscillator shows a $1/f^2$ trend. As a result, the LTI model can be used to predict the oscillator phase noise in air ambient with certain accuracy, since the nonlinearity for a low-$Q$ resonator is insignificant.

IV. CONCLUSIONS

This work presents a study of the phase noise dependency on the dc-bias $V_P$ and environmental pressure $P$. Medium $V_P$ of 25V to 45V is recognized to be the best biasing condition in vacuum due to balance between $R_m$ and resonator nonlinearity. However, in air, high-$V_P$ (90V) is suggested since the resonator nonlinearity becomes insignificant.

ACKNOWLEDGMENT

The CMOS chip fabrication was supported by the CIC and TSMC, Hsinchu, Taiwan. The authors are grateful to the Cent. for Nanotech., Materials Sci. and Microsys. of National Tsing Hua University for the use of fabrication facilities.

REFERENCES

1/f noise of quartz resonators: Measurements, modelization and comparison studies

FEMTO-ST Institute, UFC, CNRS, ENSMM, UTBM
Besançon, France
fsthal@ens2m.fr

Cibiel Gilles
Microwave and Time-Frequency
CNES
Toulouse, France

Abstract— In this paper, the description of the resonator realization and the topology of the resonator prototype is exposed. Phase noise measurements of a hundred of resonators are given. The noise results are discussed according to the position of the resonators inside the crystal block and physical analysis of the crystal (dislocation). The results are also compared according to their Q-factors measured at room temperature and at low temperature. Theoretically, the fluctuation-dissipation theorem is used in order to put numerical constraints on a model of 1/f noise caused by an internal (or structural) dissipation proportional to the amplitude and not to the speed. The order of magnitude of the noise is then discussed.

Keywords—1/f noise; quartz; resonator

I. INTRODUCTION

The Centre National d’Etudes Spatiales (CNES), Toulouse, France and the FEMTO-ST Institute, Besancon, France, investigate the origins of noise in bulk acoustic wave resonators [1]. Several European manufacturers of high quality resonators and oscillators are involved in this partnership to achieve resonators. For this investigation, quartz crystal resonators have been cut from a quartz crystal block supplied specifically for this study on 1/f noise. This crystal block was grown from a seed which originated from a previous synthetic crystal which was grown from a natural seed. This kind of synthetic crystal is usually used to grow new generations of quartz crystal blocks.

In this paper, the reader is reminded the description of the resonator realization and the topology of the resonator prototype is exposed. The resulting resonators are SC-cut with a 5 MHz resonant frequency. A comparison of these resonators is given in terms of motional parameters. Then, we report the noise measurements made on these quartz crystal resonators using an advanced phase noise measurement system. Phase noise measurements on several batches of resonators are given. The noise results are discussed according to the position of the resonators inside the crystal block, the results of physical analysis of the crystal (dislocation) and the Q-factors measured at room temperature and at low temperature. Measurements of resonator parameters have been done at low temperature in order to correlate them with noise results and possible crystal defects. Theoretically, an approach, based on the fluctuation-dissipation theorem, is used in order to put numerical constraints on a model of 1/f noise caused by an internal (or structural) dissipation proportional to the amplitude and not to the speed. The order of magnitude of the noise is then discussed.

II. RESONATOR REALIZATION

All the resonators used in this study were fabricated from a single quartz crystal block. This crystal block was obtained from a seed cut in a previous synthetic crystal which was grown using a natural seed. Its dimensions were approximately 220 mm along the Y-axis, 36 mm along the Z-axis and 110 mm along the X-axis. Two Y-cut slices have been cut before and after an oriented block used to achieve ten quartz bars. These Y-cut slices were used for the evaluation of dislocations density, by X-ray topographies. Then, fourteen quartz bars pre-oriented on the first rotation angle have been cut. The length of these bars was about 70 mm. Taking into account the width of the cutting saw, about 24 resonators could be obtained in each bar. Finally, about a hundred of resonators were used for noise measurements, the others being set apart at various stages of the fabrication process, for analysis purposes.

The final prototype of the resonator is a typical 5 MHz SC-cut resonator. The diameter of the resonator is 14 mm for a thickness of 1.09 mm. A plano-convex shape allows the energy trapping for the 3rd overtone of the slowest thickness shear mode (C-mode). A radius of curvature of 130 mm has been chosen to optimize this energy trapping according to the Tiersten-Stevens model [2]. Electrodes diameter is 8 mm. The temperature turn over point of the resonator is chosen between 80°C and 85°C by adjusting the cutting angles.

Motional parameters have been measured in the configuration of the noise bench and compared to theoretical intervals estimated using Tiersten formula [2] for squared electrodes of 8 mm and 8/√2 mm sides. Table 1 shows the comparison of these theoretical intervals with the corresponding means and standard deviations measured for a sample of 100 resonators.
TABLE I. THEORETICAL AND EXPERIMENTAL PARAMETERS OF RESONATORS (R, L, C MOTIONAL PARAMETERS, T, TURNOVER TEMPERATURE, F₀ RESONANT FREQUENCY)

<table>
<thead>
<tr>
<th></th>
<th>R (Ω)</th>
<th>L (H)</th>
<th>C (fF)</th>
<th>Tᵣ (°C)</th>
<th>F₀ (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Th. min</td>
<td>69</td>
<td>5.47</td>
<td>0.159</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Th. max</td>
<td>79</td>
<td>6.24</td>
<td>0.183</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Exp. &lt; &gt;</td>
<td>74.7</td>
<td>5.53</td>
<td>0.183</td>
<td>77.7</td>
<td>4999988</td>
</tr>
<tr>
<td>Exp. σ</td>
<td>7</td>
<td>0.14</td>
<td>6E-3</td>
<td>2.8</td>
<td>7</td>
</tr>
<tr>
<td>Exp. Min</td>
<td>65.6</td>
<td>5.18</td>
<td>0.143</td>
<td>71.1</td>
<td>49999976</td>
</tr>
<tr>
<td>Exp. Max</td>
<td>149</td>
<td>5.8</td>
<td>0.195</td>
<td>84.4</td>
<td>4999999</td>
</tr>
</tbody>
</table>

The quality factor of each resonator was then computed from its measured motional parameters (Fig. 1).

The distribution of the Q values extends over an interval [1.1×10⁶, 3.2×10⁶] with a peak clearly visible between 2.25×10⁶ and 2.75×10⁶ (Fig. 2).

III. NOISE RESULTS

Fig. 3 gives the short-term stability floors of resonators (estimated by Allan standard deviation) grouped according to the number of the initial bar in which they were cut. The stability floors span approximately two orders of magnitude. The best resonators have a short-term stability (Flicker floor) below 8×10⁻¹⁴, whereas the worst are above 10⁻¹².

IV. DISCUSSION

A. Q-factor and noise results

Fig. 4 presents the noise of the resonators according to the Q-factor at room temperature. This figure confirms previous observations that the Q factor is not really a pertinent parameter to define the stability of the best quartz crystal resonators.
The quality factors of ten resonators, chosen for their different noise levels, were then measured at 4 K. Fig. 5 presents the noise level of these 10 resonators according to their Q-factor for the A, B and C 300 modes at 4 K and for the C 300 mode at room temperature.

Fig. 5. Short-term stability of quartz crystal resonators according to their Q factor measured at cryogenic temperature (4 K).

Fig. 5 shows that the noise level measured at room temperature seems to be independent of the quality factors Q measured not only at room temperature, but also at 4 K where the effect of internal defects (if any) on Q should be noticeable. Hence, these results show no clear trend for a relationship between the noise floor of the resonators and a low temperature behavior that could have been related to internal defects in the resonance volume.

B. Dislocation measurements

X-ray topography is commonly used to observe residual dislocations inside crystals. We have therefore carried out X-ray topography experiments on thin plates (“blanks”) cut perpendicularly to the Y-axis, which corresponds to the longest dimension of the crystal block. Fig. 6 presents the results obtained with two such plates. The duration of exposure of the photos is about 5h30 with an X-ray vertical beam given by a generator of 45 kV with 25 mA.

Fig. 6. X-ray topography of Y-cut plates.

The number of dislocations, observed inside the black rectangular zone with seed excluded, is visible. It is about 1 to 3 per cm² which corresponds to a very high quality quartz crystal.

C. Theoretical approach by Fluctuation-Dissipation Theorem

The fluctuation-dissipation theorem (FDT) was used to estimate the power spectral density of thermal noise coming from fluctuations in the thickness (2h) of quartz resonators [1]. An internal friction term, \( \varphi \), is added in the formulation of the fundamental principle of dynamics for continuum media in order to obtain a \( 1/f \) spectrum at low frequencies. The mechanical displacement inside the resonator \( u(x,t) \), can be written as:

\[
\rho \ddot{u}_x = c_{22} (1 + j \varphi) \frac{\partial^2 u_x}{\partial x^2} + \eta_{22} \frac{\partial^2 u_x}{\partial x^2} \frac{\partial u_x}{\partial t} \tag{1}
\]

With \( x \) the axis along the thickness of the resonator, \( c_{22} \) the elastic constant and \( \eta_{22} \) the viscoelastic damping constant of quartz crystal. \( \varphi \) is an internal friction coefficient [3]-[4], \( \rho \) is the quartz mass per unit volume.

Searching for harmonic solutions, the complex mechanical admittance of the system is defined by:

\[
\tilde{Y}(\omega) \equiv \frac{\theta(\omega)}{\theta(0)} = \frac{F/S}{1 + (c_{22} / \eta_{22} \omega) \sin(2\omega)} \tag{2}
\]

with \( F \) the modulus of the harmonic mechanical force applied to the surface \( S \) of the electrodes (perpendicular to the thickness of the resonator) and:

\[
k^2 = \frac{p \omega^2}{(c_{22} / \eta_{22} \omega)} \tag{3}
\]

The FDT then states that the spectral power density of the thickness fluctuations in a bandwidth \( BW \), can be computed by:

\[
\frac{u_x^2(\omega)}{BW} = \frac{4k_B T}{\omega^2} \Re\{\tilde{Y}(\omega)\} \tag{4}
\]

with T the absolute temperature (in K) and \( k_B \) the Boltzmann constant (in J/K).

The assumptions \( \varphi \ll 1 \) and \( \omega \ll c_{22} \eta_{22} / \omega \) lead to:

\[
\frac{u_x^2(\omega)}{BW} \approx \frac{4k_B T}{\omega^2} \left( \eta_{22} \omega + c_{22} \varphi \right) \tag{5}
\]

Moreover, we can consider that the circular frequency at resonance \( \omega_0 \sim 1/\omega \), thus:

\[
S_y(\omega) \equiv \frac{(\delta u_x)^2}{\sigma_{y,\text{flicker}}^2} \approx \frac{u_x^2(\omega)}{(2h)^2 BW} \approx \frac{1}{\omega} \times \frac{2k_B T}{c_{22}} \left( \frac{\eta_{22} \omega}{c_{22}} + \varphi \right) \tag{6}
\]

where \( V \) is the volume of the resonator. One can then see from the previous expression that for circular frequencies lower than \( \varphi c_{22}/\eta_{22} \), the internal friction becomes dominant and gives a \( 1/f \) spectrum, with an Allan standard deviation given by:

\[
\sigma_{y,\text{flicker}} = \sqrt{\frac{2}{V} \frac{2k_B T}{c_{22}} \varphi} \tag{7}
\]

We note that \( \varphi \) could depend upon the temperature and that no assumption where made about this possible dependence. We consider here numerical values typical for a 5 MHz oscillator equipped with an SC-cut quartz crystal resonator. Due to the rotation of the axis, the 2 axis is not the usual one, so that the constants must be evaluated in the rotated basis, \( c_{22} = 115 \text{ GPa}, \eta_{22} = 1.36 \times 10^{-3} \text{ Pa s}, T = 350 \text{ K} \) and \( V = 0.104 \text{ cm}^3 \). This gives:
\[ \sigma_{y, \text{flicker}} \approx 1.06 \times 10^{-12} \sqrt{\phi} \quad (8) \]

In order to recover measured values of \( \sigma_{y, \text{flicker}} \) with this expression, we would need to have \( \phi \) between \( 10^{-2} \) and \( 10^{-4} \), which would mean that even at resonance the internal damping would be dominant over viscoelastic damping. We therefore conclude that internal damping of thickness fluctuations by any force proportional to strain and independent of frequency, may not be the dominant noise mechanism for the best SC-cut quartz resonators. However, other modes may be noisier…

Nonetheless, we try to evaluate this coefficient by the modified Granato-Lücke theory [5] of the energy loss due to some kinds of dislocation motion in the low frequency range. In [5], it is supposed first that the pinning force \( F \) of the impurity atom which arises from elastic interactions depends on the orientation of the dislocation line. Second, it is supposed that, once a dislocation has broken away from its pinning points, its motion is not necessarily limited by its line tension, but that the distance it moves may be determined by the stress field of neighboring impurity atoms. With these assumptions, the authors found an expression of the decrement for the impurity spacing controlled dislocation motion that, in the small strain amplitude limit, is given by:

\[ \Delta = \frac{6 N b N_\Lambda}{\pi c \Omega} \quad (9) \]

where \( \beta \) is a parameter having approximate value of 1.5. \( N \) is the total length of dislocation line in a unit volume of material \( \approx 2 \times \text{surface dislocation density} \). This value is of the order of \( 6 \text{ cm/cm}^2 \) judging from the X-ray image of the surface of one of the resonator (Fig. 6). \( b \) is the mean length of a Burger’s vector \( \approx 3 \times 10^{-8} \text{ cm} \). \( L_\Lambda \) is the network length \( = \sqrt{3/N} \). \( c \) is the atom fraction of impurity which must be lower than 1 ppm to get \( Q \) values as high as a few \( 10^6 \). \( \epsilon \) is the fractional difference between the radius of impurity and host atoms taken to be of the order of 20 \%. Finally, from Eq. (6), we can deduce that:

\[ \frac{1}{Q_{\text{eff}}} = \frac{1}{Q_{\text{viscous}}} + \phi \quad (10) \]

with \( \frac{1}{Q_{\text{viscous}}} = \frac{\eta_{\text{max}}}{c^2 z_2} \).

Therefore at low frequencies \( 1/Q_{\text{eff}} \approx \phi \). Hence, we attempt to identify \( \Delta \) with \( \pi \phi \) at low frequencies, in a first approximation in spite of the fact that we are not in the dominantly viscous regime. This would give \( Q_{\text{eff}} \approx 10^3 \) and \( \phi \approx 10^{-5} \) in the low frequency regime, which would be an interesting order of magnitude to attribute at least some non-negligible part of the 1/f noise to the fluctuations of thickness. However, this would also mean that at resonance:

\[ \frac{1}{Q_{\text{eff}}} + \phi \approx 4 \times 10^{-7} + 10^{-5} \approx 10^{-5} = \phi \quad (11) \]

hence that the viscous damping would not be dominant at resonant frequency which is contradictory to experimental facts.

V. CONCLUSION

The short-term stability of ultra-stable resonators has been studied according to the position of the blanks in the mother crystal. The short term stability of several resonators has been measured lower than \( 8 \times 10^{-14} \). No clear correlation between the blank position and the quality of the resonator can be seen. It is shown also that Q-factor at room temperature cannot be a precise parameter to predict the noise of a resonator. At cryogenic temperature the number of measurements is not sufficient to give a clear conclusion on the influence of the internal defects.

On the theory side, it is possible to find 1/f noise through the fluctuation-dissipation theorem, by adding a constant complex part to the elastic constant in the usual differential equation characteristic of a viscously damped harmonic oscillator. This corresponds to a frequency independent energy loss in the limit of small frequencies. The hysteretic motion of the dislocations described by a modified Koehler-Granato-Lücke theory could provide some priori description of such a loss mechanism. Indeed, it could provide an explanation for the experimental observations that the logarithmic decrement generally decreased when the dislocation density decreased when quartz were not as good as now and that sometimes a slightly higher concentration of impurity could improve the quality factor. However, numerical estimations seem to provide values that are at least an order of magnitude too high.

REFERENCES


Estimation of the light shift in Ramsey-Coherent Population Trapping

Yuichiro Yano1, Shigeyoshi Goka1 and Masatoshi Kajita2
1Graduate School of Science and Engineering, Tokyo Metropolitan University, Hachioji, Tokyo Japan
2National Institute of Information and Communications Technology, Koganei, Tokyo Japan
E-mail*: y-yano@tmu.ac.jp

Abstract—We both numerically and analytically investigate on the light shift in Ramsey-CPT resonance for compact atomic clocks. The numerical calculation of the light shift is based on a density matrix analysis in $\Lambda$-type three state model. And the analytical method leads to an estimation equation of light shift in the Raman–Ramsey scheme. The estimation equation expresses the relationship between light shift and all pulse parameters. The results show that the estimation equation was reproduced well as the numerical results.

Keywords—Coherent population trapping, Raman–Ramsey scheme, light shift, estimation equation.

I. INTRODUCTION

Atomic clocks based on coherent population trapping (CPT) resonance have attracted much attention as means of fabricating very small frequency references, such as chip-scale atomic clocks. CPT atomic clocks are in great demand for many applications, such as telecommunications, synchronization of networks and navigation systems [1]. And such clocks are required for high frequency stability.

The light shift is one of the limitations for long-term frequency stability in CPT atomic clocks [2]. The Raman–Ramsey scheme significantly reduces the light shift to one or two orders of magnitude lower than that under continuous wave irradiation [3; 4; 5]. In our previous work, we investigated the light shift in the Raman–Ramsey scheme from both theoretical and experimental sides [6]. However, the relationship between the light shift and the pulse parameters, especially in terms of the free evolution time $T$ and excitation duration time $\tau$, is still unclear. In this work, we numerically and analytically investigate the light shift in Ramsey-CPT for enhancing long-term stability of Ramsey-CPT atomic clocks. The analytical method leads to an estimation equation of light shift in the Raman–Ramsey scheme. The estimation equation expresses the relationship between light shift and all pulse parameters. The results show that the estimation equation is reproduced well as the numerical results.

II. THEORETICAL MODEL

Figure 1 (a) shows the excitation scheme using a left circular ($\sigma^+$) polarized light field on the $^{133}$Cs-D$_1$ line. In the CPT phenomenon, two-ground states of the $^6S_{1/2}$ are coupled to a common excited state of the $^6P_{1/2}$, simultaneously.

In this system, the dynamical behavior of the density matrix $\rho$ is governed by the quantum Liouville equation,

$$\frac{\partial}{\partial t} \rho(t) = i\frac{\hbar}{\hbar} [\rho, H] + R\rho,$$

where $H$ is the Hamiltonian matrix for this three level system and $R$ stands for the relaxation terms. Using the rotating wave approximation with the simplified $\Lambda$-type model shown in Fig. 1 (b), Eq. (1) can then be rewritten as
where $S_1$ and $S_2$ are the light shifts under continuous irradiation. These light shift terms are the sum of the light shift caused by the higher order sidebands. The light shifts $S_1$ and $S_2$ of the two ground states can be calculated by using perturbation method.

III. Numerical calculation method

The light shift is derived from the solution of Eq. (2). The easiest way to solve the time-dependent behavior of the density matrix is to calculate the numerical integration by finite-difference methods. However, the finite-difference methods require computation time because the time step of the finite-difference method is needed to be smaller than the relaxation time $\Gamma^{-1}$ in order to reduce computation error. In this calculation, we show the numerical solution of the density matrix using the eigenvector. Since this numerical solution is not required the numerical integration, it can reduce both computation time and error.

The number of calculation elements is a total of 9, because density matrix is Hermitian matrix of size $3 \times 3$. For the sake of simplicity of differential equation, a vector of 9 elements $\hat{\rho}$ is defined as follows:

$$
\hat{\rho} := \begin{pmatrix}
\rho_{11} \\
\rho_{22} \\
\rho_{33} \\
\Re(\rho_{12}) \\
\Re(\rho_{13}) \\
\Re(\rho_{23}) \\
\Im(\rho_{12}) \\
\Im(\rho_{13}) \\
\Im(\rho_{23})
\end{pmatrix}
$$

(6)

Using the vector $\hat{\rho}$, Eq. (2) is rewritten as follows:

$$
\frac{\partial \hat{\rho}}{\partial t} = \tilde{H}\hat{\rho}
$$

(7)

In the Raman–Ramsey scheme, $\tilde{H}$ under excitation duration time is different from that under free evolution time. The matrix $\tilde{H}$ at pulse on and off are defined as $\tilde{H}_{\text{on}}$ and $\tilde{H}_{\text{off}}$, respectively. And the vector at the boundary (vector at pulse rise $\hat{\rho}_{\text{on}}$ and fall $\hat{\rho}_{\text{off}}$) is defined as follows:

$$
\begin{cases}
\hat{\rho}_{\text{on}} = \hat{\rho}(n(\tau + T)) \\
\hat{\rho}_{\text{off}} = \hat{\rho}(n(\tau + T) + \tau)
\end{cases}
$$

where, $n$ is integer. Because $\hat{\rho}_{\text{on}}$ is vector after $T$ second from the vector at pulse fall $\hat{\rho}_{\text{off}}$, and $\hat{\rho}_{\text{off}}$ is vector after $\tau$ second from the vector at pulse rise $\hat{\rho}_{\text{on}}$, the relationship between $\hat{\rho}_{\text{off}}$ and $\hat{\rho}_{\text{on}}$ can be expressed as follows:

$$
\begin{cases}
\hat{\rho}_{\text{on}} = \exp(\tilde{H}_{\text{off}} T)\hat{\rho}_{\text{off}} \\
\hat{\rho}_{\text{off}} = \exp(\tilde{H}_{\text{on}} \tau)\hat{\rho}_{\text{on}}
\end{cases}
$$

(8)

where, exp is an exponential function of the matrix.$^1$

---

$^1$Let the initial vector is $\hat{\rho}_i$, a vector after time evolution $\hat{\rho}(t)$ is expressed as

$$
\hat{\rho}(t) = \exp(\tilde{H}t)\hat{\rho}_i
$$
rearranging the Eqs. (8), a following equation can be obtained.

\[
\left( E - \exp(\tilde{H}_{\text{off}}T) \exp(\tilde{H}_{\text{on}}\tau) \right) \tilde{\rho}_{\text{on}} = 0
\]  

(9)

where, E is identity matrix. From Eq. (9), because \( \tilde{\rho}_{\text{on}} \) is not a null vector, \( \tilde{\rho}_{\text{on}} \) is the eigenvector of the eigenvalue 0 of \( \left( E - \exp(\tilde{H}_{\text{off}}T) \exp(\tilde{H}_{\text{on}}\tau) \right) \). Since all the matrix elements are known, the vector at pulse rise \( \tilde{\rho}_{\text{on}} \) can be derived by calculating the eigenvector of the matrix. Note that the vector is normalized to satisfy the normalization condition of Eq. (3).

Since the vector at the measurement time \( \tilde{\rho}_{m} \) is a vector after \( \tau_{m} \) second from pulse rise, the vector at the measurement time \( \tilde{\rho}_{m} \) is obtained as follows

\[
\tilde{\rho}_{m} = \tilde{\rho}(n(\tau + T) + \tau_{m}) = \exp(\tilde{H}_{\text{on}}\tau_{m})\tilde{\rho}_{\text{on}}
\]  

(10)

The light shift was determined assuming that the dispersion spectrum given by the density matrix analysis is zero. The zero dispersion from Eq. (10) was found using Newton’s method.

IV. ANALYTICAL METHOD

Assuming the decoherence rate between two ground states \( \gamma_{a} \) is zero, we found that the time-dependent behavior of light shift follows Logistic differential equation [8]

\[
\frac{\partial S}{\partial t} = \frac{1}{\tau_{p}} \left( \frac{S_{21} - S}{S_{21}} \right) S
\]  

(11)

where, \( \tau_{p} \) is a pumping time that is inversely proportional to light intensity \( I \). \( S_{21} \) is a light shift under continuous irradiation (= \( S_{2} - S_{1} \)).

Under the excitation duration time \( \tau \), let the light shift at pulse rise \( (t = 0) \) is \( S_{\text{off}} \), the light shift at pulse fall \( (t = \tau) \) \( S_{\text{on}} \) is simple logistic function from Eq. (11), from \( S_{\text{off}} = S_{\text{on}} \) with \( \tau = 0 \)

\[
S_{\text{off}} = \frac{S_{21}}{\left( \frac{S_{21}}{S_{\text{on}}} - 1 \right) e^{-\tau/\tau_{p}} + 1}
\]  

(12)

And the exponential function is expressed as

\[
\exp(\tilde{H}t) = E + \frac{1}{2} \tilde{H}^{2}t^{2} + \frac{1}{6} \tilde{H}^{3}t^{3} \ldots + \frac{1}{n!} \tilde{H}^{n}t^{n} \ldots
\]

Seeing \( S_{21} \propto I \) and \( \tau_{p} \propto I^{-1} \), value of \( S_{21}\tau_{p} \) under the irradiation is valid also for \( I \rightarrow 0 \) with free evolution, Eq. (11) is given by

\[
\frac{\partial S}{\partial t} \bigg|_{\tau = 0} = -\frac{1}{\tau_{p}} S^{2}.
\]  

(13)

From Eq. (13) and \( S_{\text{on}} = S_{\text{off}} \) with \( T = 0 \), \( S_{\text{on}} \) is obtained as follows

\[
S_{\text{on}} = \frac{S_{21}}{S_{21} + \frac{T}{\tau_{p}}}
\]  

(14)

From Eqs (12) and (14), \( S_{\text{on}} \) can be rewritten as follows

\[
S_{\text{on}} = \frac{\tau_{\text{eff}}}{\tau_{\text{eff}} + \frac{T}{\tau_{p}}} S_{21}
\]  

(15)

where the factor \( \tau_{\text{eff}} \) is

\[
\tau_{\text{eff}} = \tau_{p} \left( 1 - e^{-\tau/\tau_{p}} \right)
\]  

(16)

Since the light shift at observation time \( S_{m} \) is the light shift after \( \tau_{m} \) second from pulse rise, from Eq. (11), the light shift at observation time is obtained as follows

\[
S_{m} = \frac{\tau_{\text{eff}}}{\tau_{\text{eff}} + \frac{T}{\tau_{p}}} S_{21}
\]  

(17)

This equation expresses the relationship between light shift and all pulse parameters.

V. RESULTS

A. Light shift as a function of light intensity under different excitation duration time

Figure 3 shows the light shift as a function of light intensity under different excitation duration time \( \tau \). For the sake of simplicity, the light shift in absence of the influence of observation \( S_{\text{on}} \) is plotted. When small \( \tau \) \( (\tau \ll \tau_{p}) \) is set, because \( \tau_{\text{eff}} \) approaches to \( \tau \), \( S_{\text{on}} \) can be written as follows

\[
S_{\text{on}} \big|_{\tau \ll \tau_{p}} = \frac{\tau}{\tau + T} S_{21}
\]  

(18)

From Eq. (18), the light shift under small \( \tau \) is time averaging light shift. Because the coefficient \( \tau/(\tau + T) \) is not dependent on light intensity, the light shift under small \( \tau \) is linearly proportional to light intensity and the variation of the light shift is proportional to excitation duration time \( \tau \). When large
\( \tau \) (\( \tau \gg \tau_p \)) is set, because \( \tau_{\text{eff}} \) approaches to \( \tau_p \). \( S_{\text{on}} \) can be written as follows

\[
S_{\text{on}} \bigg|_{\tau \gg \tau_p} = \frac{\tau_p}{\tau_p + T} S_{21} \tag{19}
\]

Since the coefficient \( \tau_p/(\tau_p + T) \) is dependent on light intensity, the light shift is nonlinear function of light intensity. The nonlinearity of the light shift is experimentally-ascertained in previous work [6]. The light shift in the Raman–Ramsey scheme has two characteristics. The first characteristic is that the variation in the Raman–Ramsey scheme is equal to that under continuous irradiation with low light intensity. From a physical perspective, this is because the effective light field applied in the Raman–Ramsey scheme is the same as that applied under continuous irradiation because low light intensity leads to a long coherence time. From Eq. (19), under low light intensity condition, because the second term of the denominator is negligible, \( S_{\text{on}} \) converges on \( S_{21} \). The second characteristic is that the light shift in the Raman–Ramsey scheme approaches a saturation value as light intensity approaches infinity. Under high light intensity, because the first term of the denominator of Eq. (19) is negligible, \( S_{\text{on}} \) is obtained as follows

\[
\lim_{T \to \infty} S_{\text{on}} \bigg|_{\tau \gg \tau_p} = \frac{\tau_p}{T} S_{21} \tag{20}
\]

Therefore, the variation of light shift monotonically decreases with increasing light intensity.

Though the systematic frequency shift decreases with decreasing \( \tau \), the light shift variation of the excitation duration time \( \partial S/\partial \tau \) increases with decreasing \( \tau \) because the \( \tau_{\text{eff}} \) variation increases with decreasing \( \tau \) from Eq. (16). In addition, a short duration pulse leads to a poor signal amplitude of Ramsey-CPT resonance because of increasing the population of the dark state at pulse fall. Therefore, a long excitation duration time \( \tau \) (\( \tau \gg \tau_p \)) is required for enhancing the frequency stability of Ramsey-CPT atomic clocks.

**Fig. 3.** Numerical and analytical results of light shift as a function of light intensity under different excitation duration time \( \tau \); the parameters are \( T = 800 \mu s, S_{21} = 30.9 \text{ Hz/(mW/cm}^2\text{)} \times I, \tau_p = 198 \mu s\text{-}(\text{mW/cm}^2\text{)} \times I^{-1} \), and the difference between numerical and analytical results is no more than 0.02 %.

**B. Light shift as a function of light intensity under different free evolution time \( T \)**

Figure 4 shows the light shift as a function of light intensity under different free evolution time \( T \). When \( T = 0 \) is set, the light shift becomes equal to the that under continuous irradiation \( S_{21} \)

\[
S_{\text{on}} \big|_{T=0} = S_{21} \tag{21}
\]

When \( T > 0 \) is set, because the denominator of Eq. (15) increases with increasing free evolution time \( T \), the light shift decreases with decreasing the free evolution time \( T \). Therefore, a long free evolution time is required to reduce the systematic frequency shift.

The saturation value is inversely proportional to the free evolution time \( T \) from Eq. (20). This tendency is consistent with the behavior of the light shift calculated by the adiabatic approximation [9; 10; 11]. While the light shift calculated by the adiabatic approximation diverges to infinity as the free evolution time \( T \) approaches zero, the estimation equation calculated in this work converges on finite value.

A comparison of the results of the numerical calculations performed by density matrix analysis and the estimated values from Eq. (17) shows that the relative error of the results is no more than 0.05 %. Therefore, the proposed estimation equation is useful for practical purposes.

**VI. Conclusion**

This paper presents the investigation on the light shift in Ramsey-CPT resonance for compact atomic clocks. We numerically calculated the light shift using a density matrix analysis in A-type three state model, and also analytically calculated the estimation equation of the light shift. The analytical method leads to an estimation equation of light shift in the Raman–Ramsey scheme. The estimation equation expresses the relationship between light shift and all pulse parameters. The results show that the estimation equation was reproduced well as the numerical results.

For the further work, we will experimentally measure the light shift in Ramsey-CPT resonance. And the estimation
equation is modified to include the light shift induced by a relaxation between ground states $\gamma_s$.

**ACKNOWLEDGEMENTS**

This work was supported by a Grant-in-Aid for JSPS Fellows (No. 26-6442)

**REFERENCES**


Abstract—We propose a coherent population trapping (CPT) pulsed-excitation method based on direct drive-current modulation of a vertical-cavity surface-emitting LASER for miniature atomic clocks (MACs). Because our method needs only a simple device configuration without optical external modulators, it is suitable for MACs, including chip-scale atomic clocks. Reducing the size, cost, and power consumption is also possible. Additionally, we report on experimental results for Ramsey-CPT fringes excited using the proposed method.

Keywords—coherent population trapping (CPT); miniature atomic clocks (MACs); chip-scale atomic-clocks (CSACs); pulsed excitation

I. INTRODUCTION

Miniature atomic clocks (MACs) with low power consumption and good frequency stability are required for portable equipment such as global-positioning-system receivers, in-field telecommunication devices, and various types of measuring instruments [1]. A coherent population trapping (CPT) resonance, which can be observed with a small gas-cell and a vertical-cavity surface-emitting LASER (VCSEL), are able to be used in MACs because a large microwave cavity is no longer needed [2]. Because recent CPT-based commercial MACs have relatively good frequency characteristics with small size and low power consumption, better performance is nevertheless required for commercial frequency generators.

Frequency stabilities in atomic clocks are categorized by the Allan standard deviation as either short-term or long-term. Short-term stability can be estimated by the product of $Q$ value and signal-to-noise ratio of the atomic resonance, thus, a narrow full-width at half-maximum (FWHM) and a high signal-to-noise ratio are desired [3]. For long-term stability, light shifts are well-known to be a major limitation to the long-term stability of vapor cells [4,5]. The way to improve both short-term and long-term frequency stability is to use the Ramsey-CPT resonance by employing the Raman-Ramsey scheme. The Ramsey-CPT resonance has been studied for high-performance CPT atomic clocks [6,7]. By applying the Raman-Ramsey scheme to the CPT resonance, both a narrower FWHM and a low light shift appear with the CPT resonance that improves frequency stability [8]. Because two pulsed laser beams with stable wavelengths are required in the observation of Ramsey-CPT resonances, an external optical modulator (OM) such as an acoustic optical modulator (AOM) is typically used to generate the pulsed laser lights. AOMs generally take up a large volume and consumes much power compared with the requirements of a MAC, and therefore, Ramsey-CPT resonance has not been used in MACs.

In the paper, we propose a Ramsey-CPT observation method without the external OM. In our method, pulsed laser beams are generated by a direct modulation of the VCSEL drive current. In general, modulating the drive current causes significant wavelength variation of the pulsed laser beam, therefore, a two-step pulse current method is applied to control the variation and observe the Ramsey-CPT resonance. We report the experimental results for Ramsey-CPT resonance excited by the proposed method.

II. CPT PULSE EXCITATION METHOD

The Ramsey-CPT resonance is observed by the ON-OFF-pulsed laser irradiation of the atoms [9]. The line width of the Raman-Ramsey fringe is determined by the free evolution time $T$ (pulse OFF time). Because the atoms are gradually pumped into the steady CPT state, the observation time interval $r_n$ from the beginning of irradiation needs to be as short as possible. Nevertheless, a sufficient number of atoms needs to be pumped into the CPT state by the end of the laser pulse as preparation for the Raman-Ramsey scheme. Therefore, the external OM has been used to generate the laser pulse with stable wavelength in conventional observation.

With direct VCSEL-current modulation to generate the pulsed laser, the output light intensity does vanish completely, which provides better conditions for Ramsey-CPT excitation. The power consumption is also better than when using the conventional excitation. However, the output laser wavelength is changed by the internal temperature variation in the VCSEL caused by the current modulation. Therefore, it takes time to reach the desired wavelength from the beginning of the ON pulse phase. It is also difficult to choose shorter observation time intervals. In consequence, direct modulation has been considered unsuitable to observe the Ramsey-CPT resonance.

To solve these problems, we split the VCSEL drive current into the two stages (Fig.1). In the first-stage, current $I_1$ with a duration of $\tau_1$ is set to a higher value than the normal excitation current to warm the active region of the VCSEL rapidly; hence, we can improve the rise time of the wavelength variation. In the second-stage, current $I_2$ is adjusted so that the output wavelength equals the absorption line. The laser pulse with a duration of $\tau_2$ irradiates the atoms for pumping the steady CPT
state. In this manner, we can observe the Ramsey-CPT resonance by direct-current modulation.

III. EXPERIMENTAL SETUP

The Ramsey-CPT measurement system with direct-current modulation (Fig. 2) comprises of a cylindrical gas cell with diameter of 20 mm and optical length of 10 mm containing a mixture of $^{133}$Cs atoms and Ne buffer gas at a pressure of 4.0 kPa. Its temperature was maintained at 42°C. The VCSEL operates by excitation of the $^{133}$Cs D$_1$ line with a wavelength of 895 nm, and was driven using a two-step pulse current generated by the current driver. The current values were controlled by the FPGA with D/A converter. The VCSEL was also modulated at 4.6 GHz using a RF generator via a bias tee to generate two first-order sidebands. The RF power was adjusted so that the signal-to-noise ratio of the CPT resonance is maximum. In the experiment, we chose the excited state $|F'=3>$ because the clock transition amplitude is higher than that for $|F'=4>$ when using linearly polarized laser beam [10]. A 6-mm-diameter beam was used to irradiate the gas cell; a photodiode was used to detect the light transmitted through the cell. To observe the Ramsey fringe, a sample-and-hold circuit sampled the detected signal after the observation time $\tau_m$ had passed. This signal was also used for current control of the VCSEL to maintain the desired wavelength.

IV. EXPERIMENTAL RESULTS

The transmitted light signal (Fig. 3) shows the normal one-step pulse current with the current on time $\tau_{ON}$ set at 5 ms, and the current off-time $T$ set at 600 $\mu$s. In direct modulation, multiple absorption lines generated by several sidebands of the RF modulated laser are observed depending on the wavelength variation. In Fig. 3(a), the labels $|F'=3>$ and $|F'=4>$ indicate the absorptions with the two first-order sidebands matching the frequency differences between the two ground states and the common excited states $|F'=3>$ and $|F'=4>$, respectively. The other minima of the absorption curve are the combination of career and second-order sideband or first-order and third-order sidebands matching the frequency differences between the two ground states. In direct modulation using a normal pulse current, the wavelength rise time until the desired wavelength corresponding to $|F'=3>$ from the beginning of the laser irradiation was about 2500 $\mu$s. In consequence, the Ramsey-CPT resonance cannot be observed because the Ramsey fringe disappeared in this rise time.

Figure 4 shows the transmitted light signal from the proposed two-step pulse current with $\tau_{ON} = 5$ ms, $T = 600$ $\mu$s, and the time of first-stage current was set as $\tau_1 = 25$ $\mu$s. By adding the high current stage, the rise time of the wavelength up until the desired value was significantly improved. In our method, the wavelength of the output passed through the desired value once, then reached its value again. The rise time
to the desired wavelength was about 20 μs. In comparing with the normal pulse current, the raise time was reduced by a factor of about one hundred and twenty five. By observing at the time when the desired wavelength is first reached, we can observe the Ramsey-CPT resonance.

The Ramsey fringe observed by the proposed method is shown in Fig. 5. The ordinate gives the normalized value with amplitude of each resonance. In addition to the resonance observed by the OM, the full width at half maximum (FWHM) diminishes with increasing free-evolution time. For a CPT-based clock, the contrast, which is defined as the amplitude of CPT resonance over the background signal level, is used as the signal-to-noise ratio. Figure 6 shows the dependence of the measured contrast of the Ramsey-CPT resonance on the free evolution time $T$. As $T$ increases, the contrast decreases. This is because the number of atoms in the CPT states decreases as $T$ increases by relaxing the wall collisions [11]. In the range $T < 400$ μs, the contrast is higher than that for continuous excitation of 2.7 %. The contrast at $T = 200$ μs was 3.4 %, which is 1.2 times better compared with that for continuous excitation. Figure 7 shows the dependence on the free evolution time of the measured FWHM, which decreases because the line width of the Ramsey fringe depends on the free evolution time as a function of $1/2T$ regardless of beam power. The FWHM at $T = 200$ μs is 450 Hz, which is roughly one-tenth that of the continuous excitation (4.7 kHz).

The figure of merit, which is the product of the $Q$ value and the contrast, was estimated from the measured data. Figure 8 shows the figure of merit normalized by the value of the continuous excitation. As the short-term stability can be estimated by the figure of merit, its value above unity shows an improvement in the short-term stability. The curve has a maximum value at $T = 600$ μs, which indicates a better performance by factor 5.7 compared with the conventional continuous excitation.
A CPT pulsed-excitation method was proposed based on direct drive-current modulation of a VCSEL suitable for miniature atomic clocks. Experiments showed that it was possible to observe the Ramsey-CPT resonance by the two-step pulse current to the VCSEL. By controlling the output wavelength, the figure of merit was 5.7 times better than that for conventional continuous excitation. These results indicate that our method can improve the short-term stability of CPT-based MACs.

Acknowledgment

The authors are grateful to Ricoh Company, Ltd. for providing us with the Cs-D₁ VCSEL.

References

Preliminary results of a Cs vapor cell CPT clock using push-pull optical pumping

Moustafa Abdel Hafiz1 and Rodolphe Boudot1
1FEMTO-ST, CNRS, Université de Franche-Comté,
26 chemin de l’épitaphe 25030 Besançon, France
Email: rodolphe.boudot@femto-st.fr

Abstract—This paper reports on preliminary frequency stability performances of a compact CPT-based Cs vapor cell atomic clock using the push-pull optical pumping technique (PPOP). This clock uses a single distributed feedback laser source externally modulated with a Mach-Zehnder intensity electro-optic modulation for optical sidebands generation. The clock short-term frequency stability is measured at a level of $3.0 \times 10^{-13} \tau^{-1/2}$ up to 80 s averaging time, in good agreement with the signal-to-noise ratio limit. These short-term frequency stability results are encouraging and close to those of best vapor cell atomic clocks.

I. INTRODUCTION

Vapor cell atomic clocks are an interesting technology because they combine compactness, low power consumption and excellent relative frequency stability. Over the last years, thanks to the progress of semiconductor lasers and the use of dedicated techniques, state-of-the-art laboratory-prototype vapor cell atomic clocks, based on optical-microwave double resonance technique [1], [2] or coherent population trapping (CPT) [3], have demonstrated short-term frequency stability in the $1 - 4 \times 10^{-13}$ range. These performances are about two orders of magnitude better than commercially-available Rb clocks and make them an alternative solution to bulky hydrogen masers for averaging times up to 10 000 s.

In this domain, LTF-UNINE has demonstrated a continuous-regime (CW) Rb clock with a relative frequency stability of $2.4 \times 10^{-13} \tau^{-1/2}$ up to averaging times of 10 000 s [4]. INRIM has developed a pulsed optically pumped Rb frequency standard with microwave [5] and optical detection [6] that exhibit a fractional frequency stability of $1.2 \times 10^{-12} \tau^{-1/2}$ and $1.7 \times 10^{-13} \tau^{-1/2}$ up to 10 000 s integration time respectively. Recently, a chinese group inspired by the work of Zhu et al. [7], proposed a polarization-selective method for the signal detection in such systems. This technique allowed to demonstrate resonance contrasts up to 90 % and a frequency stability of $2 \times 10^{-12}$ up to 1000 s averaging times [8]. INRIM has developed a CPT maser with a short term relative frequency stability of $3 \times 10^{-13} \tau^{-1/2}$, value close to the theoretical prediction [9]. To the best of our knowledge, the best CPT clock worldwide is the pulsed Cs vapor cell CPT clock developed in LNE-SYRTE, Paris. This clock allows the detection of high-contrast and narrow Raman-Ramsey fringes thanks to the combination of a specific lin-lin interaction scheme and a pulsed interrogation [10]. Recently, this pulsed Cs CPT clock has demonstrated a short term fractional frequency stability of $3.2 \times 10^{-13}$ up to averaging times of 1000 s [11]. A drawback of this table-top prototype clock, in terms of compactness requirements of potential industrial transfer, is the use two phase-locked lasers that make the system voluminous and complex.

In the frame of the MClocks project [12], we recently started in FEMTO-ST the development of a high-performance compact Cs CPT atomic clock based on the push-pull optical pumping technique, pioneered proposed by Jau et al. [13]. We reported the spectroscopy of high-contrast CPT resonances in Cs vapor cells [14] and demonstrated the possibility to detect high-contrast Raman-Ramsey fringes [15]. This article aims to report first frequency-stability results of the CPT-PPOP based Cs vapor cell atomic clock. A promising short-term frequency stability of $3 \times 10^{-13}$ at 1 s is measured.

II. EXPERIMENTAL SETUP

Figure 1 presents the Cs CPT clock experimental set-up.

The laser source is a distributed-feedback (DFB) diode laser tuned at 894 nm on the Cs D1 line [16]. A MZ EOM (Photline NIR-MX800-LN-10) driven at 4.596 GHz with a low noise microwave frequency synthesizer by [?] is used to generate phase-coherent optical sidebands frequency-separated by 9.192 GHz. At the output of the EOM, the optical carrier rejection is actively stabilized thanks to an original microwave synchronous detector presented in [14]. The laser beam at the output of the EOM is sent into an annex reference Cs cell. Saturated absorption spectroscopy is used to stabilize the laser frequency. At the output of the EOM, an acousto-optic modulator (AOM) is used for two main functions. Its first role is to compensate for the optical frequency shift due to the presence of buffer gas in the CPT cell. This optical
The clock resonance is well-fitted by a lorentzian function. The CPT linewidth $\Delta \nu$ is 564 Hz. The clock signal $S$, defined on Fig. 2 as $S = H - y_0$, is 0.114 V. The resonance contrast $C_r$, defined as the ratio between the CPT signal $S$ and the dc background $y_0$, is 22%. The discriminator slope $D = S/\Delta \nu$ is measured to be 0.2 mV/Hz.

Figure 3 plots the Allan deviation of the clock frequency. The latter is measured to be $3 \times 10^{-13}$ at 1 s, going down to about $3.2 \times 10^{-14}$ at 100 s. A large bump, from a few seconds to hundreds seconds averaging time, presently prevents the Allan deviation to decrease with a perfect white frequency noise slope. This is currently under investigation but could be explained by thermal effects (the laboratory room is not temperature stabilized), a non-optimized gain of the EOM bias voltage servo loop and feedback from the EOM fiber input face that perturbs the laser frequency. An additional optical isolation stage should be added in a near future. Nevertheless, these short-term frequency stability performances, comparable to those of best vapor cell atomic clocks, are very encouraging for the development of a high-performance Cs vapor cell CPT atomic clock.

Table shows main contributions to the clock short term frequency stability at $\tau = 1$ s. The noise budget is in excellent agreement with the measured clock fractional frequency stability. The clock frequency stability is currently mainly limited by the laser AM noise and the laser FM-AM noise process. The following contribution is the Dick effect. The latter will be reduced later thanks to a novel ultra-low noise frequency synthesis chain [18] that should reject the Dick effect contribution to a level lower than $7 \times 10^{-14}$ at 1 s.

### TABLE I. Main contributions to the clock short term frequency stability at $\tau = 1$ s.

<table>
<thead>
<tr>
<th>Noise Source</th>
<th>$\sigma(1 \text{s})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shot noise</td>
<td>$5.8 \times 10^{-14}$</td>
</tr>
<tr>
<td>Detector noise</td>
<td>$3 \times 10^{-14}$</td>
</tr>
<tr>
<td>LO phase noise</td>
<td>$1.1 \times 10^{-13}$</td>
</tr>
<tr>
<td>Laser AM noise</td>
<td>$2.8 \times 10^{-13}$</td>
</tr>
<tr>
<td>Laser FM-AM noise</td>
<td>$1.7 \times 10^{-13}$</td>
</tr>
<tr>
<td>Total $\sigma_{1\text{s}}$</td>
<td>$3.02 \times 10^{-13}$</td>
</tr>
</tbody>
</table>

The short-term relative frequency stability $\sigma_y(\tau)$ of a passive atomic clock is well-approximated by [17]:

$$\sigma_y(\tau) \sim \frac{\Delta \nu}{\nu_c} \frac{1}{SNR} \tau^{-1/2}$$

(1)

where $\Delta \nu$ is the clock resonance full-width at half maximum (FWHM), $\nu_c$ is the clock transition frequency, $SNR$ is the signal-to-noise ratio in a 1 Hz bandwidth of the detected signal and $\tau$ is the integration time of the measurement.

Figure 2 reports the typical clock signal for an incident laser power $P_L$ on the cell of 700 $\mu$W.

Figure 2. Clock resonance signal.

The clock resonance is well-fitted by a lorentzian function.

The CPT linewidth $\Delta \nu$ is 564 Hz. The clock signal $S$, defined

$$S = H - y_0$$

on Fig. 2 as $S = H - y_0$, is 0.114 V. The resonance contrast $C_r$, defined as the ratio between the CPT signal $S$ and the dc background $y_0$, is 22%. The discriminator slope $D = S/\Delta \nu$ is measured to be 0.2 mV/Hz.

Figure 3 plots the Allan deviation of the clock frequency. The latter is measured to be $3 \times 10^{-13}$ at 1 s, going down to about $3.2 \times 10^{-14}$ at 100 s. A large bump, from a few seconds to hundreds seconds averaging time, presently prevents the Allan deviation to decrease with a perfect white frequency noise slope. This is currently under investigation but could be explained by thermal effects (the laboratory room is not temperature stabilized), a non-optimized gain of the EOM bias voltage servo loop and feedback from the EOM fiber input face that perturbs the laser frequency. An additional optical isolation stage should be added in a near future. Nevertheless, these short-term frequency stability performances, comparable to those of best vapor cell atomic clocks, are very encouraging for the development of a high-performance Cs vapor cell CPT atomic clock.

Table shows main contributions to the clock short term frequency stability at $\tau = 1$ s. The noise budget is in excellent agreement with the measured clock fractional frequency stability. The clock frequency stability is currently mainly limited by the laser AM noise and the laser FM-AM noise process. The following contribution is the Dick effect. The latter will be reduced later thanks to a novel ultra-low noise frequency synthesis chain [18] that should reject the Dick effect contribution to a level lower than $7 \times 10^{-14}$ at 1 s.

### TABLE I. Main contributions to the clock short term frequency stability at $\tau = 1$ s.

<table>
<thead>
<tr>
<th>Noise Source</th>
<th>$\sigma(1 \text{s})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shot noise</td>
<td>$5.8 \times 10^{-14}$</td>
</tr>
<tr>
<td>Detector noise</td>
<td>$3 \times 10^{-14}$</td>
</tr>
<tr>
<td>LO phase noise</td>
<td>$1.1 \times 10^{-13}$</td>
</tr>
<tr>
<td>Laser AM noise</td>
<td>$2.8 \times 10^{-13}$</td>
</tr>
<tr>
<td>Laser FM-AM noise</td>
<td>$1.7 \times 10^{-13}$</td>
</tr>
<tr>
<td>Total $\sigma_{1\text{s}}$</td>
<td>$3.02 \times 10^{-13}$</td>
</tr>
</tbody>
</table>

The short-term relative frequency stability $\sigma_y(\tau)$ of a passive atomic clock is well-approximated by [17]:

$$\sigma_y(\tau) \sim \frac{\Delta \nu}{\nu_c} \frac{1}{SNR} \tau^{-1/2}$$

(1)

where $\Delta \nu$ is the clock resonance full-width at half maximum (FWHM), $\nu_c$ is the clock transition frequency, $SNR$ is the signal-to-noise ratio in a 1 Hz bandwidth of the detected signal and $\tau$ is the integration time of the measurement.

Figure 2 reports the typical clock signal for an incident laser power $P_L$ on the cell of 700 $\mu$W.

Figure 2. Clock resonance signal.

The clock resonance is well-fitted by a lorentzian function. The CPT linewidth $\Delta \nu$ is 564 Hz. The clock signal $S$, defined on Fig. 2 as $S = H - y_0$, is 0.114 V. The resonance contrast $C_r$, defined as the ratio between the CPT signal $S$ and the dc background $y_0$, is 22%. The discriminator slope $D = S/\Delta \nu$ is measured to be 0.2 mV/Hz.

Figure 3 plots the Allan deviation of the clock frequency. The latter is measured to be $3 \times 10^{-13}$ at 1 s, going down to about $3.2 \times 10^{-14}$ at 100 s. A large bump, from a few seconds to hundreds seconds averaging time, presently prevents the Allan deviation to decrease with a perfect white frequency noise slope. This is currently under investigation but could be explained by thermal effects (the laboratory room is not temperature stabilized), a non-optimized gain of the EOM bias voltage servo loop and feedback from the EOM fiber input face that perturbs the laser frequency. An additional optical isolation stage should be added in a near future. Nevertheless, these short-term frequency stability performances, comparable to those of best vapor cell atomic clocks, are very encouraging for the development of a high-performance Cs vapor cell CPT atomic clock.

Table shows main contributions to the clock short term frequency stability at $\tau = 1$ s. The noise budget is in excellent agreement with the measured clock fractional frequency stability. The clock frequency stability is currently mainly limited by the laser AM noise and the laser FM-AM noise process. The following contribution is the Dick effect. The latter will be reduced later thanks to a novel ultra-low noise frequency synthesis chain [18] that should reject the Dick effect contribution to a level lower than $7 \times 10^{-14}$ at 1 s.

### TABLE I. Main contributions to the clock short term frequency stability at $\tau = 1$ s.

<table>
<thead>
<tr>
<th>Noise Source</th>
<th>$\sigma(1 \text{s})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shot noise</td>
<td>$5.8 \times 10^{-14}$</td>
</tr>
<tr>
<td>Detector noise</td>
<td>$3 \times 10^{-14}$</td>
</tr>
<tr>
<td>LO phase noise</td>
<td>$1.1 \times 10^{-13}$</td>
</tr>
<tr>
<td>Laser AM noise</td>
<td>$2.8 \times 10^{-13}$</td>
</tr>
<tr>
<td>Laser FM-AM noise</td>
<td>$1.7 \times 10^{-13}$</td>
</tr>
<tr>
<td>Total $\sigma_{1\text{s}}$</td>
<td>$3.02 \times 10^{-13}$</td>
</tr>
</tbody>
</table>
IV. CONCLUSION

We implemented a CPT-based Cs vapor cell atomic clock using the push-pull optical pumping technique. The optics part of the clock, compatible with further integration, mainly combines a single diode laser, a Mach-Zehnder electro-optic modulator, an acousto-optic modulator for laser power stabilization and a Michelson delay-line system. An encouraging short-term frequency stability of $3 \times 10^{-13}$ at 1 s was demonstrated, in good agreement with the signal-to-noise ratio limit. Laser intensity effects were found to be the main limitation to the clock short-term frequency stability performances.

ACKNOWLEDGMENT

This work has been funded by the EMRP program (IND55 Mclocks). The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union. This work was partly supported by LNE, LabEx FIRST-TF and ANR-DGA ISIMAC project (ANR-11-ASTR-0004). M. Abdel Hafiz PhD thesis is co-funded by the Région Franche-Comté and the LabEx FIRST-TF (Facilities for Innovation, Research, Services, Training in Time & Frequency).

REFERENCES

Alkali Metal Source Tablet for Vapor Cells of Atomic Magnetometers

Kazuhiro Ban, Akira Terao, and Natsuhiko Mizutani
Frontier Research Center
Canon Inc.
Tokyo, Japan
E-mail: ban.kazuhiro@canon.co.jp

Kazuya Tsujimoto, Yoshikazu Hirai, Tetsuo Kobayashi, and Osamu Tabata
Graduate School of Engineering
Kyoto University
Kyoto, Japan

Abstract—An array system of optically pumped atomic magnetometers (OPAMs) that uses potassium vapor cells requires the quantity of potassium enclosed in the cells to be uniform. We describe a tablet-shaped potassium source, which generates a fixed amount of potassium, for fabricating homogeneous vapor cells. The tablet consists of a tiny microstructured plate covered with a layer of raw materials. Two sets of raw materials were chosen: one was a mixture of KCl and BaN₆, and the other was KN₃. The efficiency of potassium generation depended on the shape and size of the microstructure for each raw material, and a maximum potassium yield of 65% was observed for KN₃ tablets. The source tablets containing KCl and BaN₆ were used to build vapor cells by a microfabrication technique, and the source tablets containing KN₃ were used to introduce potassium into glass cells, which were confirmed to work in highly sensitive OPAMs.

Keywords—Optically pumped atomic magnetometer, alkali metal vapor cell, azide, microstructure

I. INTRODUCTION

Optically pumped atomic magnetometers (OPAMs) that use alkali metal vapor can measure extremely small magnetic fields without cryogenic cooling. They are used for biomagnetic measurements, such as magnetoencephalography, because their sensitivity is comparable to that of superconducting quantum interference devices [1]. We have fabricated an OPAM module that uses potassium [2] and operates under spin-exchange relaxation-free conditions. Potassium was chosen for its long T₂ time, which enables high intrinsic sensitivity [3]. In an OPAM array system, the conditions of the cells, which determine the OPAM sensitivity, must be regulated. The important conditions are the quantity of potassium enclosed in each vapor cell, the buffer gas pressure, and the residual gas and other contamination remaining in the cell. A potassium source that provides fixed amounts of potassium is needed to improve the uniformity of the amount of potassium in the cells.

Break-seal ampoules of alkali metal are used as the alkali metal source for alkali metal vapor cells, which are made of borosilicate glass. A chamber containing the ampoule is connected to a vacuum manifold with a glass tube, and is evacuated and baked for several days to remove the residual gases. The seal is broken after degassing and the alkali metal can be loaded into the cell with a torch. A sufficient amount of alkali metal is available in the source chamber when an ampoule with an appropriate size is used. An alternative low-cost source of the alkali metal is obtained by placing a mixture of alkali metal chloride and calcium grains in the source chamber and baking it to remove residual gasses. The mixture can be heated to the reduction temperature with a torch. The alkali metal vapor is generated and the CaCl₂ is left in the chamber. The alkali metal is moved toward the cell by similar torch work to that used for the ampoule source. This is described in detail in Ref. 4 for glass cells containing cesium. However, the amount of alkali metal in the cell is determined visually, and there is often no way to measure the quantity provided. We have observed that potassium vapor cells fabricated with helium buffer gas show differences in their atom density as large as an order of magnitude.

When a metal azide is used as a source material, the reaction temperature of the source material is decreased substantially. For example, the reduction of KCl with BaN₆ has the lowest reaction temperature of potassium-generating reactions [5]. A low reduction temperature is desirable for use in wafer-scale cell microfabrication processes including lithography, etching, heating, and bonding of the wafers. Liew et al. generated cesium by heating a mixture of CsCl and BaN₆ precipitated on the inner wall of the cavity fabricated by silicon-to-glass anodic bonding [6]. Thermal decomposition of alkali metal azides are also relatively low-temperature reactions. UV-assisted generation of cesium metal from cesium azide has been used to microfabricate cells by precipitating cesium on the inner wall of the cavity [7].

However, these azides are explosive powders with melting points close to their low decomposition temperatures under high vacuum. When a layer of azide in a tiny vessel or on a flat substrate is heated, thermal decomposition of the azide occurs explosively and patches of azide splash the surrounding area, which could contaminate the cell walls through which the optical pump/probe beam passes.
In this study, we focused on the potassium source and examined KN₃ and a KCl/BaN₆ mixture as source materials for potassium generation at different temperatures. It is desirable to have various raw materials available for different cell fabrication processes and different uses of the cells. The source material was precipitated on the surface of a microstructured tablet to form the alkali metal source tablet (AMST). Several types of AMSTs prevented the alkali azide from splashing during decomposition. The dependence of the potassium yield on the material of the microstructure and its size parameters was studied. The AMST containing KCl/BaN₆ was used as potassium source for a microfabricated potassium cell [8], whereas the AMST containing KN₃ was used for making glass cells that contain fixed amounts of potassium for high-sensitivity magnetometers.

II. EXPERIMENTAL METHODS

A. Materials

Two kinds of potassium source that decompose thermally when heated under a high vacuum and generate potassium were compared.

1) KCl/BaN₆:
The reaction of KCl/BaN₆ generates potassium in a two steps reaction. First, BaN₆ decomposes at 120 °C [5]. The Ba then reduces KCl at temperatures above 200 °C generating potassium.

\[
\text{BaN}_6 \rightarrow \text{Ba} + 3 \text{N}_2↑
\]

\[
2 \text{KCl} + \text{Ba} \rightarrow 2 \text{K} + \text{BaCl}_2
\]

KCl and BaN₆ were purchased from Wako Pure Chemical Industries, Ltd. The source material solution was a 1:5 mixture of 3.5 M KCl solution and 0.35 M BaN₆ solution. The solutions were passed through a membrane filter (0.22 μm pore size) before mixing.

2) KN₃:
Thermal decomposition of KN₃ starts at about 350 °C [9].

\[
2 \text{KN}_3 \rightarrow 2 \text{K} + 3 \text{N}_2↑
\]

KN₃ was purchased from Select Lab Chemicals GmbH. A 5.6 M KN₃ solution was prepared and similarly filtered.

B. Microstructures

The microstructured bare tablets were chosen according to the following requirements. They are inert to the reagents, are heat resistant up to about 500 °C, have low outgas on heating, high thermal conductivity, and high electrical resistivity. The high resistivity is necessary because if a piece of highlyconductive material drops into the cell, this could increase the magnetic thermal noise caused by the randomly moving free electrons. Based on these requirements, porous alumina and Si micropillars (Si pillar) were selected.

Porous alumina was purchased from MITEC Co., Ltd. The void fraction was about 30% and the average pore sizes were 2.1–580 μm. Si pillars were fabricated from a Si wafer by photolithography and Deep reactive ion etching (D-RIE). The diameters of the pillars and interstice distance among the pillars were 20–640 μm, and the void fraction was about 30%. Additionally, the Si pillar tablet had an exterior wall to contain the aqueous solution. All bare tablets and Si wafer (flat substrate) were cut into 7 × 7 × 1 mm pieces.

C. Preparation of AMSTs

To remove impurities and improve the hydrophilicity of the surface, the bare tablets were ultrasonically cleaned sequentially with acetone, isopropyl alcohol, and MilliQ water for 5 min each. After drying, the hydrophilicity of the surface of the bare tablets was increased by UV/O₃ treatment.

The source material was precipitated on the bare tablet with a two-step process: aqueous solution (15 μL) was placed on a microstructured bare tablet. The same process was performed for a flat Si substrate control. The water was removed under vacuum (below 10 Pa). This process was repeated six times for KCl/BaN₆ solution, and four times for KN₃ solution.

D. Evaluation of potassium generation from AMSTs

One AMST were placed in the bottom of each of several test tubes attached to a vacuum manifold equipped with a turbo molecular pump (Fig. 1) [8]. One of the test tubes was used to monitor the temperature of the tablet. The temperature was measured with a thermocouple attached with ceramic paste to a flat substrate, which was placed at the bottom of the tube. The upper openings of the tubes were sealed by using a gas torch. The flat bottoms of the test tubes provided good thermal conduction from the tube to the AMSTs. The bottom of every test tube was inserted into a hollow (15 mm deep) in a stainless steel plate on a ceramic heater plate. To increase the thermal conductivity between the stainless steel plate and the glass tube, the gap was filled with alumina powder (ALO16PB, Kojundo Chemical Laboratory Co., Ltd.). A quadrupole mass spectrometer (RGA200, Stanford Research Systems, Inc.) attached to the vacuum system was used to measure changes in the partial pressures of H₂O, N₂, and O₂.

To remove water from the surface of the glass and the AMSTs, the manifold and the tubes were covered with the heating apparatus and ribbon heaters and were baked at 115 °C until the chamber pressure decreased to less than 10⁻⁵ Pa. The
temporal change of the partial pressure of H₂O and O₂ were monitored to ensure that the baking was sufficient. The temperature of the ceramic heater was raised slowly to initiate the thermal decomposition. For KCl/BaN₆, the temperature of the ceramic heater was set at 400 °C for 2 h, whereas the reading of the interior thermocouple was about 280 °C. For KN₃, the temperature of the ceramic heater was set at 500 °C for 2 h, whereas the reading of the interior thermocouple was about 350 °C. Above the reaction temperature, an increase in the chamber pressure and formation of a metal film or droplets on the glass tube wall above the heated portion were visually confirmed. The pressure of N₂ gas was used as an indicator to show that the decomposition of BaN₆ or KN₃ was continuing. When the vacuum pressure fell below 10⁻⁴ Pa, the reaction was regarded to have finished. The glass tubes containing the potassium film and droplets were separated and the potassium in each of the tubes was dissolved in MilliQ water. The potassium concentration in the aqueous solutions was quantified by using inductively coupled plasma-atomic emission spectrometry (CIROS CCD, SPECTRO Analytical Instruments GmbH).

III. RESULTS AND DISCUSSION

To determine the optimal parameters for the microstructure, the effect of the size parameters on the potassium yields for the four types of AMST was examined. The yield was defined as the ratio of the amount of metallic potassium film formed by decomposition to the amount of potassium precipitated on the bare tablet: 2 mg of potassium in 90 μL of KCl/BaN₆ solution, and 13.2 mg of potassium in 60 μL of KN₃ solution, respectively.

A. Potassium generation from KCl/BaN₆ AMSTs

For the thermal decomposition and reduction of the KCl/BaN₆ system, porous alumina AMSTs and Si pillar AMSTs showed similar dependence on the size parameter, namely the pore sizes and pillar diameters (Fig. 2). The potassium yield was much higher on microstructures than on the flat substrate except for the alumina with very small pores. The highest yield of the porous alumina AMSTs was 23.5% on the tablet with an average pore size of 170 μm and was three times higher than the yield of 7.7% on a flat substrate. The highest yield of the Si pillar AMST was 34.4% with a diameter and interstice distance of 320 μm, and was nearly five times higher than the yield on the flat substrate.

To understand the underlying mechanism for these yield curves, the AMSTs were observed by scanning electron microscopy (SEM) and SEM-energy dispersive X-ray spectroscopy (EDS; XL-30, Philips). Figure 3 shows the SEM images for the Si pillar AMSTs. For the AMST with 320-μm-diameter pillars, which produced the highest yield, uniformly precipitated raw materials on the top, side, and bottom surfaces of the pillars were observed indicating no phase separation. In contrast, for the AMST with 20-μm-diameter pillars, which resulted in the lowest yield, there were KCl crystals larger than 100 μm on top of a group of pillars. This indicated that the mixture was phase separated during the drying process and only a small amount of the BaN₆ molecules are near the KCl crystal. The separation may result from the difference in the solubility between the two raw materials and the quick precipitation caused by vacuum evacuation.

Similarly, the formation of KCl crystals is observed on alumina substrates with small pores. To locate where the BaN₆...
was precipitated, cross-sections of the porous alumina AMST and the flat substrate reference were analyzed with SEM-EDS (Fig. 4). For the alumina substrate with 170 μm pores, which resulted in the highest yield for porous alumina, the distribution pattern of the potassium atoms was similar to the pattern of the barium atoms. This indicated that the KCl and BaN₆ were precipitated as a mixture. In contrast, for the substrate with the 2.1 μm pores, which produced the lowest yield, the barium atoms were uniformly distributed but the potassium atoms were concentrated at the top of the microstructure. The non-uniform distribution could decrease the yield. For the flat Si substrate in Fig. 4 (c), the potassium gathers on the surface and some of the barium remains in the potassium layer. The probability that KCl comes into contact with BaN₆ and is reduced is higher than for alumina with 2.1 μm pores. These observations indicate that the potassium yield and the similarity of the KCl and BaN₆ distribution show a strong correlation for the AMST loaded with KCl/BaN₆, suggesting that the yield is limited by the degree of spatial segregation of the two reagents.

The AMST made of 170 μm porous alumina was used for cell microfabrication taking advantage of its low temperature reaction. The cell showed a sharp absorption line and functioned as a magnetometer [8].

B. Potassium generation from KN₃ AMSTs

The yield dependence on the microstructure parameter for AMSTs containing KN₃ is plotted in Fig. 5, and shows substantial differences between the types of substrate. The Si pillar AMSTs had a lower yield than the flat substrate used as the reference. In contrast, all the porous alumina AMSTs had a higher yield than the flat substrate. The highest yield was 65% with an average pore size of 60 μm.

The following type of AMSTs produced a low yield and showed similar behavior: Si pillars, flat Si substrate, alumina with the largest pores, and alumina with the smallest pores. A white precipitate was deposited on the sidewall near the AMST in the tube (Fig. 6 (a)). This might be caused by small particles of KN₃ escaping from the microstructure before decomposition and sticking to places where the reaction temperature was not reached.

In contrast, the AMST with 60 μm pores generated many metal droplets on the sidewall of the tube above the heated area and no white precipitate was observed (Fig. 6 (b)). This suggested that the AMST prevented the scattering of KN₃ and promoted the generation of potassium. The porous alumina AMST surface changed from white to gray after the thermal decomposition. This color change could be caused by some of the generated potassium infiltrating the alumina substrate and could explain the upper bound of the yield.

IV. APPLICATION OF AMSTS IN A MAGNETOMETER

In the previous section, we compared the potassium yields of the two types of AMSTs. There was a large difference in the amount of potassium in the source material solutions used for making AMSTs. For both AMSTs, because it is preferable to have the maximum amount of potassium atoms, the aqueous solutions used for the precipitation were nearly saturated. For the KCl/BaN₆ solution, the low solubility of BaN₆ limits the amount of the KCl that can be used, whereas the high solubility of KN₃ allows a 5.6 M solution to be used. Thus, AMSTs fabricated with KCl/BaN₆ contained 2 mg of potassium, and the AMSTs fabricated with KN₃ contained 13.2 mg.

KCl/BaN₆ may be suitable for supplying sub-milligram quantities of potassium in microfabricated cells because the reaction temperature is low allowing a wider temperature range to be used for the fabrication process. For introducing potassium into glass cells where the volume of the cell is larger and more than a milligram of potassium is needed, KN₃ may be more suitable. Glass vapor cells were fabricated with a 60 μm porous alumina AMST containing KN₃ and their sensitivity in an OPAM was measured.

A. Fabrication of potassium cells

Borosilicate glass cells (Japan Cell Co., Ltd.) were hollow cuboids, and had two chambers: a potassium reservoir (3.5 × 11 × 8 mm), and a measurement chamber (8 × 11 × 8 mm). Potassium cells were fabricated from glass cells attached to each test tube of the vacuum manifold with a straight, narrow tube (Fig. 7). The AMST was covered with niobium foil to heat it with RF wave irradiation. One or two faces of the AMST remained open as the potassium outlet. To remove water from the glass surface and the AMSTs, the vacuum manifold including the tubes was covered with ribbon heaters and baked at 250 °C for 24 h under high vacuum.
The RF heating of the niobium foil induced potassium generation in the AMST. While the measurement chamber of the cell was heated, the potassium on the narrow tube wall was loaded into the reservoir by a gas torch. The manifold was filled with helium and nitrogen (24:1) as a buffer gas and quenching gas, and the total pressure of the buffer gas was 200 kPa at room temperature. Each narrow tube was softened with a gas torch, and the potassium-loaded cells were released from the manifold and permanently sealed.

Twelve potassium cells were fabricated by using AMSTs. Deconvolution analysis was applied to the absorbance spectrum measured around the D1 line to estimate the density of potassium. For the 12 potassium cells, the potassium vapor densities were within 50–70% of the density empirically derived by Killian [10].

B. **OPAM measurements**

A potassium cell in an oven (Fig. 8) was placed at the center of a three-layer-metal magnetic shield with a shielding factor of $10^4$ at 30 Hz [11]. The cell was heated to 180 °C under a flow of hot air to increase the atomic density. Potassium atoms were spin-polarized by a circularly polarized pump laser beam supplied by a laser diode with an external cavity (Tiger, Sacher Lasertechnik) tuned to the center of pressure-broadened D1 line of potassium atoms (770.1 nm). A linearly polarized probe laser beam supplied by another laser diode (Lion, Sacher Lasertechnik) was slightly blue-detuned from the D1 resonance (769.9 nm) and crossed the pump beam at right angles in the middle of the cells.

The polarization rotation of the probe beam was measured as the change in intensity differences between the vertical and horizontal polarization components split by the polarized beam splitter. The change in intensity differences was measured with a balanced amplified detector (2007 Nirvana, New Focus) and stored to a personal computer via an A/D converter. The signal corresponded to the magnetic field perpendicular to both the pump and probe beams. The weak external dc magnetic fields passing through the shields were canceled by three-axes Helmholtz coils set up around the cell. A static magnetic field was applied along the pump beam direction and the resonance frequency of the OPAM was tuned to 10 kHz.

The fabricated potassium cells were put into the magnetometer and good sensitivity was demonstrated. Figure 9 shows a plot of the sensitivity spectrum of the magnetic field at 10 kHz resonance, showing a sensitivity of 3.3 fT rms/Hz$^{1/2}$. Other cells produced a sensitivity of 3.3–3.8 fT rms/Hz$^{1/2}$ at 10 kHz, demonstrating the small variation in sensitivity among these cells.

V. **CONCLUSION**

We proposed AMSTs as alkali metal sources for vapor cells. Two types of potassium source material were used to make the AMSTs: KCl/BaN$_6$ and KN$_3$. We confirmed that AMSTs with suitable microstructural parameters are useful for potassium source, which enables the use of the otherwise explosive thermal decomposition of azides. As a result, conditions for generating potassium efficiently under high vacuum were found. The AMSTs containing KN$_3$ on porous alumina with 60 μm pores were used to fabricate the potassium vapor cells at...
low temperatures. The vapor cells were tested in an OPAM and showed high sensitivity comparable with the sensitivity required for bio-imaging systems. The next step is to fabricate an OPAM array system for biomagnetic measurements and imaging that uses these uniform alkali metal vapor cells as sensor heads.

ACKNOWLEDGMENT

This work was supported, in part, by the Innovative Techno-Hub for Integrated Medical Bio-imaging Project of the Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

A part of this work was conducted at the Kyoto University Nano Technology Hub for the "Nanotechnology Platform Project" sponsored by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

REFERENCES

Alkali Metal Consumption by Discharge Lamps Fabricated from GE-180 Aluminosilicate Glass

C.M. Klimcak, M. Huang, and J. C. Camparo
Physical Sciences Laboratories
The Aerospace Corporation
2310 E. El Segundo Blvd., El Segundo, CA 90245
james.c.camparo@aero.org

Abstract — Alkali rf-discharge lamps provide the light for optical pumping in vapor-cell atomic clocks and magnetometers. Traditionally, the discharge lamp’s envelope has been fabricated from alkali-resistant Corning 1720 aluminosilicate glass, and such lamps have demonstrated decade-long continuous operation. Specifically, the diffusion of alkali atoms into this glass during lamp operation has been shown to be sufficiently slow that lifetimes in excess of ten years can be obtained using only moderate initial alkali fill levels (< 400 μg). However, Corning 1720 glass is no longer being manufactured. It is therefore important to identify alternative glass types that offer comparable alkali resistance, and that are readily available. Although Schott 8436 aluminosilicate glass has been shown to be a suitable substitute, it would be advantageous to identify an alternate glass type produced in higher volume. One alternative, which is manufactured in very high volume for use in automotive lamps, is GE-180 glass. Although this glass also offers high alkali resistivity due to its aluminosilicate composition, the rate of diffusion of alkali atoms into this material when employed as the glass envelope of an rf-discharge lamp has not been measured. We present here our initial results, obtained by Differential Scanning Calorimetry (DSC), of the rate of Rb consumption by discharge lamps manufactured with GE-180 glass. Our results suggest that GE-180 is an excellent substitute for Corning 1720 glass in rf-discharge lamps. We also discuss some unusual issues that have been observed during our DSC measurements, and we outline some new methods for attaining accurate alkali lamp fills using the DSC technique.

I. INTRODUCTION

Low pressure alkali rf-discharge lamps are important components in vapor-cell atomic clocks and alkali magnetometers [1,2]. The lifetime and even the performance of these devices can be governed by the lifetime and operating characteristics of the lamp. Discharge lamp lifetimes were thoroughly investigated a number of years ago, and they were shown to be controlled by diffusion of alkali metal into the lamp’s glass envelope [3,4]. The diffusion process traps the alkali in the glass, rendering it unavailable for generating the alkali emission lines necessary for optical pumping [5]. Consequently, continuous long-term operation of a lamp in either a vapor-cell atomic clock or alkali magnetometer could result in the loss (consumption) of all the free alkali in the lamp, thereby making the lamp useless for optical pumping.

For the specific case of Rb rf-discharge lamps made of Corning 1720 glass, alkali consumption has been shown to follow a time dependent rate law [1,2], where the total amount of Rb consumed after operating a lamp for time t, C(t), is given by

\[ C(t) = A + B \sqrt{t} \]  

where A is an initial (very fast, but usually minor) chemical consumption term that occurs at lamp startup, and B is a diffusive consumption term that lies in the range of 0.1-10 μg/hr^1/2. The diffusive consumption coefficient B should depend primarily on the composition of the glass forming the lamp’s envelope. However, to some extent it will also depend on the specific fabrication processes employed by the lamp manufacturer, and the specific operating conditions of the lamp. It can be shown that B is directly proportional to the diffusion coefficient in Fick’s Law.

Routinely, discharge lamp envelopes have been constructed from Corning 1720 aluminosilicate glass, due to its superior resistance to chemical attack by alkali vapor; and such lamps have demonstrated decade-long continuous operation [6]. Unfortunately, this particular type of glass is no longer available, necessitating the search for alternatives. Schott 8436 aluminosilicate glass has been shown to be an excellent substitute, because its Rb mass consumption coefficient has been determined to be very low, even less than that of Corning 1720 [3]. Another readily available aluminosilicate glass that is produced in high volume for automotive lamp applications is GE-180 glass. Although its aluminosilicate composition suggests that it should offer high resistivity to alkali attack in rf-discharge lamps, there have been no measurements of its A or B consumption coefficients.

We are conducting a two-year experimental program to determine these consumption coefficients for GE-180 glass by operating 12 rf-discharge lamps and periodically measuring their Rb content. Briefly, our approach is identical to that of prior investigators: the mass of free, undiffused metallic Rb in a lamp is measured periodically to determine the mass of Rb in the lamp that has not been consumed during the lamp’s operation. A plot of mass consumed versus lamp operational time then allows us to fit the data to Eq. (1), and determine the A and B coefficients. The mass of free metallic Rb in a lamp is measured using a Differential Scanning Calorimeter (DSC) [7]: the amount of heat required to melt the free metallic Rb in the lamp is measured, \( Q_{\text{fusion}} \), which is equal to the product of

This work was funded by U.S. Air Force Space and Missile Systems Center under Contract No. FA8802-14-0001.
the specific enthalpy of fusion for Rb, \( \Delta H_{\text{fusion}} \), and the mass of free Rb, \( M_{Rb} \), contained in the lamp:

\[
Q_{\text{fusion}} = M_{Rb}\Delta H_{\text{fusion}}.
\]  

In the DSC technique, one heats the lamp’s glass bulb, raising its temperature linearly in time through the Rb melting transition, while monitoring the rate of heat flow into the lamp. If the lamp initially contains solid metallic Rb, then the observed heat flow rate, \( dQ/dt \), will rise when the Rb begins to melt. The flow rate will increase above a baseline flow rate up to a maximum, and will then return to that baseline after all the Rb has undergone the phase change, generating a peak in the observed heat flow rate as a function of time. The integrated area under a plot of calorimetric heat flow versus time, which exceeds the baseline heat flow rate, is \( Q_{\text{fusion}} \). Rubidium that has diffused into the lamp’s envelope is distributed within the glass, and these dispersed atoms do not undergo a phase transition, and consequently do not contribute to \( Q_{\text{fusion}} \). The total amount of Rb that has been consumed by the lamp after operating for a particular time is equal to the difference between the lamp’s initially measured Rb mass and the mass left in the bulb at time \( t \). A series of consumption measurements, taken at several intervals over a two year period, will allow us to determine \( A \) and \( B \) for our GE 180 rf-discharge lamps.

II. EXPERIMENT

A. Overview

We acquired twelve Rb rf-discharge lamps fabricated from GE-180 glass along with standard rf-discharge drivers for the lamps [8]. Rubidium consumption was derived from measurements of \( M_{Rb} \) in the lamps at several times during their operating history. The free Rb mass was measured using a Perkin-Elmer, Diamond Model, power-compensated DSC, whose calorimeter head assembly was modified to accommodate our discharge lamps [9]. The modifications included removal of the plastic, thermally-insulated, spring-loaded calorimeter head cover, as well as the rotating plate sample and reference compartment assembly. These items were replaced with a machined Delrin cap possessing two internal counter-bored chambers for housing lamps, and which have a slightly greater width and height than the lamp dimensions. This Delrin cap was used to cover the lamps after placing them on the calorimeter pans, and it was fastened to the calorimeter baseplate with a threaded rod and wing nut via the center threaded hole that had been used to attach the rotating compartment assembly. A cylindrical stainless steel container open at one end was filled with insulating foam, leaving an internal open volume for the Delrin cap. This is inverted over the cap to create a shroud that provides additional thermal and environmental isolation. In our system, a Styrofoam enclosure is purged with dry nitrogen, and placed over the shroud to prevent condensation of ambient water vapor during measurements.

Very high calorimetric sensitivity is achieved with the power-compensated DSC by directly measuring the difference in heat flow between our sample-under-test (i.e., the lamp that contains the Rb charge) and a reference lamp that is close in total mass to the sample but lacks Rb [9]. The lamp manufacturer potted our lamps into metal bases, for insertion into the lamp exciter electronics. These metal bases were not removed from the lamps during the measurements. Rather, the lamps were inverted in the calorimeter measurement pans, with the metal base of the lamp upward, and the flat window of the lamp downward resting directly on the DSC pan. A diagram of the modified calorimeter head assembly showing the configuration of the sample-under-test and reference lamps within the Delrin cap (but without the inverted cylindrical shroud or purged Styrofoam enclosure) is shown in Fig. 1.

To begin the DSC measurement, it is necessary to condense all the Rb in the lamp to a single contiguous spot on the face of the lamp. This “drive-down” procedure was performed by placing the lamp in an oven which had a small, conically-shaped cold finger touching the face of the lamp. The temperature of the cold finger was maintained well below the temperature of the oven (200 °C) with a refrigerator in order to condense the Rb to a single contiguous spot approximately centered on the lamp’s face. The temperatures of the oven and cold finger were maintained for a twelve hour
period to ensure that all of the Rb had deposited onto the face of the lamp. A successful condensation typically yields a disc of solid Rb approximately 1 to 1.5 mm in diameter and approximately centered on the lamp’s face without any evidence of Rb droplets elsewhere; droplets are revealed by inspection with a 10X hand magnifier. After condensing the Rb, the sample-under-test and reference lamps are placed in the calorimeter pans, and equilibrated at 0°C for at least 30 minutes to ensure that the Rb is in a solid phase at the start of the DSC melting scan. Although Rb melts above room temperature (i.e., 39.3°C), we have routinely observed supercooling of the rubidium [10-12], which necessitates a low temperature soak to solidify the Rb following drive-down.

All Rb mass measurements were corrected for calorimeter aging using a heat flow calibration factor that was acquired by measuring melting and solidification phase transitions of a known mass of indium metal encapsulated in an aluminum pan and provided by Perkin-Elmer. Indium is a reference standard commonly used for the calibration of calorimetric heat flow. Although a major calibration with Indium is performed annually on this instrument using a software-driven calibration procedure embedded within the instrument’s control software, additional indium measurements were performed after each set of Rb lamp measurements to compensate for any calorimeter aging that may have occurred since the instrument’s last annual calibration. The most recent major instrument calibration was performed on 27 January 2014. At that time, the calorimetric heat flow was calibrated to yield an absorbed heat during an indium melting scan of 28.450 J/g, the accepted value of the specific enthalpy of fusion of indium. The specific enthalpy of fusion computed from all subsequent indium measurements, which were acquired with this exact same indium sample, during each lamp measurement interval over an entire year of operation was 28.475 ± 0.055 J/g. Therefore, the average heat flow correction that was applied during 2014 was no more than 0.1% (i.e., 28.475 – 28.450 J/g), which indicates that our calorimeter is an extremely stable instrument even with the calorimeter head modifications. The precision of the indium measurements was also excellent as indicated by the 1-sigma uncertainty, which was less than 0.2%.

B. Melting and Solidification Curves

Both melting and solidification phase transitions were considered for their use in this investigation. Specifically, rather than measuring the heat absorbed during melting, one could also measure the heat liberated during solidification. While either approach can be used to determine the Rb mass, there are advantages and disadvantages to each method. In the following discussion we will point out some qualitative features of the two processes that affect the appearance of the observed DSC phase transition, and the ability to measure the rubidium mass in the lamp.

The melting transition of a pure solid at a constant pressure always occurs at the same temperature, which makes it very easy to set up the temperature range for repetitive DSC (melting) phase-transition scans. One needs only to cool to a sufficiently low temperature between each scan in order to ensure that the Rb has solidified prior to performing a repeat scan. The minimum temperature width of the melting transition, observed as a plot of measured heat flow versus time, is governed by the rate of heat transfer within the Rb ingot, and is independent of the temperature scan rate (assuming that the scan rate is slow enough to minimize excessive instrument distortion). This is the most straightforward case, and is the easiest to experimentally implement. However, it produces a phase transition curve (i.e., dQ/dt vs. t or equivalently dQ/dT vs. T) that has a relatively large temperature width, and therefore the greatest interference from two effects: 1) a background slope in the dQ/dt baseline, which arises from uncompensated differences in heat capacity between the sample-under-test and the reference, and 2) random baseline drift due (for example) to electronic circuitry effects.

A solidification phase transition can be very different. Solidification of a liquid phase occurs only after achieving the growth of a Rb cluster of sufficient size to nucleate crystallization [10]. The subsequent rate of growth of the cluster determines the rate of solidification, and therefore the temperature width of its DSC phase-transition curve. While melting is dependent upon the rate of energy absorption and its dispersion throughout the sample, solidification is dependent upon the rate of energy dissipation. However, if a Rb sample is significantly supercooled[10-12], as is typically observed with Rb solidification in these discharge lamps, then it is already at a temperature where the solid is the lowest energy phase; the fact that the liquid doesn’t solidify, and liberate Qfusion, is due to the lack of a suitably-sized nucleation center. Once a nucleation center forms, it can trigger very rapid solidification of the sample with the rate of solidification now being determined only by the rate at which atoms can physically line up into their correct crystallographic positions, and not on their energy thermal energy conduction rate. Consequently, solidification transitions of supercooled liquids will generally occur over a much narrower temperature range than the corresponding melting transitions, yielding much sharper transitions with considerably improved signal to noise ratio and reduced interference from both background baseline slope and random baseline drift. These significant advantages are gained only at the expense of reducing the temperature scan rate, and increasing the measurement time in order to minimize convolution of the narrow solidification transition with the DSC’s response function.

A complication of the solidification technique is that the growth of the nucleating cluster in a supercooled liquid may not always occur at the exact same “phase-transition” temperature, because solid phase growth may be dependent upon the chance contact with a heterogeneous nucleating surface; and this contact could easily vary from solidification scan to solidification scan. Consequently, the range of temperature over which the solidification commences is typically much larger than the temperature width of the melting phase transition. Once initiated, however, solidification occurs rapidly over a very narrow temperature range. These considerations require one to scan over a very wide temperature range (several times the melting transition width) in order to assure capture of the very narrow solidification transition, while maintaining a scan rate that is slow enough to avoid distortion of the solidification phase-
transition $dQ/dt$ curve due to the finite response time of the DSC. Balancing these constraints can be problematic, particularly if the solidification temperature varies considerably from scan to scan.

Notwithstanding the difficulties of solidification DSC scans, the reduction of DSC baseline interference encouraged us to utilize the solidification approach as an adjunct for determining Rb mass in our lamps. The potential variability of the solidification temperatures is, of course, a concern for a long-term investigation, particularly if the mechanism for nucleation center growth is dependent upon total lamp operating time and/or lamp operation significantly affects the DSC solidification transition curve. We were also concerned that solidification of a supercooled liquid is a non-equilibrium process, which might not be suitable for the present GE-180 lamp-glass investigation. Consequently, we opted to measure both melting and solidification transitions of Rb in our study. The dual measurements assure that accurate mass measurements are performed over the entire duration of the GE-180 glass qualification program, while potentially enabling improved measurements with the use of narrow solidification phase transitions.

Typical phase transitions observed with our lamp bulbs during heating and cooling DSC scans (at a linear temperature scan rate) are shown by the red and blue curves in Fig. 2. In the figure, we have superimposed two sequentially scanned transitions to illustrate the significant DSC phase-transition curve difference between melting and solidification.

![Figure 2: Superimposed melting and solidification phase transitions for Rb showing the dramatic width reduction produced by the solidification of a supercooled liquid. In this example, the melting transition occurred at $T = 39.5 \, ^\circ\text{C}$ and the solidification at $T = 8 \, ^\circ\text{C}$. These transitions have been displaced on the temperature axis and their peak heights adjusted to full scale on this plot for ease of comparison. The absolute value of the integrated area under each peak was in agreement within <1 %.

As shown in Fig. 2, heating produces an endothermic phase transition curve, indicating the heat absorbed by the melting sample. Cooling produces an exothermic transition curve, indicating that heat evolves in the sample’s solidification. The integrated areas under these heat-flow curves correspond to the total absorbed and evolved heats of fusion, $Q_{\text{fusion}}$, from which we determine the Rb mass. The reported uncertainties in our integrated areas correspond to a single standard deviation, which derives from five separate integrated area determination. Although the difference in $Q_{\text{fusion}}$ between melting and solidification for Fig. 2 was very small, we have every so often observed differences as large as 5%, with the larger value of $Q_{\text{fusion}}$ almost always associated with the solidification phase transition. We do not presently understand the cause of these large (but infrequent) discrepancies. However, we do note that the average discrepancy between the melting and solidification values of $Q_{\text{fusion}}$ is only about 1%, which is a value less than our Rb mass measurement uncertainty (for the melting phase transition alone).

C. Data Taking

Since diffusion follows a $t^{1/2}$ dependence, it is important to take data at more closely spaced time intervals during the beginning of the measurement program, when the consumed mass is expected to vary fastest with time. At the start of the program we therefore selected the following data acquisition intervals (in weeks of lamp operating time): 0, 2, 4, 8, 12, 16, 20, 26, 32, 38, 46, 54, 62, 72, 82, 92, and 104 weeks. At the present time we are beginning to acquire measurements after approximately 21 weeks of operation. The one week time slip in our schedule was due to the presence of an additional set of lamp measurements that was inserted in our schedule between the 4th and 5th measurements in this series. As indicated above, we began this measurement program with twelve discharge lamps, but after only a few weeks of operation we found that three of the lamps could no longer have their Rb masses measured; this was due to the presence of a large exothermic baseline offset that preceded the melting transition in these lamps’ DSC scans. This offset was about three times larger than the peak height of the melting transition curve, and it produced a baseline curvature that prevented an accurate determination of the phase-transition $dQ/dT$ area. Similarly, an endothermic offset was observed in these lamps during their solidification DSC scans. The cause of these baseline offsets is unknown, and it is presently under investigation. Occasionally, an offset coincided with the observation of physical motion of the molten pool of rubidium within the lamp envelope, causing the liquid pool to spread over the surface of the lamp during the measurement. Repeated attempts to measure these lamps after re-condensing the rubidium to a contiguous spot following a failed measurement attempt were not successful. Consequently, we removed these lamps from the GE-180 glass investigation, leaving nine discharge lamps in our long-term study. Further details and discussion of this anomalous baseline offset, including speculation about its origin, are provided in the Appendix.

The following procedure was used each time that lamp operation was interrupted for a calorimetric measurement:

1. Rubidium in the lamp was condensed onto the face of the lamp in a contiguous spot.

2. Six melting and six solidification transitions were measured for each lamp. Melting transitions were

...
acquired at a temperature scan rate of 0.25 °C/min and solidification transitions at a rate of 0.1°C/min.

3. Each of these phase transition curves (a total of 12 curves for each lamp) was integrated five different times, using slightly different temporal (i.e., temperature) intervals. This allowed us to compute five mass estimates, and an “integration uncertainty” for each of our phase-transition curves.

4. Rubidium was evaporated from the lamp face and condensed onto the back face of the lamp (i.e., that part of the lamp in the metal base).

5. Steps 1-3 were repeated a 2nd time. Thus, at each point in time, and for each lamp, 120 different estimates of $Q_{\text{fusion}}$ were obtained.

Using these 120 $Q_{\text{fusion}}$ estimates, we computed the average Rb mass in the lamp along with a standard deviation. These masses, however, will not be reported here. Instead, we report the more meaningful A and B coefficients for Rb mass consumption.

After correcting the mass measurements for the calorimeter calibration factor, we determined the consumed mass of Rb. Consumption was computed separately for the melting and the solidification experiments. Our results for the set of nine lamps are shown in the composite set of graphs of Fig. 3. The ordinate in each graph is the consumed Rb mass in micrograms, and the abscissa is the lamp operating time in days. The solid circles correspond to masses derived from melting scans, and the open circles from solidification scans. The average discrepancy between the two sets of consumed masses is about 1%, which was within our measurement error. Consequently, we combined the results from the heating and cooling scans for each lamp. We performed a least square fit of the combined, consumed mass data for each lamp using the expected consumption model given by Eq. (1). This yielded a pair of best fit A and B coefficients for each of the nine lamps, which we present in Table 1, and the consumption curves for these A and B coefficients are shown as the dashed lines in Fig. 3.

Table 1: Best fit A and B coefficients derived from a least square fit of the consumption data for each lamp.

<table>
<thead>
<tr>
<th>Lamp</th>
<th>A (μg)</th>
<th>B (μg/day$^{1/2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.90</td>
<td>0.634</td>
</tr>
<tr>
<td>2</td>
<td>-0.01</td>
<td>0.791</td>
</tr>
<tr>
<td>3</td>
<td>-2.47</td>
<td>0.930</td>
</tr>
<tr>
<td>4</td>
<td>0.97</td>
<td>0.624</td>
</tr>
<tr>
<td>5</td>
<td>-2.49</td>
<td>0.789</td>
</tr>
<tr>
<td>6</td>
<td>2.84</td>
<td>0.431</td>
</tr>
<tr>
<td>7</td>
<td>0.06</td>
<td>0.608</td>
</tr>
<tr>
<td>8</td>
<td>2.32</td>
<td>0.656</td>
</tr>
<tr>
<td>9</td>
<td>-2.15</td>
<td>0.670</td>
</tr>
<tr>
<td>Average</td>
<td>0.107</td>
<td>0.681</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.084</td>
<td>0.141</td>
</tr>
</tbody>
</table>

Figure 3: The consumed Rb mass for each of the nine lamps in this study as a function of the lamp’s operating time. Error bars (± 2 μg) have been provided only for lamp 1 to reduce clutter in the figure. The error bars for the other lamps are the same. The dashed lines are based on least squares fits to the A and B coefficients of Eq. (1) for each lamp, and are listed in Table 1.
Consequently, there is considerable scatter in the consumption data, as evidenced in Figs. 3 and 5. Clearly, the present data cannot be used to confirm the expected square root dependence of Rb consumption on time, and thus confirm the diffusive loss mechanism for GE-180 glass. Data acquisition over a longer term will be required to do so. In the dashed line of Fig. 6 we have extrapolated the projected Rb consumption of a generic lamp out to two years using the diffusive consumption model and our derived family-wide A and B coefficients. We have also plotted our averaged consumption data (solid circles) as well as a linear least square fit of the measured data (dotted line), which represents a simple linear loss rate with time. An error bar has been placed in the figure at 730 days (the termination of our measurement program), which corresponds to 3.3 times our present Rb consumption measurement uncertainty. This extrapolation indicates that by the end of our measurement program we should be able to clearly differentiate between a diffusive consumption model of Rb loss in GE-180 glass (with its square root dependence on time) and a linear consumption rate model – so long as our measurement uncertainty and average loss coefficients remain unchanged over the measurement program’s duration.

In conclusion, our experiments have confirmed that Rb mass consumption measurements for rf-discharge lamps, derived from melting and solidification DSC transitions, are in good agreement with one another up to our measurement uncertainty (i.e., ± 2 μg), even though the liquid phase is supercooled for the solidification DSC scan. To improve the determination of Rb consumption coefficients we averaged the melting and solidification measurements for each lamp, and also averaged the coefficients of all lamps to determine “family wide” GE-180 glass Rb consumption coefficients. These family-wide coefficients and their standard deviations allow us to compute the initial Rb fill level that would be needed for a GE-180 lamp to achieve a specified lifetime with any selected confidence. Specifically, the minimum initial fill required to achieve a certain lifetime is equal to the sum of the average consumption over that lifetime (computed from the consumption law and its family-wide average coefficients) plus an additional amount related to “Rb loss margin.” This
margin is obviously related to the level of reliability for achieving a certain lamp lifetime and Gaussian statistics. As an example, our family-wide consumption coefficients have allowed us to compute initial Rb lamp fills that would be required to achieve lifetimes of 10, 15, and 20 years with a 99.9% confidence; these are listed in Table 2.

**Table 2:** Fill level required to achieve various lifetimes with a 3-sigma or 99.9% confidence interval.

<table>
<thead>
<tr>
<th>Lifetime, yrs.</th>
<th>Initial Rb Mass in GE-180 Glass Lamp, µg</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>73</td>
</tr>
<tr>
<td>15</td>
<td>88</td>
</tr>
<tr>
<td>20</td>
<td>101</td>
</tr>
</tbody>
</table>

**IV. APPENDIX**

In this appendix, we address the anomalous baseline offset that was observed with the three lamps that were removed from our GE-180 lamp-glass qualification study. These lamps began behaving anomalously a few weeks after commencing lamp operation. A typical DSC scan for one of these lamps, in which the Rb sample was first melted and then cooled, is shown in by the blue (lower) plot in Fig. A1. The endothermic melting transition occurs in the midst of an exothermic baseline shift, and the subsequent exothermic solidification transition occurs in the midst of an endothermic baseline shift. The red (upper) curve on this plot displays the programmed temperature during the DSC scan. The melting and solidification transitions lie approximately in the middle of these baseline offsets, and there is a change of slope before and after the transition.

This heating/cooling curve has also been plotted versus the sample-under-test temperature, and is shown in Fig. A2. The heating and cooling curves exhibit hysteresis, which does not fully recover back to the original baseline. Note also that the solidification transition exhibits a moderate amount of supercooling (i.e., the phase-transition temperature for solidification is lower than the phase transition temperature for melting). The observed heat flow offset on melting is exothermic, indicating an apparent step-down in heat capacity with increasing temperature.

**Figure A2:** The heating/cooling curve from Fig. A1 plotted versus sample temperature instead of time. The arrows indicate the time of the scan.

The knee regions of the exothermic and endothermic offsets both precede (in time) their respective 1st order phase transitions suggesting that pre-melting and pre-solidification phenomena are contributing to these offsets. Pre-melting of a solid is a well-known effect that can be caused by the presence of impurities, interfacial curvature, and solid-phase disorder (due to the presence of strain, point defects, dislocations, and polycrystallinity), phenomena that all lower the melting point of a solid [13]. In our case, the most likely cause of pre-melting is the presence of impurities in the sample that can produce a lower melting eutectic mixture. All three of the anomalous lamps were filled from the same batch of Rb on the same fill manifold. If the Rb metal were impure, then it is likely that all three lamps would be “contaminated,” and might behave similarly. Other lamps from this batch have not been tested.

We have often observed significant displacement of the Rb in these anomalous lamps, and have even filmed Rb spreading over the lamp face during melting. The Rb in these lamps does not always spread over the face of the lamp though. In some cases, examination of an anomalous lamp after a DSC scan has revealed the presence of two circular discs of rubidium that were displaced slightly from each other, as if an upper disc of rubidium had slipped over a lower fluid surface. This can be another effect of pre-melting.

At the present time, our working hypothesis is that the anomalous behavior of these lamps is caused by an impurity that forms a lower melting eutectic mixture, which pre-melts below the normal melting point of Rb. The presence of this pre-melted layer during the DSC scan can modify thermal gradients as well as melt flow into the Rb condensed phase to.pdf
produce a baseline offset along with motion of the Rb pool in the lamp. However the possibility of a previously unseen thermodynamic effect caused by the interaction of rubidium with the conditioned surface of the lamp exists. Recent experiments that we have performed with thin rubidium films on the face of these lamps have exhibited similar unexplained heat capacity steps. Clearly, the actual physical mechanism for generating these offsets has not yet been unambiguously identified, and so we are continuing our studies of the anomalous lamp behavior.

REFERENCES

Mercury Lamp Studies in Support of Trapped Ion Frequency Standards

L. Yi, E. A. Burt and R. L. Tjoelker
Jet Propulsion Laboratory
California Institute of Technology, Pasadena, USA
Email: Lin.Yi@jpl.nasa.gov

Abstract— the mercury linear ion trap frequency standard (LITS) [1] at JPL continues to advance with multiple applications. In particular, the outstanding long-term stability [2] and practicality of the ground-based clock have attracted significant interests for time-keeping and metrology. However, the mercury RF discharge lamp used for optical pumping and state detection may limit the ultimate stability performance of the clock [3-4], constraining even broader application. For mercury ion frequency standards, the operational lamp behavior is described by the ratio of useful light at 194nm and unwanted background light at 254nm (194/254). This ratio has been observed to depend on several factors Increasing the 194nm output decreases optical pumping times and an increase of the 194/254 ratio improves the clock signal-to-noise ratio (SNR). These improvements lead to an improvement in clock short-term stability and enable the use of an even broader range of local oscillators.

We have carried out several experiments to unfold the relationship between the 194/254 and the fabrication factors: buffer gas pressure, lamp ID, and quantity of mercury. The quantitative results may be used to improve the process of lamp fabrication for mercury ion frequency standards. The research here may also shed light on other lamp-based applications.

Keywords—atomic frequency standards, RF discharge lamp, deep ultraviolet, vacuum ultraviolet, lamp fabrication, buffer gas, optical pumping

I. INTRODUCTION

The mercury linear ion trap frequency standard (LITS) at JPL [1] have multiple applications. In particular, the outstanding long-term stability and practicality of the ground-based clock LITS9 [2] have attracted significant interest for time-keeping and metrology. However, the mercury RF discharge lamp used in these clocks for optical pumping and detection may limit the performance of the clock [3-4], which consequently constrains broader application.

The operational lamp behavior is described by the ratio between useful vacuum ultraviolet (VUV) light at 194nm and unwanted deep ultraviolet (DUV) light at 254nm (194/254). Previously this ratio has been observed to depend on several fabrication factors: buffer gas (argon) pressure, lamp inner diameter (ID) and quantity of mercury. However, the quantitative relationship is not known well, which constrains the fabrication optimization and repeatability. An increase of the 194/254 would improve the clock signal-to-noise ratio (SNR) and decrease optical pumping time. This leads to an improvement in clock short-term stability and/or enables the use of a local oscillator having lower cost and performance.

We have carried out several experiments examining the 194/254 ratio and the fabrication factors: argon pressure, lamp ID, quantity of mercury. They include:

1. Lamp tubes with different IDs are connected to a vacuum system, where the argon pressure can be tuned. A Mercury discharge is generated in the tubes and light is guided via a piece of deep UV fiber to a grating spectrometer to give the 194/254. Optimal argon pressure is determined in these “open” lamps with different IDs.

2. Sealed (standalone) lamps with different IDs (10 mm, 4mm, 2mm) are fabricated under the same argon pressure (1.0 Torr). The light is sent through free space to a spectrometer and each lamp is tested with a range of operating parameters to achieve an optimal 194/254 ratio.

3. A double-bulb lamp was fabricated providing a means to control the Hg level in the active plasma. The light of the bulb under study is sent via free space to a spectrometer to give the 194/254 ratio.

II. EXPERIMENTS AND RESULTS

A. Argon pressure effect on DUV/VUV spectrum

The deep ultra-violet spectrum control effects are observed while varying the buffer gas (argon) pressure. The mercury vapor (natural abundance) inside the tubes is not saturated. Tests are performed in different tube diameters attached to the ultra-high vacuum system pumped by a turbo molecular pump. The light is guided by a short piece of DUV fiber sealed at the end of the tube. The light coupled out by the fiber is analyzed by a grating-spectrometer. The power of light at 194nm and 254nm is recorded while argon pressure changes from a few hundreds of milli-Torr to a few Torr. A convectron gauge is used to monitor the argon pressure and is calibrated before and after the tests.

From the data, we observe that for both 194nm and 254nm, a DUV radiation peak exists at specific argon pressures. The peaks for these two wavelengths are separated so that there
exists an argon pressure point/range where 194/254 ratio is optimum. This argon pressure point/range depends on several possible experimental parameters: e.g. tube diameter, mercury vapor pressure and electrical excitation configuration. The data in Figure 1 was performed in a glass manifold with the only 1mm (inner diameter) tube attached. Figure 2-4 are performed in a different glass manifold with 5mm, 3mm, 1mm tubes attached.

Figure 1: Observation of the 194 and 254 nm spectrum in a 1 mm diameter tube. 35W RF forward power is recorded. The y axis are in arbitrary unit. The argon pressure is estimated according to gauge calibration data taken before the experiment and assuming a constant pumping speed.

Figure 2: Observation of the 194(red) and 254nm(black) spectrum peaks and ratio(blue) in a 1mm diameter tube driven with about 35W forward RF power.

Figure 3: Observation of 194nm (red) and 254nm(blue) light in a 3 mm diameter tube with about 35W forward RF power.

Figure 4: Observation of 194nm (red) and 254nm (black) light in a 5mm diameter tube. Convectron gauge reading is recorded during pressure changing. Tube diameter is 5mm with about 35W forward RF power. Light 194nm (in red) and 254nm (in black) are recorded.

In a lamp system open to a larger vacuum system, the configuration provides an accurate Argon pressure measurement, although the plasma volume is not confined tightly, especially in the axial direction of the tube. The amount of mercury may vary during the experiment. However, when comparing the 254nm and 194nm under similar conditions for each tube diameter, there exists an optimal
argon pressure with maximized 194nm/254nm ratio favorable for LITS operation. Larger diameter lamps work better at lower argon pressure.

B. Diameter effect on DUV/VUV spectrum

$^{202\text{Hg}}$ enriched lamps with different diameters are built and evaluated. The buffer gas argon pressure is 1.0 Torr. The 254nm/194nm ratio is tested in inductive resonators with different lamp driving parameters (power, frequency). The frequency ranges from about 160MHz to 200MHz, corresponding to the dial reading from 0.0 to 9.9 (x axis of Figure 5 and 7). The forward power ranges from 5W to 25W, corresponding to the dial reading from 0.0 to 9.9 (Figure 6, 8, 9).

For the 10mm inner diameter lamp (Figure 5-6), the minimal 254nm/194nm is about 36.

For the 4mm inner diameter lamp (Figure 7-8), the minimal 254nm/194nm is about 23.

For the 2mm inner diameter lamp (Figure 9), only different power level is tested with the resonant frequency. The lamp driver is modified to deliver more power. The forward power ranges from 15W to 40W, corresponding to the dial reading 0.0-9.9. The minimal 254nm/194nm is about 25.

Within the scope of this experiment, when the 10mm bulb is fabricated together with other smaller diameter bulbs, the 254nm/194nm is not as good as smaller bulbs.

It is known that smaller diameter lamp requires higher electron temperature, i.e. a higher starting voltage and higher power to sustain the discharge. In addition, due to the higher energy of the electrons and ions, the Hg/Hg+ implantation into the fused silica is more probable. In this case, even though the smaller diameter lamp exhibits better 254nm/194nm ratio, the lifetime of the small bulb may be shorter.

Another aspect is the length of the lamp. In these standalone lamps, the length is about 25mm. The current RF resonator extends the field to the entire length so that the confined, full-volume plasma is generated.
C. Hg quantity effect on DUV/VUV spectrum

The amount of Hg involved in the plasma is another important aspect for the DUV/VUV spectrum. A 2-Bulb lamp is fabricated (Figure 10) so that one can control the amount of Hg in the plasma in one bulb by wall temperature control in the other bulb. It is also possible to control and observe argon plasma discharge (Figure 11) rather than Hg discharge.

![Figure 10: 2-Bulb lamp, inner diameter 10mm with partial Hg plasma (blue) generated in one bulb.](image)

![Figure 11: 2-Bulb lamp, inner diameter 10mm with Ar plasma (pink) generated in one bulb.](image)

In the case of argon discharge, the 254/194 is 19, but the absolute 194nm output is about 0.014, which is much smaller than the case in Hg discharge.

Figure 12 and 13 show that there is not much difference as long as Hg plasma (blue) is generated. However, when only argon plasma (pink) is generated, the 254nm/194nm can be very favorable but the absolute power of 194nm is very low. This is an explanation for our previous work [4], the DUV spectrum control in capillary lamps. When argon dominates in the plasma as a result of low quantity of Hg, the Hg ionization process also happens in the bulb and generates 194nm light. Due to the small amount of mercury, the absolute power of the 194nm light is low. Since neutral mercury in the bulb is also low and does not dominate the plasma, the plasma can still keep power and charge balance with dominating Ar/Ar+ and a small portion of Hg/Hg+. In this case the ratio of 254nm/194nm can reach a very low level as in the extreme condition in [4]. In other words, most the mercury atoms available in those capillary tubes are ionized. The same behavior is also observed when a standard 10mm bulb reaches the end of life. This is because that the mercury is depleted by implantation into substrate, and more argon is involved in the ionization process. This breaks the original Hg/Hg+ power and charge balance in the plasma and the 254nm/194nm also becomes very low (<10). In most cases, the lamp does not last very long and eventually reaches the end of life.
As an on-going effort to improve mercury discharge lamp fabrication the quantitative relationship between the light spectrum and the fabrication parameters are studied.

Due to the complexity of the discharge plasma, there are still several unknowns. First, what is the actual argon pressure after the lamp is sealed and how does it evolve over time? How does the lamp vacuum evolve? Non-invasive laser spectroscopy can be used to measure the argon pressure dependent collision shift to the optical transitions of the neutral mercury atoms. Other electromagnetic wave spectroscopy (from radio frequency to optical frequency) can also probe the pressure of background gases. Second, what is the optimal lamp internal diameter when considering both the lifetime and the 254nm/194nm ratio? Third, what is the mathematical model for the plasma when the quantity of Ar, Ar⁺, Hg, and Hg⁺ varies over time? Fourth, what is the mathematical model for plasma when the discharge volume is smaller than 1mm, which is categorized in the micro-plasma regime? These questions are currently under investigations.

ACKNOWLEDGMENT

The authors gratefully acknowledge the experimental support from W. Diener, J. Gonzalez, R. Hamell, A. Kirk, and B. Tucker. The authors also appreciate discussions with J. Prestage, T. Le and T. Bandi. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Copyright 2015 California Institute of Technology. Government sponsorship acknowledged.

REFERENCES


Multimode SiC Trampoline Resonators Manipulate Microspheres to Create Chladni Figures

Hao Jia†*, Hao Tang†*, Philip X.-L. Feng*

Electrical Engineering, Case School of Engineering, Case Western Reserve University, Cleveland, OH 44106, USA
†Equally Contributed Authors.  *Emails: hao.jia2@case.edu; hao.tang5@case.edu; philip.feng@case.edu

Abstract—This digest paper describes the experimental demonstration of two dimensional (2D) microscale ‘Chladni figure’-like patterns [1] of populations of microspheres in liquid using SiC micromechanical resonators. SiC square trampoline resonators (size: 50μm×50μm) exhibit appreciable high frequency multimode resonances when operating in liquid. We are able to manipulate relatively small (1.7μm in diameter) and large (7.75μm in diameter) microspheres by resonant excitation of the trampolines and create a series of 2D microscale Chladni patterns (corresponding to specific mode shapes), from simple to complex. This SiC resonator platform may offer new opportunities for microparticle patterning by taking advantage of its straightforward device fabrication and engineerable multiple modes, and further facilitate cell manipulation, cellular interaction and behavior controlling, and other biophysical and biomedical studies in liquid.

Keywords—Chladni figures; silicon carbide (SiC); resonator; trampoline; multimode; microbead; microsphere

I. INTRODUCTION

In recent years, there are increasing interests and needs in patterning and manipulating micro/nanoparticles and more interestingly, biological cells on device surfaces. Previous techniques are either slow or require pre-patterned structures or substrates to transfer patterns [2-7], such as soft lithography [2], chemical surface patterning [4], dip-pen nanolithography (DPN) [5]. Optical tweezers [7] can manipulate single cells in a ‘non-contact’ fashion with high force and spatial resolution, but time consumption and laser induced heating effects may limit itself in manipulating cells in large quantities. Surface acoustic wave (SAW) methods [8-10] can manipulate cells in a very fast manner with high controllability, but only simple patterns (e.g., dots and straight lines) are usually achieved and the device structures are relatively bulky and often less easy-to-fabricate.

Microscale Chladni figures, first demonstrated using microcantilevers [11], have stimulated the interests of manipulating micro/nanoparticles using mechanical resonances. Resonant microelectromechanical systems (MEMS) with 2D surfaces, can offer larger capturing area and more importantly, diverse 2D resonance modes even in water, which could thus be even more attractive than 1D-microcantilevers. However, to date no 2D Chladni patterns have been demonstrated with MEMS resonators, and no biological particle and cell manipulation, no patterning and sorting have been explored using vibrating flexural MEMS resonators, even though they possess fast responses, high controllability and flexibility in mode engineering.

In this work, we take an initial step toward pursuing the generation and control of 2D microscale Chladni patterns on the top surface of SiC trampoline resonators by leveraging their robust multimode behavior in liquid. SiC material is selected not only for its excellent mechanical (e.g., high elastic modulus, $E_Y$~450GPa), optical (wide bandgap of $E_g$>2.3eV), and thermal properties (thermal conductivity of $\kappa$=320-490 W/[m⋅K]), but also for its high biocompatibility [12, 13], which makes SiC-based resonating platforms well suited and attractive for biological studies in liquid.

Fig. 1: SiC square trampolines with corner tethers. Optical images show (a) a SiC trampoline resonator (50μm×50μm×0.2μm) with tether length of ~4μm and (b) a SiC trampoline resonator (50μm×50μm×0.4μm) with tether length of ~2μm. Scale bar: 20μm. Fabrication process based on SiC-on-Si platform is illustrated in (c), which consists two steps: (i) focused ion beam (FIB) patterning on the pre-deposited SiC layer to define trampoline geometries; (ii) HNA etching to release the trampoline structure.

II. DEVICE FABRICATION

SiC square trampolines with corner tethers have been fabricated. Trampoline structure is selected based on the following considerations: (i) 2D multimode resonances can
be excited (even in liquid) so as to achieve 2D microscale patterns; (ii) 2D surface offers larger spatial area (compared with 1D cantilevers/beams) that plenty of microspheres can be delivered even though device sizes are small; (iii) device surfaces can be flat enough to avoid gravity-induced movement; (iv) device geometries can be highly engineerable so that both relatively simple and more complex Chladni patterns can be possibly obtained.

Two types of trampoline resonators (50µm ×50µm) are fabricated. Ones as shown in Fig. 1a are with a thinner SiC layer ~200nm (red colored) and longer tethers of ~4µm (which can demonstrate multimodes much closer to an ideal trampoline with less coupling from surrounding suspended structures). Ones as shown in Fig. 1b are from a thicker SiC layer ~400nm (green colored) and shorter tethers of ~2µm (which can demonstrate more complex multiple modes strongly coupled with surrounding suspended structures).

The devices are fabricated based on a SiC-on-Si platform as shown in Fig. 1c, which involves a simple two-step fabrication process: (i) Trampoline structure is patterned on a pre-deposited SiC layer using focused ion beam (FIB), with SiC layer milled through, leaving Si substrate exposed for the next-step etching. (ii) HNA etching (modified recipe 10% HF: 70% HNO₃=1:1) is blended to fast etch the silicon substrate, making the trampoline structure fully suspended. After the fabrication process, the whole chip is mounted onto a piezoelectric actuator in a package and totally immersed in DI water. An optical window is employed to seal the aquatic environment and flatten the liquid surface.

A piezoelectric disk actuator excites the flexural-mode resonances more strongly than laser does in liquid so as to efficiently drive the microspheres. We sweep the driving frequency from 5MHz to 100kHz (from high to low, for convenience of introducing the mode shapes), with a step of 100kHz within ~5–1MHz, and a step of 50kHz within 1MHz–100kHz. We take time-lapse images and record videos by using a high-resolution CCD camera during the frequency sweeping process.

**III. EXPERIMENTAL TECHNIQUES**

**A. Experimental Apparatus and Measurement System**

Figure 2 illustrates the schematic of the experimental apparatus and measurement system, consisting of an optical excitation/detection module, a piezoelectric actuation chip module, a time-lapse imaging module, and a microparticle injection module (details will be shown in Section III.B).

We first measure the multimode resonances of the SiC trampoline resonators in DI water using the optical actuation/detection module, in which a 405nm blue laser is employed to thermally excite the device resonances while a 633nm red laser is focused on device surface to detect the flexural-mode vibrations of each mode interferometrically [12, 13]. By implementing this module, we are enabled to: (i) characterize the multimode resonances of such device platform in fluid (determining resonance frequencies and quality factors); (ii) predetermine the piezoelectric driving frequencies for the future experiments.

**B. Micro-Injection Module**

Silica (SiO₂) microspheres (1.7µm and 7.75µm in diameters) are used in this work. For small microspheres (1.7µm in diameter), we implement a custom-built oil pressure-controlled micro-injection module to deliver them onto device surfaces in large amount (as shown in Fig. 3). Microspheres are initially sonicated in DI water and sampled by a pressure-controlled micropipette (tip diameter ~40µm). When delivering microspheres, micropipette is first focused under the microscope objective by adjusting a XYZ stage, then the chip is raised up (by another XYZ
stage) to the same focal plane with target device next to the pipette tip. By adjusting the oil-pressured pump/syringe, the injection speed and distance can be precisely controlled so that microspheres are continuously delivered to the device surface while still movable in the aquatic environment.

C. Optical Tweezers

We have also attempted to deliver much larger (7.75μm, suggesting cell biology applications) SiO₂ microspheres onto the devices to realize manipulation of large microparticles using such resonating platform. However, since such large microspheres cannot be delivered in large amount (considering their weights), we implement an optical tweezer system using the existing He-Ne (633nm) laser in the optical actuation/detection module since the laser is highly-focused under a 50× microscope objective.

As shown in Fig. 4, we are able to optically tweezer individual 7.75μm microspheres onto the SiC trampoline resonators along a tether (~4μm-wide) narrower than the microsphere diameter. Such movement can be precisely controlled by moving the laser spot 1μm per step. We observe that, while also a nice microparticle manipulation technique and precisely controllable, optical tweezers can be rather time consuming, especially when dealing with manipulation of large populations of microparticles.

![Fig. 4: Optical tweezers delivering 7.75μm-diameter microspheres onto the SiC trampoline resonator (shown in Fig. 1a) along one of its narrow tethers. Delivery process is shown in (a), (b), (c), (d). (e) Microsphere patterning using such optical tweezers demonstrating a letter “F” on the SiC surface.](image)

IV. RESULTS AND DISCUSSIONS

A. Calibration of Multimode Resonances in Water

We first characterize the multimode resonances of SiC trampoline resonators (shown in Fig. 1a and b) in water using optical actuation/detection module (Fig. 5). To ensure that the multiple peaks in the frequency spectra are resonance modes, we vary the detecting laser (red spot shown in the insets) to different locations on device (center, side and corner) and off device. The black curves show flat background spectra based on which we are able to distinguish at least 5 resonance modes within 100kHz–5MHz. Such robust high-frequency multimode resonances of SiC trampoline resonators in water helps to predetermine the piezoelectric driving frequencies for the next steps and facilitate the efficient manipulation of microspheres into multiple 2D Chladni-like patterns.

![Fig. 5: Multimode resonance characteristics of the SiC square trampoline resonators (Fig. 1a and b) in water. At least 5 flexural modes within 100kHz–5MHz are distinguishable when the detecting (red) laser is focused at device center (red curve), side (blue curve) and corner (green curve). Black curves show a flat background when focusing off the device.](image)

![Fig. 6: Manipulating a number of 1.7μm-diameter microspheres using the SiC square trampoline resonator shown in Fig. 1a, creating a series of simple and complex 2D microscale Chladni figures or patterns. Microspheres are observed to stabilize on the nodes of the simulated mode shapes listed on the left. We are able to resolve 4 flexural modes of the SiC trampoline resonator at 4.5-4.6MHz, 2.3-2.4MHz, 0.6-0.7MHz and 0.3-0.4MHz, which agree with the resonance spectrum in Fig. 5a.](image)
B. Manipulating Small Microspheres into Chladni Patterns

We first deliver a number of 1.7μm microspheres onto the SiC trampoline resonator (in Fig. 1a) using the microinjection module. To efficiently drive the microspheres to resolve 2D microscale Chladni patterns in water, a piezoelectric actuator is excited from 5MHz to 100kHz, with a step of 100kHz within 5–1MHz, and a step of 50kHz within 1MHz–100kHz. The real-time distributions of the microspheres are captured in high-resolution imaging by using the time-lapse imaging module. In this case, images are taken under 10× objective because 1.7μm microspheres are much more distinguishable compared with using 50× objective (color difference may exist). In this case, images are taken under 10× objective because 1.7μm microspheres are much more distinguishable compared with using 50× objective (color difference may exist). We are able to manipulate these 1.7μm microspheres using the SiC trampoline resonator. The time-lapse images indicate that the movement of the microspheres follows individual Chladni patterns (towards the nodes of the mode shapes) when the driving frequency is continuously swept from 5MHz to 100kHz. We also observe that some of these results agree with previous observations in Fig. 6, such as the higher mode at ~4.6MHz, cross “×” mode at ~2.3MHz and the fundamental mode at ~0.3MHz. These results prove that such microparticle manipulation techniques may be versatile for particles with micrometer diameters and possess potentials for cell biology and biomedical studies. The microspheres are observed to quickly stabilized on the nodes (blue colored area) of the mode shapes (in the case of the first mode, microspheres are almost removed from the surface in response to the vibrations of whole device area). Both complex and simple patterns (such as cross “×” and line “/”) are observed in a fast manner within 10s. To our knowledge, it is the first time that 2D microscale Chladni patterns are observed at high frequency and in real-time using such resonating platform.

C. Manipulating Large Microspheres into Chladni Patterns

We also attempt to manipulate large microspheres using such SiC trampoline resonator, suggesting relevant potential biological applications, such as cell manipulation and patterning and cellular interaction and behavior controlling. By using the optical tweezers, we are able to successfully deliver four 7.75μm-diameter microspheres onto the device (also as in Fig. 1a) along a narrow tether. In this case, time-lapse images are taken under 50× objective since the 7.75μm microspheres are visible enough, and the image quality can also benefit from higher magnification. As shown in Fig. 7, we are able to manipulate these 7.75μm microspheres using SiC trampoline resonator. The mode shapes are observed in a fast manner within 10s. To our knowledge, it is the first time that 2D microscale Chladni patterns are observed at high frequency and in real-time using such resonating platform.

D. Chladni Patterns Showing Complex Coupled Modes

We further extend the experiment with the other SiC trampoline resonator (shown in Fig. 1b), which is engineered to be able to demonstrate a series of coupled modes (resonance modes of the whole suspended membrane are presented instead of the trampoline area only). As shown in Fig. 8, we are able to observe at least 6 2D microscale Chladni patterns corresponding to 6 coupled modes at high frequency when the driving frequency is swept from 5MHz to 100kHz, which also agree well with the resonance spectrum in Fig. 5b. Mode shapes are simulated and best corresponded to each pattern. Such results are very interesting because: (i) design of such resonating platform requires awareness of coupled modes (tether geometries, or the etching window should be optimized for both multimode characteristics and microsphere diameters); (ii) 2D Chladni patterns of mechanically coupled modes are also useful for microparticle manipulation and patterning.

However, instead of stabilizing along the nodes like in Fig. 6 and Fig. 7, microspheres are observed stabilized along the antinodes of the mode shapes this time. Possible
exotic device shapes can be envisioned to explore various biophysical and biomedical studies at chip scale. Since controlling cellular interaction and behavior, and other types of resonance modes, and further stimulate explorations of simple device fabrication and engineerable multiple MEMS resonating platform may offer new opportunities for mode shapes based on Chladni patterns. Such versatile SiC microspheres and resolve both simple and complex coupled modes. We are able to manipulate relatively small (1.7 μm in diameter) and relatively large (7.75 μm) microspheres tend to be influenced by the top layer of the inner vortices, which deserve further analysis in the future.


We acknowledge the support from the National Science Foundation through grant ECCS-1408494, the Case School of Engineering, the Swagelok Center for Surface Analysis of Materials (SCSAM), and the Materials for Opto/Electronics Research & Education (MORE) Center at Case Western Reserve University (CWRU), and a CSC Fellowship (No. 201306250042).

ACKNOWLEDGMENT

We acknowledge the support from the National Science Foundation through grant ECCS-1408494, the Case School of Engineering, the Swagelok Center for Surface Analysis of Materials (SCSAM), and the Materials for Opto/Electronics Research & Education (MORE) Center at Case Western Reserve University (CWRU), and a CSC Fellowship (No. 201306250042).

REFERENCES

Calibrating Temperature Coefficient of Frequency (TC\textsubscript{f}) and Thermal Expansion Coefficient (\(\alpha\)) of MoS\textsubscript{2} Nanomechanical Resonators

Rui Yang*, Zenghui Wang, Philip X.-L. Feng*
Department of Electrical Engineering and Computer Science, Case School of Engineering
Case Western Reserve University, Cleveland, OH 44106, USA
*Email: rui.yang@case.edu; philip.feng@case.edu

Abstract—We report on experimental study and calibration of temperature coefficient of frequency (TC\textsubscript{f}) and laser heating effect in two-dimensional (2D) resonators, and through which, a new method for determining thermal expansion coefficient (\(\alpha\)) in 2D crystals. We measure the resonance characteristics of the ‘drumhead’ nanomechanical resonators based on molybdenum disulfide (MoS\textsubscript{2}) using sensitive laser interferometry. By tracking the resonance frequency at varying temperatures, we obtain high TC\textsubscript{f} up to -0.396%/K. Combining with finite element modeling (FEM), we extract \(\alpha\) of 3.7\textpm 4ppm/K. We also measure the laser power dependence of the resonance frequency (up to -4.116%/mW), demonstrating the potential of this type of devices for sensitive thermometers and laser power meters.

Keywords—Resonator; TC\textsubscript{f}; Thermal Expansion; MoS\textsubscript{2} 2D Nanoelectromechanical Systems (NEMS); Optical Interferometry

I. INTRODUCTION

Nanoelectromechanical resonators based on molybdenum disulfide (MoS\textsubscript{2}) two-dimensional (2D) crystal exhibits attractive properties such as high frequencies and potential for frequency scaling due to desirable mechanical properties and unconventionally high strain limits [1,2]. These MoS\textsubscript{2} resonators are interesting for a number of sensing applications due to their low mass and large surface area, and pressure sensing capability of these devices have been explored [3]. Ultrathin MoS\textsubscript{2} resonators also present high potential for frequency-shift-based temperature sensing applications because the resonance frequency shifts with the thermally induced strain. Further, as today’s MoS\textsubscript{2} resonators’ vibrations are mostly transduced via optical detection schemes, and because laser heating affects the resonance characteristics of micro/nanoelectromechanical systems (MEMS/NEMS) [4], it is necessary to calibrate the laser heating effect on MoS\textsubscript{2} resonators. Here we show measurement of the temperature coefficient of frequency (TC\textsubscript{f}) through which we extract thermal expansion coefficient (\(\alpha\)) of the MoS\textsubscript{2} device, and calibration of the resonance frequency (\(f\textsubscript{res}\)) shift under different laser power levels. This type of device is interesting for temperature sensing as previously realized in MEMS devices [5], as well as photodetectors that have been demonstrated using MoS\textsubscript{2} transistors [6,7] and heterojunction p-n diodes formed by MoS\textsubscript{2} and other 2D materials [8]. As we can measure frequency and time with exceptionally high accuracy, and our device has high TC\textsubscript{f} and high laser power dependence, this technique could lead to the development of ultrasensitive temperature sensors and laser power meters.

II. DEVICE AND MEASUREMENT TECHNIQUES

We pattern circular microtrenches on SiO\textsubscript{2} (290nm)/Si wafer using photolithography and buffered oxide etching (BOE). Then MoS\textsubscript{2} crystals are mechanically exfoliated onto the microtrenches, and the candidate devices are identified under an optical microscope. Suspended MoS\textsubscript{2} crystal flakes that either fully or partially cover the microtrenches could serve as candidates for resonance measurements. The MoS\textsubscript{2} flake thickness is first estimated from optical contrast and then measured with atomic force microscope (AFM).

Fig. 1. FEM simulation of the MoS\textsubscript{2} resonators with the MoS\textsubscript{2} crystal (a) & (b) fully-covering, and (c) & (d) partially-covering the microtrenches. (a) & (c) Mode shapes of the fundamental resonance mode for respective devices. (b) & (d) Temperature profiles of the devices under 633nm laser heating.

Resonance mode shapes of the devices simulated by finite element modeling (FEM, using COMSOL) are shown in Fig. 1a (for fully-covered device) and Fig. 1c (for partially-covered device), with the dependence of the resonance frequency on temperature considered in the model. The laser heating effect is also simulated by assuming a Gaussian laser beam with spot size of 1\textmu m focused on the suspended MoS\textsubscript{2}, inducing temperature change. Fig. 1b and Fig. 1d show the temperature profiles of the devices under 1mW laser power, showing maximum temperature rise of \(-28\text{–}30°C\).
The resonance characteristics of the MoS$_2$ resonators are measured with a specially-engineered optical interferometry system using 633nm He-Ne red laser with spot size of ~1μm. The undriven thermomechanical motion modulates the reflected light intensity, which can be recorded with a photodetector connected to a spectrum analyzer, and details of the measurement system can be found in previous work [9,10]. We also measure the driven resonances by opto-thermally exciting the device motion with an amplitude-modulated 405nm blue diode laser close to the device. A radio-frequency (RF) network analyzer sweeps the driving frequency and modulates the laser intensity, as well as detects the resonance signal [9,11]. Both thermomechanical and driven resonances could be used to monitor the resonance frequency under different temperature and laser power levels, while we mostly use driven resonance in this work, mainly because thermomechanical resonance requires long average time, and at certain temperature and small laser power, thermomechanical resonance becomes hardly visible. All measurements are performed in moderate vacuum (~10$^{-4}$–40mTorr).

### III. RESULTS AND DISCUSSIONS

#### A. TC$_f$ Measurement and Extraction of $\alpha$

We measure the resonances of the devices at varying temperatures to determine the TC$_f$, and we limit the temperature range to be within ±10°C to minimize the effect of surface adsorption [12] and other effects due to excessive heating/cooling. The temperature of the device is controlled with a Peltier heater/cooler, and the temperature ($T$) reading is performed with a diode temperature sensor [13]. Fig. 2a and 2c show the measured resonances of a fully-covered multilayer (56nm thick) MoS$_2$ resonator at 20.7°C and 42.4°C, respectively, where $f_{res}$ clearly decreases at higher temperature. The resonances are fitted with a driven damped harmonic resonator model [1], allowing us to determine $f_{res}$ and quality ($Q$) factor. In the experiment, we vary the temperature with small steps, and sweep the temperature both up and down to check if there is any hysteresis, which has been observed in the measurement with large temperature range [12]. By summarizing the relative $f_{res}$ shifts at different temperatures (using the 20.7°C data as reference), we observe linear TC$_f$ with no hysteresis, and extract TC$_f$ of -0.396%/K (Fig. 2b). The very large TC$_f$ makes the MoS$_2$ resonators promising for highly responsive temperature sensors.

With the measured TC$_f$, we further perform FEM to simulate the temperature effect on resonance frequency and extract $\alpha$ through comparison of the experimental data with simulation results. The fundamental-mode $f_{res}$ of the drumhead resonator can be determined by [1]:

$$f_{res} = \frac{k d}{\pi} \sqrt{\frac{D}{\rho d^4}} \left[ \frac{(kd)^2}{2} + \frac{\gamma d^2}{4D} \right]$$  \hspace{1cm} (1)

where $\rho$ is areal mass density, $k$ is a numerically-determined modal parameter, $d$ is the diameter of the microtrench, $\gamma$ is the tension in the device (in N/m), and $D$ is the flexural rigidity of the suspended MoS$_2$ device, which could be expressed as:

$$D = \frac{E t^3}{12 (1-\nu^2)}$$  \hspace{1cm} (2)

where $E$ is the Young’s modulus, $t$ is the MoS$_2$ thickness, and $\nu$ is the Poisson ratio.

When the temperature of the substrate changes, due to the difference in $\alpha$ between SiO$_2$ and MoS$_2$, the strain in the suspended MoS$_2$ changes, which shifts the resonance frequency according to (1). In FEM simulations, we vary the thermal expansion coefficient and plot the resonance frequency shift at different temperatures (Fig. 2b). The simulated resonance frequency shifts are compared with the experimental data, and the thermal expansion coefficient that makes the simulated resonance frequency shift closest to the experiment is determined as the $\alpha$ of the device, which is 3.7ppm/K for the fully-covered device. It is worth noting that in the model, we only consider the thermal expansion effect on resonance frequency, and neglect the effect of Young’s modulus change with temperature. This is based on the assumption that the relative change in Young’s modulus with temperature should be small because the initial $E_y$ is high and the temperature range is small; while we assume the initial tension in the device is small, thus the thermally-induced tension by stage heating will result in more significant change in the total device
tension. Another assumption we make is that the clamping between MoS2 and SiO2 is ideal without any sliding, which should also hold because the small temperature range we use in the experiment should not induce excessive deformation.

The temperature-dependent resonances are also measured for a partially-covered MoS2 resonator (Fig. 2f inset). Fig. 2d and 2f show the resonances at 31.9°C and 39.1°C, respectively, and we fit these curves to a Lorentzian function with a phase shift. We observe that the \( Q \) is only ~10, which is much lower than the fully-covered device. This could possibly be attributed to the different damping mechanisms or surface adsorbates [3]. This device shows TC of \(-0.209\%/K\), which is smaller than the fully-covered device. From the summarized frequency shift data and FEM simulation, we also extract \( \alpha \) in this device to be 4.0ppm/K, which is quite close to the value obtained for the fully-covered device, with a difference of only 7.5%. This proves that monitoring resonance shift is a reliable method to determine \( \alpha \) of 2D resonators.

**B. Laser Power Dependence of MoS2 Resonators**

We also measure the dependence of \( f_{\text{res}} \) on 633 nm detection laser power. When the laser is focused on the device, the reflected laser power is modulated by the device motion and can thus be used for resonance detection, while the absorbed laser power could increase the device temperature. It is different from stage heating which heats up the substrate, and then both the MoS2 device and the substrate will reach the same temperature given sufficient time for heat transfer. This is because of the following. First, the laser is a Gaussian beam which has a finite spot size, thus the absorbed heat is different at varying positions relative to the laser. Second, laser heating first heats up the suspended MoS2 and then the heat is transferred to the substrate; limited by the thermal conductivity, there is a temperature gradient. The thermal expansion is then different at different positions of the MoS2 and the substrate. We simulate the laser heating effect with FEM for a 56nm-thick MoS2 suspended on a 5um microtrench. We have previously investigated simulation of Joule heating of MoS2 transistors with FEM [14], while for suspended MoS2 resonators, the simulation is changed to laser heating effect with different heat conduction. The temperature profiles for fully and partially covered devices are shown in Fig. 1b and 1d, presenting that the highest temperature is usually the center of the laser spot, and the temperature lowers toward the substrate.

In the experiment, we start from small 633nm laser power of 350\( \mu \)W, which still results in observable driven resonance, as shown in Fig. 3a and Fig. 3d, for the same resonators shown in Fig. 2b and 2e inset, respectively. Then we gradually increase the laser power and monitor the resonances (Fig. 3c and 3f). We observe that \( f_{\text{res}} \) decreases as we increase the laser power. Similar to TC of \( f_{\text{res}} \) shift with laser power also exhibits linear relationship, as shown in Fig. 3b and 3e, demonstrating \( f_{\text{res}} \) shift with 633nm laser power of \(-4.116\%/mW\) and \(-1.529\%/mW\), respectively. We also limit the laser power to be smaller than 1.05mW on device in order to prevent the high power density of the focused laser from damaging the device. The device resonance frequency is quite responsive to laser power, demonstrating their potential for applications as ultrasensitive laser power meters.
0.35mW) with 3c (Q≈95.5 at 1.05mW), and comparing Fig. 3d (Q≈26 at 0.35mW) with 3f (Q≈11 at 1.05mW). This is also consistent with the effect that lower stage temperature leads to higher Q, by comparing Fig. 2a (Q≈528 at 20.7°C) and 2c (Q≈272 at 42.4°C), though for the partially-covered device changing temperature does not change Q much. To closely examine the effect of temperature and laser power on Q, we summarize the Q under different conditions in Fig. 4, with Fig. 4a and 4b showing the fully-covered device in Fig. 2a-2c, and Fig. 4c and 4d showing the partially-covered device in Fig. 2d-2f. Fig. 4 further proves that Q is in general decreasing with both higher temperature and higher laser power, which will require further study to determine its origin.

C. Calibration of the Effect from Different Laser

We further calibrate the effect from both 633nm red laser and 405nm blue laser on the resonance frequency shift of the MoS2 resonators, using both undriven thermomechanical resonances and driven resonances (Fig. 5). We measure the driven resonance frequency shift at different red and blue laser power, thermomechanical resonances with blue laser driving off resonance, and resonances with laser on different positions on device. We find that the 405nm laser has very little heating effect (green curve) compared to the 633nm laser, since we defocus it and move it away from the device to prevent overheating. The $f_{res}$ dependence on the 633nm laser power is similar for different blue laser driving power, and for both driven and thermomechanical resonances, showing that using driven resonances to measure TC is relatively accurate. The red laser position will affect the trend of $f_{res}$ shift with red laser power, because the laser heating depends on position.

IV. CONCLUSIONS

In summary, we have measured the TC/f of MoS2 NEMS resonators suspended on circular microtrenches, and calibrated the resonance frequency shift with different laser power. The resonances are measured with laser interferometry as the stage temperature or laser power is changed. Thermal expansion coefficient is extracted through comparing the experimental data with FEM. These devices show very high TC/f values, and clear, strong laser power dependence, making them promising for ultrasensitive thermometers and laser power meters.

ACKNOWLEDGMENT

We thank J. Lee, K. He, J. Shan for helpful discussions and technical support. We thank support from Case School of Engineering, National Academy of Engineering (NAE) Grainger Foundation Frontier of Engineering (FOE) Award (FOE 2013-005), the CWRU Provost’s ACES+ Advance Opportunity Award, the National Science Foundation CAREER Award (ECCS-1454570), and the CSC Fellowship (No. 2011625071). Part of the device fabrication was performed at the Cornell NanoScale Science and Technology Facility (CNF), a member of the National Nanotechnology Infrastructure Network (NNIN), supported by the National Science Foundation (Grant ECCS-0335765).

REFERENCES

Performance Evaluation of CMOS-MEMS
Thermal-Piezoresistive Resonators
in Ambient Pressure for Sensor Applications

Jung-Hao Chang1, Cheng-Syun Li2, Cheng-Chi Chen2, and Sheng-Shian Li1,2
1 Dept. of Power Mechanical Engineering, 2 Inst. of NanoEngineering and MicroSystems
National Tsing Hua University, Hsinchu, Taiwan
E-mail: g123869443@hotmail.com

Abstract—In this work, we report a thermally driven and piezoresistively sensed (a.k.a. thermal-piezoresistive) CMOS-MEMS resonator with high quality factor in ambient pressure and with decent power handling capability. The combination of (i) no need of tiny capacitive transducer’s gap spacing thanks to thermal-piezoresistive transduction, (ii) the use of high-Q SiO2/polysilicon structural materials from CMOS back-end-of-line (BEOL), and (iii) the bulk-mode resonator design leads to resonator Q more than 2,000 in ambient pressure and 10,000 in vacuum. Key to attaining sheer Q in ambient pressure relies on significant attenuation of the air damping effect through thermal-piezoresistive transduction as compared to conventional capacitive resonators which necessitate tiny transducer’s gap for reasonable electromechanical coupling. With such high Q and inherent circuit integration capability, the proposed CMOS-MEMS thermal-piezoresistive resonators can potentially be implemented as high sensitivity mass/gas sensors based on resonant transducers. The resonators with center frequency around 5.1 MHz were fabricated using a standard 0.35 μm 2-poly-4-metal (2P4M) CMOS process, thus featuring low cost, batch production, fast turnaround time, easy prototyping, and MEMS/IC integration.

Keywords: CMOS-MEMS, Micro-resonators, Thermal actuation, Piezoresistive sensing, High Q, Micro-electro-mechanical Systems, Resonant transducers, Mass sensors

I. INTRODUCTION

In the past decade, air pollution and environmental particulates are identified as the main triggers of lung cancers due to modern industrialization, especially the particulate matter with the diameter less than 2.5 μm (i.e., PM 2.5). Such small particles would easily penetrate into the human lungs, and thus the lung cancer rate fast grows worldwide in recent years. As a result, sensors capable of detecting the particulate matter are of great importance to monitor the air quality. Since the conventional capacitive MEMS resonators not only alleviate the air damping effect due to no tiny gap spacing but also consist of CMOS-compatible structural materials, such as polysilicon and silicon dioxide, to benefit inherent MEMS/IC integration. On the other hand, although the piezoelectric resonators possess excellent electromechanical coupling and relatively higher Q in air, the issue of CMOS compatibility still places a bottleneck for monolithic integration with circuits. In contrast, the recent development of the thermal-piezoresistive resonators only alleviate the air damping effect due to no tiny gap spacing but also consist of CMOS-compatible structural materials, such as polysilicon and silicon dioxide, to benefit inherent MEMS/IC integration. Furthermore, the size scale-down of the resonators has less impact on the thermal-piezoresistive transducers as compared to other transduction mechanisms which require considerable transducer areas. Therefore, the thermal-piezoresistive transduction becomes a simple and effective mechanism to couple the energy between mechanical and electrical domains.

The prior thermal-piezoresistive resonators are often made by silicon carbide or single crystal silicon to attain high Q while the heater for driving and the piezoresistor for sensing are also formed using the same structural materials via doping 0[3][4], which substantially limits their design flexibility as well as integration with CMOS circuitry. Thanks to the recent...
advancement on the CMOS-MEMS technology through which MEMS resonators can be monolithically integrated with amplifier circuits to realize a single-chip configuration [5]. Furthermore, the thermal-piezoresistive resonators could also be realized through the CMOS-MEMS technology [6]; unfortunately they suffer low $Q$ (only 589) due to the post-CMOS process issues.

To address the low-$Q$ issue, we take advantage of the previously developed CMOS-MEMS resonators with capacitive driving and piezoresistive sensing [7], where a record-high $Q$ among CMOS-MEMS resonators is reported. By means of optimization of the BEOL structural materials, the measured $Q$ of the CMOS-MEMS resonator is greater than 15,000 in vacuum [7]; however, the resonance vanishes in ambient pressure owing to the air damping effect caused by its tiny capacitive gaps. To greatly alleviate damping in air, the resonator is re-designed in this work to remove the small gap spacing, while instead adopting thermal driving and piezoresistive sensing for operation in ambient pressure.

In this paper, we propose a thermal-piezoresistive resonator fabricated by a standard 0.35 μm 2P4M CMOS-MEMS platform, successfully demonstrating $Q$ up to 2,000 in air, robust transduction, and decent power handling capability. With such features, the resonators developed in this work can potentially be implemented as mass sensors capable of detecting particulate matters.

II. DEVICE DESIGN AND OPERATION

Fig. 1(a) presents the SEM photo of the proposed CMOS-MEMS thermal-piezoresistive resonator, including (a) standalone device SEM view, and (b) FIB cut cross-sectional view of the thermal beam.

MEMS thermal-piezoresistive resonator. To attain high $Q$, low-loss structural material from the dielectric (i.e., SiO$_2$) layers of CMOS BEOL is utilized, which also provides electrical insulation for the embedded heater and sensor with flexible routing capability. In addition, we adopt the bulk-mode resonator design which features larger mechanical stiffness, higher quality factor, and better power handling capability as compared to its flexural-mode version. Compared with the capacitively-driven and piezoresistively-sensed resonator developed in [7] (cf. Fig. 1(b)), the proposed thermal-piezoresistive resonator without tiny gap spacing (cf. Fig. 1(a)) greatly mitigates the air damping effect. To eliminate the feedthrough level in the thermal-piezoresistive resonators, two polysilicon layers (i.e., 2P) are used to serve as the heater (Poly-1) and sensor (Poly-2), respectively, thus enabling a decoupling scheme of heating and sensing elements, as shown in Fig. 1(c), to prevent the resistive feedthrough. Based on the material properties of the two polysilicon layers, Poly-1 layer possesses low resistance and high current density capability, which is suitable for heater operation. In contrast, Poly-2 features high gauge factor (or piezoresistive coefficient), being a good candidate to act as piezoresistors. To operate the resonator in Fig. 3, one dc current ($I_D$) flows through Poly-1 heater to generate biasing for the thermal drive while another dc current ($I_S$) is applied into Poly-2 sensor to provide the resistive readout. Such a measurement scheme enables single-ended, single-to-differential, differential-to-signal, and fully differential test configurations, thus allowing us to explore the optimized operation of the proposed device.

III. EXPERIMENTAL RESULTS

The resonator was fabricated by a TSMC 0.35 μm 2P4M process, followed by a maskless metal wet etching process [7]. The SEM views of the thermal-piezoresistive resonator are shown in Fig. 2 where (a) presents the fabricated resonator by the use of the CMOS-MEMS platform while (b) shows a focus-ion-beam (FIB) cut cross-sectional view on the thermal actuator beam in which the Poly-1 heater and Poly-2 piezoresistor are clearly seen. The resonator is placed in
ambient pressure with the measurement setup depicted in Fig. 3 under a fully differential scheme. Fig. 4 presents the measured frequency characteristics in air and vacuum environments using the differentially thermal drive and differentially piezoresistive sense. $Q$-factor in vacuum is six times higher than that in air while $Q$ of 2,000 in ambient pressure is still very attractive among all CMOS-MEMS counterparts. The transmission loss between vacuum and air (around 10 dB) can be compensated using higher de-power operation. Furthermore, the power handling of the resonator is also characterized as shown in Fig. 5, indicating the resonator possesses a decent power handling capability in air. Compared with the flexural-mode capacitive resonators where the input power of -20 dBm introduces strong Duffing effect due to capacitive nonlinearity, the proposed thermal-piezoresistive resonator sustains +6 dBm input power without significant nonlinear effect. Finally, Table 1 provides comparison of the state-of-the-art CMOS-MEMS resonators with various transductions.

IV. CONCLUSIONS

This work presents a thermally actuated and piezoresistively sensed CMOS-MEMS resonator with $Q\sim2,000$ in air and decent power handling capability. With a fully differential measurement scheme, the resonance peak is clearly resolved in the transmission measurement without the need of de-embedding the feedthrough signal. Compared to the capacitive type CMOS-MEMS resonators, this work offers a CMOS-compatible dc bias without tiny transducer’s gap, which can potentially be implemented as highly sensitive mass/gas sensors in the future.

Acknowledgment

The CMOS chip fabrication was supported by the CIC and TSMC, Hsinchu, Taiwan. The authors are grateful to the Cent. for Nanotech., Materials Sci. and Microsyst. of National Tsing Hua University for the use of fabrication facilities.

REFERENCES

Comparison of Acoustic Wave Pressure Sensors for TPMS applications

Manohar B. Nagaraju *, Suresh Sridaran †, Andrew R. Lingley *, John D. Larson III †, Brian P. Otis *, Richard C. Ruby †
* Electrical Engineering, University of Washington, Seattle, WA, USA
† Wireless Semiconductor Division, Avago Technologies, San Jose, CA, USA

Abstract—We study and compare the pressure sensitivities of different area Rayleigh Lamb wave (RL) mode (S\textsubscript{0} mode) and FBAR resonators. The studied RL-mode and FBAR resonators operate at 785MHz and 628MHz respectively. The resonators are fabricated on a released membrane with AlN as the piezoelectric layer. The resonators are hermetically sealed and the manufacturing process uses standard micromachining techniques throughout. The devices exhibit a pressure sensitivity over a range of 15 – 80 psi, suitable for Tire Pressure Monitoring Systems (TPMS). The sensitivities of different area resonators are compared.

I. INTRODUCTION

Micro-electromechanical systems (MEMS) based pressure sensors have gained a significant place in the sensor market, mainly for microphone applications. Pressure sensors are also being increasingly used and explored for new applications; in automotive (TPMS systems), mobile devices and wearables for indoor navigation, industrial and aerospace industries. In this study, we explore acoustic based pressure sensors for TPMS applications.

A TPMS system has become an essential component in modern vehicles as stipulated by the National Highway Traffic Safety Administration (NHTSA) in 2006 in the United States. Following similar legislations in other countries, the global demand for TPMS is expected to be 54.77 million sets in 2017 with an average annual growth rate of 16.83%. We have previously demonstrated a fully integrated FBAR-based pressure sensor with the sensor, processing circuitry, a frequency reference for communication and the transmitter in a single hermetically sealed die [1]. In this work, we demonstrate the use of Rayleigh Lamb wave (RL) mode resonators as a pressure sensor to improve the sensitivity for TPMS applications; and compare the performance of different area resonators.

The rest of the paper is organized as follows: Section II discusses the theory, design and processing of the RL-mode resonator. Section III presents the measured results and Section IV concludes the paper.

II. THEORY, DEVICE DESIGN AND FABRICATION

Lamb Wave Resonators (LWR’s) exploit the advantages of a robust design with low sensitivity to technological tolerances as in SAW resonators and the IC compatible manufacturing process of FBAR resonators. Electrical excitation of the wave is typically achieved, as in the SAW case, by Inter-Digitated-Transducers (IDT’s). The lowest order symmetric lamb wave mode (S\textsubscript{0}) is used extensively in research due to its high velocity, low dispersion and moderate piezoelectric coupling [2].

The structure of a Interdigitated RL-mode resonator is shown in the Figure 1. Interdigitated electrodes form the input and output for the resonator. The electrodes are anchored on all sides to provide a sealed cavity on one side of the membrane for a differential pressure sensing. The resonant frequency is dependent on the spacing between the electrodes (d) and the acoustic phase velocity (V\textsubscript{a}(t)), for a given plate thickness, and is given by equation 1. Figure 2 shows the measured |S21| characteristics of a Lamb-wave resonator.

\[
f_{\text{resonant}} = \frac{V_a(t)}{2d} \tag{1}
\]

A COMSOL FEM simulation is used to illustrate the different modes in the resonator. Figure 3 and Figure 4 shows the in-phase and out-of-phase S1 modes for the resonator at 785MHz and 804MHz respectively.
Acoustic resonators (FBAR and lamb-wave resonators) have been exploited for their usefulness as mass [3] and pressure sensors [1] [4]. The direction of propagation of the acoustic wave provides significant differences in the sensing methodology and performance between the FBAR and lamb-wave resonators.

For a given plate thickness, the phase velocity of the acoustic wave propagating in the piezoelectric membrane is related to the stiffness matrix, \( c_{eff} \) and the mass density \( \rho \) of the membrane as:

\[
V_a = \sqrt{\frac{c_{eff}}{\rho}}
\]  

(2)

The stress, strain and the electric field relationship for a piezoelectric membrane is defined as in equation 3, where \( 'c' \) is the stiffness co-efficient matrix at constant electric field and \( 'e' \) is the piezoelectric stiffness. Recent studies have proposed a non-linear model to model the second harmonic (H2) and intermodulation distortion (IMD3) data simultaneously [5] and is modeled as in equation 4.

\[
T = c^E S - eE
\]  

(3)

\[
T = c^E S - eE + \Delta T
\]  

(4)

\[
\Delta T = -\delta_1 eS^2 + \frac{\delta_2 e^2 E^2}{2} + \frac{\delta_3 c^E S^2}{2}
\]  

(5)

Application of pressure on one side of the membrane creates a bending stress in the membrane, changing the acoustic phase velocity \( V_a \) and hence the resonant frequency. Compared to a FBAR, since the direction of propagation of the acoustic wave and the stress change are in the same plane, the pressure to frequency deviation sensitivity is higher. Considering only the 2\(^{nd}\) order terms in equation 5 and using stiffness coefficient matrices \( c_1 \) and \( c_2 \) for the first and the second order effects, the stress in the membrane is

\[
T = c_1 S - c_2 S^2
\]  

(6)

The piezoelectric membrane has an initial stress, \( S_{int} \), developed by the manufacturing process. A bending stress in
the membrane due to a pressure input results in an additional stress \(S_{\text{ext}}\), and can be expressed as
\[
T = c_1(S_{\text{int}} + S_{\text{ext}}) - c_2(S_{\text{int}} + S_{\text{ext}})^2
\]  
\(T\) is equivalent to the operating point and \(S_{\text{ext}}\) is a small signal input (external pressure) on the piezoelectric membrane. Expanding and rearranging equation 7, we get,
\[
T = (c_1 + 2c_2S_{\text{ext}})S_{\text{int}} + c_2S_{\text{int}}^2 + c_1S_{\text{ext}} + c_2S_{\text{ext}}^2
\]  
(8)

For small stress input, the 3\(^{rd}\) and the 4\(^{th}\) terms are negligible. Comparing with the equation 6, the effective stiffness matrix \((c_{1})\) is approximately equal to
\[
eff = c_1 + 2c_2S_{\text{ext}}
\]  
(9)

The change in the effective stiffness coefficient translates to a resonant frequency drift due to the applied pressure (external pressure input), forming a pressure-to-frequency transducer.

**B. Processing**

The fabrication of the RL-mode resonator follows the standard processing methodology as in a FBAR resonator. The cavity under the RL-mode resonator is defined by etching into silicon. The cavity is filled with a sacrificial oxide and polished back to the silicon surface. The 1\(\mu\)m bottom metal layer is deposited first and acts as the grounding electrode. A 1\(\mu\)m thick Aluminium Nitride (AlN) piezoelectric layer is deposited on top followed by a 1\(\mu\)m top metal deposition. The top metal is patterned to form the inter-digitated fingers for excitation of the RL mode. A layer of gold is patterned on top of this to create the pads.

A lid wafer is processed with standard micromachining techniques and through-wafer vias are etched to make external connections. The base wafer is then diffusion bonded to the lid wafer to form a sealed cavity at low pressure on top of the RL-mode resonator. A standard Deep-Reactive-Ion-Etching (DRIE) process is then carried out, with the sacrificial oxide as the stop layer, to provide a channel for the pressure input. A HF dip of the wafer removes the sacrificial oxide and releases the RL-resonator

**III. Measured Results**

Figure 5 shows a photograph of the RL-mode die. The die was mounted on a PCB and placed in a pressure chamber. The frequency response (\(|S21|\) characteristic) of the test devices was measured using a vector network analyzer (E5071C) while applying a differential pressure across the piezoelectric membrane, controlled by a pressure regulator.

An RL-mode die with an area of \(147k\mu m^2\) was first used for the pressure response experiment. Figure 6 shows the frequency shift of the \(S_1\) in-phase mode (785MHz) with an increase in pressure. The sensitivity is higher (16.4ppm/psi) in the low-pressure range from 15psi to 25psi and reduces to 6.35ppm/psi over the range of 25psi to 80psi. The non-linear behavior (differing slopes in the high pressure and low pressure region) needs further exploration to reduce the calibration procedure. The response bottoms out at 80psi. The sensitivity is higher than the sensitivity of 2.2ppm/psi observed in a FBAR device [1]. The frequency shift can be mapped to pressure values using a 3-point (parabolic fit) calibration scheme and figure 7 shows the measured pressure vs the pressure input. The maximum error in the pressure measurement is ±0.2psi.

The pressure response of RL-mode devices for different areas (105k\(\mu\)m\(^2\) and 73.5k\(\mu\)m\(^2\)) was characterized as shown in the figure 8. The pressure response showed a similar characteristic as the larger area device, with a lower sensitivity. The 105k\(\mu\)m\(^2\) device had a sensitivity of 13.25ppm/psi in the low-pressure range from 15psi to 25psi and a sensitivity of 5.77ppm/psi over the range of 25psi to 75psi. The sensitivity values in the same pressure range was 11.2ppm/psi and 5.33ppm/psi for the 73.5k\(\mu\)m\(^2\) device. The FBAR resonant frequency pressure response, monitored through an oscillator...
Fig. 8. Performance comparison of different area RL-mode devices as pressure sensors

circuit, is also plotted for comparison.

IV. CONCLUSION

We have demonstrated the feasibility of RL-mode resonators as pressure sensors for TPMS applications. The pressure sensitivity is higher for $S_1$ mode Lamb wave compared to the FBAR mode. We achieve this in a wafer scale fabrication process using standard micromachining process throughout, making it viable for mass production. The non-linear behavior (differing slopes in the pressure response) needs further exploration.

REFERENCES


Micromechanical Piezoelectric-on-Silicon BAW Resonators for Sensing in Liquid Environments

Abhinav Prasad, Ashwin A. Seshia
Nanoscience Centre, Department of Engineering
University of Cambridge
Cambridge, UK CB4 1AZ
ap676@cam.ac.uk

Jérôme Charmet
Department of Chemistry
University of Cambridge
Cambridge, UK CB2 1EW

Abstract—This paper reports micromachined piezoelectric-on-silicon bulk acoustic wave resonators operating at a nominal frequency of approximately 3.15 MHz in fluidic media. Electrical measurements of the open-loop response of the resonators when one of the resonator surfaces is submerged in water indicate high quality factors in the range of 110-190. These values of quality factor are at least an order of magnitude higher than the flexural mode counterparts. The resonators are further exposed to Glycerol-Water mixtures of varying viscosity-density resulting in characteristic negative resonant frequency shifts. Experimental values are compared with a simplified liquid loading model and an agreement of up to 13% for highest and within 3-4% for lowest glycerol concentrations is established. These devices due to the relative ease of operation in liquid environments, scalability, high quality-factors and high mass-sensitivity have the potential for integration with microfluidics and electronics in order to realize an integrated platform for biochemical sensing and analysis.

Keywords—BAW resonator; piezoelectric transducer; MEMS; liquid sensing;

I. INTRODUCTION

The development of miniaturized sensors for probing small volumes of liquid samples is of interest for a multitude of applications. In the case of point-of-care (POC) diagnosis, for example, small volumes of liquid sample interfaced to an appropriate sensor can be screened for a panel of disease markers [1]. Another application for probing the physical and chemical properties of tiny liquid samples is driven by the chemical and pharmaceutical industry for the development of functional liquids, inks or appropriate drug targets. Hence, an important aspect enabling such a technology is the development of sensors that are small, inexpensive, perform real-time measurements and allow straightforward integration with signal conditioning electronics.

Micromachining technology has enabled batch manufacturing of such micro-/nano-scale sensors which can transduce a signal arising from the interaction of the sensing element with a target analyte to an electrical or optical signal. Among the various transduction methods researched, mechanical transducers, which are based on the measurement of the change in the mechanical properties due to the interaction with the environment, have shown promising progress in recent years. Moreover, these types of transducers can be readily integrated with the signal conditioning electronics. In the case of dynamic mode mechanical transducers, also called acoustic resonators, the mechanical structure is vibrated at its resonance frequency and any modulation of frequency response due to surface interactions can be recorded electrically.

These transducers have demonstrated a high degree of sensitivity in vacuum, and in some cases even achieving atomic resolution [2]. However, when these transducers are operated in air or liquid, which is critical for the development of Lab-on-Chip (LOC) platforms, they suffer degradation in the performance due to a) high level of ambient damping, and b) interference of the liquid sample with signal transduction. Bulk acoustic wave (BAW) resonators demonstrate higher quality factors which is an important metric defining the performance of such devices, due to reduced effect of viscous losses as compared to their flexural counterparts [3], and, hence, are ideal candidates for liquid based operations. One such approach, for example, involves mechanical coupling of identical BAW resonators where liquid is interfaced to one resonator and the electrical signals are transduced from the other device [4]. However, the design of the fluidic interface is relatively more complex and a modest reduction in overall sensitivity with the coupled approach is observed as well.

To overcome the liquid related challenges more effectively, piezoelectrically transduced micromechanical square BAW resonators are reported here which demonstrate high quality factors of 110-190 in water. Experiments for different Glycerol-Water mixtures were performed without liquid sample interference with the electrode structure or the signal conditioning electronics. A simple liquid loading model was further derived from [5] and was utilized to compare the experimental values with theoretically predicted values.

II. DEVICE DESCRIPTION

A. Fabrication of square-shaped resonators

The piezoelectric resonators used in this study were fabricated through industrial Multi-User MEMS processes (MUMPs). The resonators fabricated in this process were built into the device layer of a Silicon-on-Insulator (SOI) wafer with thin films of AlN deposited and subsequently patterned. On top of AlN film, Cr/Al metal pads were sequentially evaporated. The resonators are square shaped and are suspended from the
substrate through T-shaped anchors (Fig. 1(a)). As the structures were released from the substrate though bulk-micromachining, etch cavities (trenches) exist on the bottom side of the resonator structures. Some of the design, material and device parameters are mentioned in Table 1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device Silicon layer thickness (h)</td>
<td>10µm</td>
</tr>
<tr>
<td>Piezoelectric film (AlN) thickness</td>
<td>0.5µm</td>
</tr>
<tr>
<td>Metal electrode (Cr/Al) thickness</td>
<td>1µm</td>
</tr>
<tr>
<td>Resonator edge length (L)</td>
<td>1400µm</td>
</tr>
<tr>
<td>Square-extensional frequency (air)</td>
<td>~3.16 MHz</td>
</tr>
<tr>
<td>Quality factor (air)</td>
<td>2220</td>
</tr>
<tr>
<td>Quality factor (water)</td>
<td>110-190</td>
</tr>
</tbody>
</table>

B. Acoustic mode of operation

The resonators are actuated in the first extensional mode, also known as the square-extensional (SE) mode. This mode shape can be approximated by the superposition of displacement functions described in [6]. COMSOL simulations of the square-resonator for the SE mode is shown in Fig. 1(b). Applying the Rayleigh-Ritz theorem on the mode displacement functions and using a 1 degree-of-freedom mass-spring-damper lumped element model, the frequency $f_0$ can be expressed as $f_0 = \frac{1}{2\pi} \sqrt{\frac{k_0}{m_0}}$ where effective mass $m_0$ is $\rho_S h L^2$ and effective stiffness $k_0$ is $\pi^2 E_S h$ while $\rho_S$ is the density and $E_S$ is the Young’s modulus of Silicon. For an addition of small mass $\Delta m$, change in frequency $\Delta f$ is expressed by the relation $\Delta f = -\frac{f_0}{2 m_0} \Delta m$.

III. ELECTRICAL CHARACTERIZATION

Piezoelectric resonators are actuated and sensed using a two-port configuration as shown in the Fig. 2. A pair of Cr/Al metal electrodes is used for actuation while the remaining pair of electrodes is used for sensing. The device layer Si is grounded to reduce feedthrough parasitic. When an alternating voltage ($V_a=225 mV$) is applied between the actuating electrodes and device Si, a vertical electrical field is generated across the piezoelectric AlN film which produces an in-plane excitation force that mechanically deforms the underlying device layer through the piezoelectric effect. Governing relationships for such an excitation are elaborated in [7]. Transmission parameters ($S_{21}$) of the resonator are measured using a Network Analyzer, and the magnitude and phase frequency plots are recorded. In Fig. 3, transmission magnitude and phase plots in air are shown for one of the resonators. Quality factor and resonance frequency is extracted using the 3dB technique. Quality factor of the SE mode for the piezoelectric resonators are lower as compared to equivalent electrostatic resonators (~10000) [8]. In this case, the quality factor in air is potentially limited by dissipation at the interfaces of the device layer stack and electrical loading due to electrode resistivity [9].

Fig. 1. (a) Optical micrograph of a square resonator (L=1400µm), (b) COMSOL simulation for Square-extensional mode of vibration.

Fig. 2. Circuit schematic for two-port actuation and sensing for square resonators.

Fig. 3. Transmission magnitude and phase plot of the square resonator in air.
IV. LIQUID MEASUREMENTS

A. Liquid containment

One of the many challenges of interfacing electrostatic resonators to liquid is the possibility of liquid penetrating the transduction gap (between electrode and resonator proof mass) and shorting the device, or at least deteriorating the response to an extent such that electrical characterization is not possible. Poor electromechanical coupling in case of electrostatic resonators and high associated capacitive feedthrough parasitic make the detection of already damped peaks very difficult. Piezoelectric resonators propose simple solutions to these liquid related challenges. As there are no suspended electrode structures (electrodes are patterned on the top surface instead), the liquid flowing between the transduction gaps to electrodes simply does not arise. Due to the existence of a 400µm back-etch cavity on the backside of the resonator, it is possible to interface to liquid samples via the bottom surface. Good wetting properties of the cavity side walls resulting in strong capillary forces helps the liquid stay within the cavity and, thus, prevents the liquid from creeping through the gaps around the resonator to the opposite side where electrodes and wire bonds are situated (Fig. 4). Hence, back etch cavities of volumes of about 1µL can be used as “liquid reservoirs”. Also, due to more efficient electromechanical coupling in the case of piezoelectric resonators, even when the feedthrough parasitic is high, the resonance peak in most cases is distinct enough to enable effective electrical characterization.

B. Liquid loading

When laterally vibrating resonator’s surface comes in contact with liquid media, an evanescent shear wave is generated at the resonator-liquid boundary. The effective penetration depth \( \delta_l \) of such an evanescent wave is defined by the following expression [10]:

\[
\delta_l = \sqrt{\frac{\eta_l}{\pi f_l \rho_l}}
\]

where, \( \eta_l \) is liquid dynamic viscosity, \( \rho_l \) is liquid density and \( f_l \) is the operating frequency of the resonator in the liquid. Thus, for an operating frequency of 3.15 MHz, the penetration depth can be calculated to be around 318 nm for water solutions. It may be assumed that the liquid in this layer is effectively probed by the resonator, and, hence, contributes to mass loading. Using the Rayleigh-Ritz principle, an approximate expression for the effective mass of such a layer for SE mode can be calculated [5]:

\[
\Delta m_l = \rho_l \delta_l L^2 = (\sqrt{\frac{\eta_l \rho_l}{\pi f_l}}) L^2
\]

Hence, the change in resonant frequency of the device due to liquid mass loading is given by:

\[
\Delta f = -\left(\frac{f_l}{2m_0}\right) \Delta m_l
\]

It is important to note here that we have assumed that there is no perturbation in the strain energy of the resonator-liquid system. There are two important conclusions from the expression listed in Eq. 3:

- Change in frequency is dependent on the viscosity-density product of the liquid media.
- Irrespective of the height of liquid column above the resonator surface, only a thin layer of fluid participates in the liquid mass loading. However, this is an ideal scenario where leakage of vibrational energy in the longitudinal direction in form of compressional waves is ignored [11].
C. Experimental results

To establish the performance of the piezoelectric resonators in liquid, several experiments involving DI water and Glycerol-Water mixtures of varying concentrations (0%, 4%, 10%, 16%, and 20% by w/w) were conducted. Quality factors of 110-190 in DI water were observed for various resonators. This variation can be attributed to different surface conditions for different resonators and varying leakage of vibrational energy due to out-of-plane displacements in different resonators.

Before starting the liquid experiments, resonator bottom surfaces were treated with N₂ plasma for 60 seconds in a Femto plasma cleaner from Diener Electronic to remove organic contaminants and as well as to enhance the wetting properties of the cavity. Before introducing liquid samples, base-line frequency readings were taken. After every experiment, resonators were flushed with DI water three times to remove adsorbed glycerol on the exposed surfaces. Transmission magnitude plots for the resonators are shown in Fig. 5 which demonstrate that the negative frequency shift observed depend on the liquid mixtures. As the amount of glycerol increases in the mixtures, the density and the dynamic viscosity increase. This effect coupled with the negative shift in the frequency increases the penetration depth, thereby, increasing the liquid mass loading. The theoretical liquid loading values assume that the liquid completely wets the bottom surface. Experimental values match theoretical values closely at low glycerol concentration and increase to maximum deviation of 13% from the predicted values for 20% glycerol-water mixtures. This deviation is attributed to the fact that the residual glycerol present from the previous steps which could not be removed through repeated flushing steps. This issue can be overcome by integrating the resonator with a microfluidic flow-cell which will enable prolonged flushing of the surface. A good fit between mean frequency shift and square root of viscosity-density product is obtained for the experimental data providing validation for the analytical model presented in Eq (3).

![Fig. 5. Magnitude plot of piezoelectric resonator for the resonator in air, when one surface is in contact with DI water and when one surface is in contact 16% w/w Glycerol-Water mixture.](image)

![Fig. 6. Resonant frequency shift response for the square-extension mode (experiment, theory) with different Glycerol-Water mixtures in contact with one surface of the resonator.](image)

![Fig. 7. Plot of mean frequency shift vs. viscosity-density product for different mixtures.](image)

V. CONCLUSIONS

Micro-/nano-scale resonant sensors co-integrated with fluidic handling and electronics have been proposed as potential candidates for various applications in biochemical sensing and environmental monitoring. For devices operating in liquid media, it is essential to ensure that degradation in performance associated with liquid operation is mitigated. Piezoelectric-on-Si BAW resonators presented in this paper demonstrate high quality-factors in liquid due to lower dissipation as compared to flexural mode microcantilevers and improved liquid handling as compared to electrostatic BAW resonators. When exposed to water-glycerol mixtures of varying viscosity and density, a predictable change in resonant frequency is observed. Further, these resonators can be scaled to increase mass sensitivity. Follow-on experiments to test the response of similar appropriately functionalized devices to biomolecular analyte binding in buffer will be carried out as a next step towards the realization of a label-free biosensor platform.
ACKNOWLEDGMENT
The authors would like to thank the Cambridge Trusts and the W.D. Armstrong fund for financial support.

REFERENCES
Stress sensitivity coefficients of HBAR

Baron Thomas, Petroni Valerie, Martin Gilles, Combe Guillaume, Clairet Alexandre, Dulmet Bernard
Time & Frequency department, FEMTO-ST, UMR CNRS-UFC-ENSMM-UTBM 6174, 26 Chemin de l’Epitaphe, 25030 Besançon Cedex, France
E-mail: thomas.baron@femto-st.fr

Lesage Jean-Marc
DGA – Information Superiority, DGA, French MoD, Bruz, France

Laroche Thierry, Ballandras Sylvain
Frec|n|sys, Temis Innovation, 18, rue Alain Savary, 25000 Besançon, France
E-mail: sylvain.ballandras@frecnsys.fr

Abstract—Vibration sensitivity is an important specification for oscillators dedicated to space or airborne systems. Vibration sensitivity can be due to the resonator, the oscillator loop or non-oscillator components like wire, for instance. Commonly, the main source of acceleration sensitivity is due to the resonator. Active compensation can be used to decrease this effect, but such systems are not easily miniaturized. This paper presents computations of the stress sensitivity coefficients of frequency for the high-overtone bulk acoustic resonators and the design of a simple packaging to minimize vibration sensitivity. The final goal is to control vibration sensitivity of the high-overtone bulk acoustic resonators with dedicated packaging. The computed results are compared to experimental ones. The agreement between theoretical and experimental results is about 50%.

Keywords—High-overtone Bulk Acoustic Resonators (HBAR) component; oscillator; vibration sensitivity; packaging.

I. INTRODUCTION

One of the challenges of frequency sources dedicated to space and airborne systems is the control of the acceleration sensitivity of the oscillator arising from shocks and vibrations. Until now, acoustic resonators such as Bulk Acoustic Wave (BAW) or Surface Acoustic Wave (SAW) resonators present a g-sensitivity around $5 \times 10^{-10}/g$ for SAW operating in the 300-600 MHz range and around few $1 \times 10^{-9}/g$ for BAW in the range 10-100 MHz in the best case. As the operating frequency of acoustic wave devices tends to increase, new resonator principles have been investigated recently. Particularly, High-overtone Bulk Acoustic Resonators (HBAR) combining GHz-range operation capabilities with maximum quality factor $Q$ achievable along this principle have been investigated and new (two-port) resonator architectures have been proposed [1]. Intrinsic temperature compensation of such resonators had been demonstrated [2] [3]. A previous work has demonstrated experimentally a low vibration sensitivity of High-overtone Bulk Acoustic Resonators (HBAR) [4]. The corresponding experimental results show a global vibration sensitivity of $3.9 \times 10^{-11}/g$ for HBAR based on AlN on Sapphire, and 2.6 or 2.9$\times 10^{-9}/g$ for HBAR based on LiNbO$_3$ piezoelectric layer on Quartz or LiTaO$_3$. To reproduce and improve such results, the resonator design must be supported by accurate computations. However, the calculation of the stress sensitivity of HBAR resonators imposes some theoretical developments and the implementation of an ad hoc simulation tool.

This paper describes the theoretical approach used to calculate the HBAR stress sensitivity coefficients of frequency. To validate the approach, we compare the results of computation in the case of HBARs with a quartz substrate, due to the knowledge of non-linear elastic constant. The computation consists of one calculation of the stress sensitivity coefficients of frequency and one computation of stress field taking account the packaging with PCB. The experimental setup is explained, and finally, a comparison between theoretical and experimental results along each direction of space is done.

II. STRESS SENSITIVITY COEFFICIENTS OF FREQUENCY

HBARs combine the outstanding properties of the strong coupling coefficient of the deposited piezoelectric thin film and of the high intrinsic quality substrates. The piezoelectric film and the two electrodes on opposite sides are used as a transducer whereas the acoustic energy is mainly trapped in the substrate, Fig. 1. Resonance frequencies correspond to integer numbers of half wavelengths in the entire thickness. Unlike Film Bulk Acoustic Resonator (FBAR) and Solidly Mounted Resonator (SMR) in which only odd overtones exist, both odd and even overtones are compatible with resonance mode electrical and mechanical boundary conditions. For more details, the reader can consult [5].

This work was partly supported by the RAPID project ORAGE under grant #092906659#.
This work was partly supported by the French RENATECH network and its FEMTO-ST technological facility.
This project has been performed in cooperation with the Labex ACTION program (contract ANR-11-LABX-0001-01).
COMSOL computations have been performed on the supercomputer facilities of the Mésocentre de calcul de Franche-Comté.

978-1-4799-8866-2/15/$31.00 ©2015 IEEE 214
A. Theorie

A first attempt to calculate the stress sensitivity coefficients of thin film resonators (FBARs) has been proposed by Masson et al. [6] but this approach is not easily adaptable to the treatment of HBAR resonators which are composed of several crystalline materials. In this approach, we consider the effects related to elastic constants. The tridimensional variational formulation for elastic problem can be written in Lagrange coordinate:

\[
\frac{\delta \omega_{\alpha}}{\delta a_{i}} \cdot A_{ijkl} \cdot \frac{\delta u^{\alpha}_{j}}{\delta a_{k}} + \rho \cdot \omega_{\alpha} \cdot \frac{\partial u^{\alpha}_{0}}{\partial t} \cdot dV = \alpha \cdot \delta u^{\alpha}_{i} \cdot F_{i} \cdot dV + \delta u^{\alpha}_{\beta} \cdot T_{\beta} \cdot n_{j} \cdot dS
\]

(1)

Index 0 correspond to the non-perturbed state, while the displacement is denoted by \(u_{\alpha}\), the surface forces by \(T_{\beta} n_{j}\), the volume forces by \(F_{i}\), the pulsation by \(\omega_{\alpha}\), and the mass density by \(\rho\).

Then, let us consider the Sinha-Tiersten perturbation method [7] [8] [9] which involves the computation of static \(U_{ijkl}\) and dynamic \(H_{ijkl}\) terms.

\[
\Delta \omega = U_{ijkl} H_{ijkl}
\]

(2)

\[
H_{ijkl} = \left( \delta_a \delta_{ijkl} \delta_{mn} + C_{ijkl} \delta_{jmn} + C_{ijkl} \delta_{jmn} + C_{ijkl} \delta_{jmn} \right) \Omega
\]

(3)

\[
U_{ijkl} = \frac{1}{2} \rho_{ijkl} \cdot \frac{\partial u^{\alpha}_{0}}{\partial t} \cdot dV
\]

(4)

We adapt the method to carry all necessary integrations across the different layers of the stacked HBAR structures. Assumption of constant stress in the various layers is done allowing the derivation of HBAR stress sensitivity coefficients of frequency \((\alpha_{mn})\). It is possible to consider vertically—inhomogeneous stresses in our model by increasing artificially the number of layer of the HBAR stack. Only mechanical terms are taken into account and the contribution of the piezoelectric constants were deliberately omitted in the perturbation equations since the electromechanical coupling rarely approaches the percent in HBARs.

B. HBAR stress sensitivity coefficients of frequency computation

All non-linear constants of material were found in Error! Reference source not found.. The knowledge of LiNbO₃ and quartz non-linear constants will allow us to validate the computation of the stress—sensitivity coefficients of frequency of HBARs. We firstly compute these coefficients and with the separate computation of stress field, we will obtain computational results suitable for comparison with experimental results.

Fig. 2 shows the evolution of stress sensitivity coefficients for the case of HBAR based on (YX)/163° LiNbO₃ piezoelectric layer and (YX)/35°/90° quartz substrate. The stress sensitivity coefficients are slowly changing along the overtone number but we note a maximum variation near the piezoelectric layer fundamental resonance. For other overtones, stress sensitivity coefficients are almost constant and remain close to the coefficient of the single substrate plate. Since the stress coefficients obtained using the proposed calculation are very close to those obtained with the calculation of BAW \(\alpha_{mn}\) on single-crystal substrates, usual low sensitivity cuts for BAW can be used also for HBAR.

C. HBAR vibration sensitivity computation

Firstly, we compute stress field applied in quartz substrate with COMSOL, as shown in Fig. 3. We take into account PCB
of FR-4 with fixed boundary all around the PCB, the rectangular alumina which used to decreased stress field into the resonator and only the substrate of the HBAR for this computation. We take into account the cut orientation of the quartz substrate.

Due to the assumption of the computation of stress coefficient sensitivity, we assume a homogenous stress in all HBAR. Table II synthetized values of stress field calculate with COMSOL on the active volume of HBAR. The active volume of HBAR corresponds to the volume under the electrodes due to energy confinement at high frequency [11].

**TABLE II. STRESS TENSOR ALONG THE THREE COORDINATES OF SPACE FOR SOLICITATION OF 1G**

<table>
<thead>
<tr>
<th>Stress tensor (Pa)</th>
<th>Component of stress tensor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>X axis</td>
<td>2.72</td>
</tr>
<tr>
<td>Y axis</td>
<td>-111.6</td>
</tr>
<tr>
<td>Z axis</td>
<td>-34.01</td>
</tr>
</tbody>
</table>

Based on the computed results of stress tensor and of the stress–sensitivity coefficients of frequency, we easily calculate the vibration sensitivity of our oscillator by multiplying the stress tensor with the stress–sensitivity coefficient of frequency. The results are on table III.

**TABLE III. VIBRATION SENSIBILITY OF PACKAGED HBAR CONSTITUTED BY (YXLT)35°/90° QUARTZ SUBSTRATE AND (YXL)/163° LiNbO3 LAYER TRANSDUCER**

<table>
<thead>
<tr>
<th></th>
<th>X axis</th>
<th>Y axis</th>
<th>Z axis</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1.68e⁻⁹/g</td>
<td>-3.68e⁻⁹/g</td>
<td>1.15e⁻⁹/g</td>
<td>3.36e⁻⁹/g</td>
</tr>
</tbody>
</table>

III. **HBAR Resonator Characterization**

A. **HBAR oscillator**

We used a Quartz/LiNbO₃ HBAR resonator to realize an oscillator. A filter is inserted in the oscillator loop to select a specific overtone among the many ones accessible. The oscillator loop is also made with a low noise RF amplifier and the output is extracted by a coupler. HBAR oscillator has been built in a package 3×4 cm², as shown in Fig. 4. The HBAR resonator was conditioned onto a dedicated PCB and alumina.

B. **HBAR vibration sensitivity measurement**

Typical frequency shifts in oscillators due to vibration are on the order of 10⁻⁸ to 10⁻¹⁰ per g (acceleration of gravity near the earth’s surface) [12]. The acceleration sensitivity of an oscillator was explained in detail by Filler, [13]. Two main equations can help us to determine the vibration sensitivity of our oscillator:

\[
\Gamma_i = 10^{\frac{\chi_i}{2B}} \frac{2\nu}{\sqrt{\nu_0^3/BW}}
\]

\[
\Gamma = \sqrt{\Gamma_1^2 + \Gamma_2^2 + \Gamma_3^2} \cdot x = x', y', z
\]

\(\chi\) is the root-mean-square value of vibration, \(\Gamma_i\) is the component of acceleration sensitivity vector in the \(i\) (\(i=\text{X}, \text{Y}\) and \(Z\)) direction, \(\nu\) and \(\nu_0\) are respectively the Fourier frequency and the frequency of the oscillator, \(BW\) is the bandwidth of vibration and \(L_f\) is the phase noise.
The measurements of g-sensitivity have been achieved in all space directions on a test bench applying random vibrations in the 10 - 2000 Hz frequency range with 5 and 7 g rms intensity levels respectively. Random vibrations are applied vertically, and the oscillators were rotated in different position to achieve three directions. In Fig. 5, the position for X axis is shown.

The impact of acceleration on the phase noise of the HBAR-stabilized oscillator is illustrated in Fig. 6 showing the evolution of the phase noise when submitted to the above-mentioned perturbation. LiNbO₃/Quartz HBAR oscillators operating at 690 MHz experience a g-sensitivity of $2.59 \times 10^{-9}/g$. The g sensitivity along X axis is 4.03$\times 10^{-10}/g$, along Y axis is 2.2$\times 10^{-9}/g$, and along Z axis is 1.28$\times 10^{-9}/g$.

IV. DISCUSSION

The ratio between theoretical and experimental vibration sensitivity is 1.5. This result validates our approach to design HBAR integrated with the electronics of the oscillator to minimize the vibration sensitivity of the system. Indeed, some assumptions are done to calculate the stress sensitivity coefficients of frequency of the HBAR. One assumption consists in considering homogenous stresses in the device. It is possible to improve the accuracy of the computation by splitting the substrate into different layers, but the first result shows enough accuracy, and the control of the vibration sensitivity can be done only with the improvement of the stress induced by the packaging on the resonator.

The knowledge of the stress sensitivity coefficients of frequency allows us to work on the static stress induced by the approach of the packaging. Thus, it is possible to minimize some stress tensor components in function of the stress sensitivity coefficients of frequency of the HBAR. In the case of this paper, the HBAR presents negligible stress sensitivity coefficients of frequency for the yz and xz components, but high values for the y, z and xy components. So, some attention needs to be drawn on this stress tensor component for the design of the integration. Moreover, the y component of the stress--sensitivity coefficients of frequency is opposite to the xy component of the stress sensitivity coefficients of frequency. An approach to compensate both stress tensor components can be used to minimize the vibration sensitivity.

V. CONCLUSION

This paper presents the computation of the stress sensitivity coefficients of frequency of high-overtone bulk acoustic resonators. From these stress sensitivity coefficients of frequency, a simple static study of the packaging of the oscillator allows us to compute the vibration sensitivity of the system. The results are compared to experimental results. The theoretical and experimental results show an agreement of 50%.

REFERENCES

A 400μW Differential FBAR Sensor Interface IC with digital readout

Manohar Nagaraju *, Suresh Sridaran †, Andrew Lingley *, Reed Parker †, Richard Ruby †, Brian Otis *
* Electrical Engineering, University of Washington, Seattle, WA, USA
† Wireless Semiconductor Division, Avago Technologies, San Jose, CA, USA

Abstract — A low-power sensor interface IC suitable for a differential frequency measurement application is demonstrated. The circuit is used in a FBAR sensor system which includes a sensor and a reference FBAR. The sensor signal is processed and a digital output representing the sensor input is transmitted using a two wire serial interface. The architecture is entirely digital and provides a digital output while also providing a power interface architecture to process the sensor information on chip and provide a digital output while also providing a true differential frequency measurement.

I. INTRODUCTION

Thin Film Bulk Acoustic wave Resonators (FBAR’s) provide a very high quality factor of the order of a few thousand. The communication industry has used the high quality factor and the small footprint (< 1 mm³) of FBAR’s to its advantage, mainly in front-end filters, duplexers and oscillators. The f-Q product - one of the resonator performance metrics - is 10¹³ and comparable to a quartz crystal. This enables the design of ultra low power oscillators [1] and transceivers [2] with performance comparable to quartz based designs. New applications in wireless sensor nodes, ‘smart dust’ can benefit greatly due to a miniaturized, low-power, high-frequency solution for frequency generation using FBAR-based oscillators.

FBAR’s are also increasingly finding new applications as a sensing element. The resonant frequency of an FBAR is dependent on the thickness of the piezoelectric layer and the acoustic phase velocity through the bulk of the membrane. This property has been exploited for their usefulness in mass and pressure sensors. We demonstrated a fully-integrated mass [3] and pressure sensor [4] which integrates the sensor, processing IC, a frequency reference for a communication interface and a transmitter in a single hermetically sealed die. Since an IC wafer is used as the hermetic lid, it can theoretically use any type of IC process, reducing development cost and increasing design flexibility. A 0.6μm CMOS process was used for demonstration. In this work, we demonstrate a low-power interface architecture to process the sensor information on chip and provide a digital output while also providing a true differential frequency measurement.

The rest of the paper is organized as follows: Section II discusses the theory and practical considerations in designing interface circuitry for FBAR sensors. Section III discusses the proposed architecture and outlines the design of the key circuit blocks. Section IV presents the measured results and Section V concludes the paper.
C. Issues in FBAR-based Sensors

FBAR provides high sensitivity to any added mass/pressure, however there are practical limitations in using an FBAR sensor. The resonant frequency of the FBAR is sensitive to environmental variables like humidity, temperature, package-level stresses etc. Temperature-compensated FBARs have a temperature dependence of 50 ppm over −20 to 100 °C [3]. Additionally, the FBAR resonant frequency can drift by 100 ppm due to aging and stress. This greatly degrades the accuracy of the sensor. To cancel drift mechanisms to first order, we integrate two matched oscillator structures with two separate resonators (sensor and reference FBAR) in close proximity and use one of them as a reference to track the frequency change due to unwanted variables. The sensor input is applied to only the sensor FBAR. A differential frequency measurement then cancels the unwanted frequency drift to first order. The reference FBAR also provides a frequency reference for a communication interface.

III. PROPOSED ARCHITECTURE

Figure 1 shows the architecture of the proposed sensor interface IC. At the core of this architecture is the presence of a reference and sensor FBAR to perform a differential frequency measurement and canceling out any measurement inaccuracies due to temperature, package stress etc. The two FBAR’s are chosen to be 628 MHz FBAR’s. The oscillator interface with the FBAR resonator directly impacts the resolution of the sensor.

The FBAR resonant frequency is monitored with a pierce oscillator structure as shown in the Figure 2. The design goal of the oscillator is to reduce the integrated jitter to improve the sensitivity of the sensor. The far-off phase noise in an FBAR-based sensor is not a concern due to the high-Q of the resonator and the relatively long sampling time in a sensor measurement.
of the order of ms. However, close-in phase noise performance dominates the total integrated noise in the system and sets the resolution of the sensor. AM-PM (amplitude modulation to phase modulation) conversion arising from non-linear device parasitic dominates the close-in phase noise generation in an FBAR oscillator. A non-linear compensation capacitor was added to reduce the close-in phase noise [6].

**Architecture description:** The sensor oscillator frequency is estimated using counter based logic. The number of edges in the sensor oscillator is counted for a fixed duration. This duration is defined by counting a set number of edges from the reference oscillator. Any common mode change results in a drift of both the sensor and the reference oscillator. The output of the counter thus remains constant. However, a differential input resulting from a sensor variable (mass or pressure) change affects only the sensor oscillator and changes the counter output. Thus a change in the output of the counter directly provides the sensor response.

The measurement window for the counting operation is fully programmable through a serial interface. The integration time trades power consumption (active duty cycle) for the minimum noise (hence resolution) in the measurement of the sensor. Typically, integration time for minimum Allan deviation is in the range of $10 - 100$ ms for FBAR oscillators. The start-up time of a FBAR oscillator ($10\mu$s) is negligible compared to the time required for minimum Allan deviation. The counting operation is run for 1/2 cycle of the reference clock and the counter output is transmitted serially for the other 1/2 cycle of the reference clock. A sync-code (EB90) is used at the beginning of the serial transmission for synchronization at the receiver.

**Asynchronous counter:** The differential frequency measurement translates to a multiple clock domain operation. A signal crossing a clock domain, appears to the circuitry in the new clock domain as an asynchronous signal. This results in a metastable state in the first storage element (flip-flop) in the new clock domain. Systems with multiple clock domains handle this with specialized synchronization circuits that greatly increase the power consumption of the system [7]. In the proposed sensor system, the total duty-cycled power consumption is of the order of a few $\mu$W.

An asynchronous counting scheme is employed to save power without synchronizing the two clocks. The sensor oscillator is followed by a chain of dividers (divide by 2) and the output of the dividers is latched as a count value at the positive and negative edge of the divided reference clock. The difference between the latched counts provides an estimate of the number of edges in the given time period from which the sensor frequency could be estimated. Figure 3 shows the block diagram of the counting operation. The design of the first divider is critical since the input frequency is high (628MHz). A True Single Phase Clock (TSPC) flip flop based divider which incorporates high speed and low power consumption is used for the first divider. The power consumption of the subsequent dividers is negligible and standard static CMOS based circuits are used.

**Quantization Error:** A digital counting scheme inherently results in a quantization error in the frequency measurement. The division ratio or the counting period dictates the quantization error. The minimum quantization error with a modulo-N programmable frequency divider in the current architecture is given by equation 4. For a 628MHz FBAR, a divide ratio of $2^{26}$ translates to a resolution of 0.05ppm with a 53ms interval between successive measurements. The resolution is limited by the divide ratio rather than by the minimum noise (Allan deviation) in the proposed system.

$$Q_e = \frac{f_{\text{sensor}}}{2^N - 1}$$  \hspace{1cm} (4)

**IV. Measured Results**

The sensor interface IC was fabricated in a 0.13$\mu$m CMOS process. The IC was interfaced with a FBAR pressure sensor (Figure 4). The pressure sensor was processed as in [4] without the lid circuitry. The sensor interface IC consumes 530$\mu$A from a 0.75V supply with the digital processing consuming 120$\mu$A. For one cycle of measurement and transmission (100ms), with a duty cycle ratio of 1%, the average power consumption of the IC is 4$\mu$W.

The response of the pressure sensor was characterized by mounting the chip on a measurement PCB. The PCB was placed in a pressure chamber controlled by a pressure regulator. The programmable divider was programmed for a counting
Fig. 5. Transfer function (digital code vs pressure input) for the sensor interface IC

Fig. 6. Pressure sensor calibration curve from 20 to 80 psi

Fig. 7. Allan Deviation measurement from the sensor oscillator

The clock frequency of close to 10Hz. The serial output data consists of a 16-bit frame sync code $EB90$ followed by a 29-bit digital value representation of the sensor frequency.

Figure 5 shows the digital output code as the pressure inside was increased from 20psi to 80psi. The digital output codes represent the sensor frequency measurement. We observe a linear response and the pressure sensitivity is calculated to be $1.6$ ppm/psi.

The digital output code was then mapped to pressure values using the slope of the transfer function. Figure 6 shows the measured pressure as a function of the pressure input. The maximum error in the pressure measurement is $\pm 0.53$ psi.

The resolution of a FBAR sensor is determined by the minimum detectable frequency shift, which is specified by an Allan deviation measurement in an oscillator-based sensor interface. Figure 7 shows the Allan deviation measurement. The minimum detectable frequency shift is $3.9$ ppb at an integration time of 100ms. However, in the proposed digital sensor interface, the quantization error is limited by the measurement window. With a maximum divide ratio of $2^{26}$ and a 628MHz FBAR for deriving the counter clock, the minimum achievable resolution is 60ppb. This translates to a resolution of 0.037psi. Higher division ratios improves the resolution, however the increase in the noise of the system limits the resolution and the power consumption.

V. CONCLUSION

This paper demonstrated a low power digital interface circuit for operation with FBAR-based sensors. The instantaneous current consumption of the interface circuit is 530µA and operates off a 0.75V supply. The low supply voltage and the low instantaneous power consumption makes it feasible to operate the system with standard coin-cell batteries. The architecture is scalable to advanced process nodes and will benefit from process scaling. The digital output from the sensor interface IC is compatible for further digital signal processing.

ACKNOWLEDGMENT

The authors would like to acknowledge the efforts of Dorie Delapena for bonding services.

REFERENCES

Dual-mode NEMS self-oscillator for mass sensing

Guillaume Gourlat, Marc Sansa, Guillaume Jourdan, Patrick Villard, Gilles Sicard, Sébastien Hentz.
Univ. Grenoble Alpes, F-38000 Grenoble, France
CEA, LETI, MINATEC Campus, F-38054 Grenoble, France
guillaume.gourlat@cea.fr

Abstract—We report the first experimental demonstration of a heterodyne self-oscillator operating alternatively on the first and second flexural mode of a silicon NEMS resonator. This architecture features a downmixing scheme where the NEMS motion-induced piezoresitive signal at 25 MHz and 70 MHz is shifted down to few tens of kHz thus reducing the bandwidth constraint on the electronics. In closed loop operation, the oscillator presents excellent frequency stability, identical to the one obtained in PLL operation. While monitoring successively the two modes of the oscillator, mass addition on the NEMS was simulated by electrostatically-induced frequency shifts. This self-oscillator scheme represents low complexity and power saving architecture compatible with the readout of dense sensor arrays required in applications such as mass sensing.

Keywords—NEMS, Nanoelectromechanical system, resonators, oscillators, self-oscillators, heterodyne scheme, noise measurement, mass sensors, mass spectrometry, silicon resonator.

I. INTRODUCTION

Nano Electro Mechanical Systems (NEMS) constitute a promising solution for mass sensing applications [1] [2], which requires very high capture efficiency of the analyte only achievable by the increase of the sensing area brought by the co-integration of arrays of sensors with CMOS circuitry [3]. Two architectures have been reported in the literature. The homodyne self-oscillating (SOL) scheme [4] is the most compact, but is very sensitive to parasitic coupling and is not able to handle multimode operations that are necessary to deduce the mass of an individual particle landing on an unknown position of the NEMS. On the other hand, the Phase Lock Loop (PLL) handles multimode resonators with robustness vis-à-vis the parasitic capacitances but it implies power consuming and bulky circuitry, which does not scale favorably for large arrays of sensors. In this context, we present an improved heterodyne self-oscillator [5] capable of handling multimode operation with reduced footprint and consumption of the self-oscillator scheme when co-integrated in CMOS technology.

In this paper, characterization results of the silicon resonator designed for mass sensing applications will first be presented. Then, a complete description of the heterodyne self-oscillator scheme will be given and the frequency stability results of the open and closed loop operations on the first two modes of the resonator will be discussed. Finally, mass landing events will be simulated by electrostatically induced frequency shifts [6] and alternatively monitored during multi-mode operation of the self-oscillator.

II. DEVICE DESCRIPTION AND TRANSDUCTION METHOD

A. Device and fabrication

The nanomechanical resonator used here [7] is a monocrystalline silicon doubly-clamped beam with compliant anchors for enhanced dynamic range. The resonant beam is 160 nm thick, 300 nm wide and 10 µm long. As presented in Fig. 1, either actuation electrode actuates the beam and the in-plane motion is transduced in the electrical domain thanks to two piezoresistive nanoscale gauges disposed in bridge configuration for background cancellation. The device operates under vacuum (typically 10⁻⁵ mbar) at room temperature, around 297 K.

B. Transduction method

The device embeds a detection scheme compatible with a heterodyne transduction method in which the nanogauges in bridge configuration are biased with an alternative signal at a frequency close to the nominal resonant frequency of the resonator [8]. Therefore, the strain induced in the gauges mixes the biasing signal down to a low frequency at the NEMS output. This method is compatible with off chip bonding and an experimental setup built with discrete equipment.

Fig. 1. SEM image of the suspended in-plane doubly-clamped silicon NEMS resonator. The vibrating beam is 10 µm long, 300 nm wide and 160 nm thick. The left anchors and nanogauges are shorted. The rightmost nanogauges are independent of the right anchor and biased with complementary signals.
II. EXPERIMENTAL METHODS

The NEMS silicon resonator is excited in the open loop and closed loop scheme. We denote \( \omega_b \), \( \omega_0 \), and \( \Delta \omega \) the resonator nominal resonant frequency, the high frequency bias and the intermediate low frequency respectively.

III. HETERODYNE OSCILLATOR ARCHITECTURE

The concept of self-oscillating structure consists in feeding back to the input of the resonator a fraction of its own output while respecting a set of conditions known as the Barkhausen criterion (1). We define the open loop transfer function \( |H_{OL}(j\omega)| \) as the product of the resonator \( |H_{res}(j\omega)| \) and the electronics \( |H_{elec}(j\omega)| \) transfer functions. Since the detection method employed mixes the motion signal down to an intermediate frequency (IF), the resulting signal cannot be directly fed back to the input of the resonator. Instead, this signal has to be shifted up by mixing it with a signal (LO) synchronous to the one that was used during the downmixing scheme (AC Bias). The complete oscillator loop is composed of an amplifier featuring a built-in tunable band pass filter, a low pass filter and a passive mixer as depicted in Fig. 2. Outside the main feedback loop, we have a programmable oscillator with two outputs whose relative phase shift can be precisely set. The AC Bias signal is applied to the piezoresistive nanogauges while the phase shifted AC Bias signal drives the mixer LO input.

\[
|H_{OL}(j\omega)| = 1 \quad \text{and} \quad \arg(H_{OL}(j\omega)) = 0 \quad [2\pi] \quad (1)
\]

This heterodyne scheme allows independent tuning of the loop gain and phase. The gain condition is set by the tunable low noise amplifier while the phase is set outside the loop by tuning the phase shift between the nanogauges biasing signal and the LO input of the mixer.

A. Open loop characterization

The oscillator is first studied in open loop by setting the switch position to (a) as described in Fig. 2. The NEMS resonator response is plotted as a function of frequency for the first mode (Fig. 3a) and second mode (Fig. 3b) of operation. The nominal frequency and quality factor for each mode are respectively fitted as \( \omega_0/2\pi = 26.093 \text{ MHz}, \omega_0/2\pi = 72.048 \text{ MHz}, Q_1 = 6300, Q_2 = 4400 \). The gauges were biased with 750mV amplitude while the AC and DC drive levels were set respectively to 100 mV and 100 mV (mode 1) or 1 V (mode 2). We estimated respectively for mode 1 and 2 a required gain of \( G1 = 724 \) and \( G2 = 1219 \) to ensure oscillation during closed loop operations.

B. Closed loop operations

Closed loop operation of the oscillator is obtained by setting the switch position to (b) as described by Fig. 2. The AC Bias signal frequency was set close to the previously extracted resonator nominal frequency at \( \omega_0/2\pi = 25.959 \text{ MHz} \) and \( \omega_0/2\pi = 71.965 \text{ MHz} \). The low noise amplifier gain is set to around 1400 following the open loop estimation and taking into account that the upmixing implies a supplementary attenuation of 1/2. The phase condition is then met by monitoring the amplitude of oscillation while sweeping the phase shift between the AC bias and the LO signal sent to the mixer.

The self-oscillations shown in Fig. 4, recorded at the output, are stable at 98.3 kHz with 148 mV for the first mode and at 88.5 kHz with 122 mV for the second mode. We notice that the first mode is exactly at the expected intermediate frequency \( \Delta \omega_1 = \omega_0 - \omega_0 \) whereas the second mode is 6 kHz

Fig. 2. Diagram of the oscillator setup used on the first and second mode of the silicon resonator. Switch position (a) and (b) corresponds respectively to the open loop and closed loop scheme. We denote \( \omega_b, \omega_0 \), and \( \Delta \omega \) the resonator nominal resonant frequency, the high frequency bias and the intermediate low frequency respectively.

Fig. 3. Open loop response of the NEMS silicon resonator for the first mode (a) and second mode (b). Blue dot and red line are respectively the experimental amplitude data and its fit, and they are related to the left y-axis. The green dashed line which relates to the right y-axis corresponds to the experimental phase.

Fig. 4. Oscilloscope acquisition (a) and fast fourier transform (b) of the low intermediate frequency signal after amplification and filtering in closed loop for mode 1 in blue and mode 2 in red.
off of its estimated frequency $\Delta \omega_2 = \omega_{02} - \omega_{02}$. We explain this discrepancy by a slightly off gain parameter for the second mode.

As shown in Fig. 5, the oscillator frequency and its downmixed part can be tuned as a function of the phase shift introduced between the nanogauges AC Bias signal and the LO signal. The AC Bias phase was kept at zero while introducing a phase shift in the LO signal. The Barkhausen criterion states that the oscillations can build up only if at $\omega = \omega_{osc}$ the open loop gain is higher than 1 and the phase is equal to zero or a multiple of $2\pi$. The oscillator amplitude follows a parabolic trend, plotted in solid green line in Fig 5, explained by the quadratic dependence of the NEMS module (2) to a supplementary phase shift $\psi$, introduced by the electronics.

$$|H_{nems}(j\omega)| = |H_{nems}(j\omega_0)| \times (1-\psi^2)/2 \quad (2)$$

The oscillator frequency shift, plotted in Fig 5 as blue dots, is proportional to the phase shift introduced between the bias and LO signal with a factor of 29.9 Hz/deg for the first mode and a factor of 86 Hz/deg for the second mode related to the theoretical factor $\omega_{01,2}/(720Q_1,2)$. It is worth noticing that self-oscillations are triggered for a broad phase window of the LO signal which depends on the overall open loop gain: the higher the gain, the wider the window.

C. Stability of the heterodyne self-oscillator frequency

Frequency stability is a key parameter for mass sensing applications as it sets a limit to discriminate mass induced frequency jumps from the noise floor. The oscillator architecture should not degrade the resonator intrinsic frequency stability. Allan deviations of the resonator in open loop were compared to the one of the closed loop oscillator. The NEMS frequency stability was acquired in open loop by a lock in amplifier (LIA) driving the resonator at the resonance frequency while monitoring the phase variation of the demodulated output signal. Then, the frequency stability of the heterodyne oscillator in closed loop was characterized by locking a PLL on the intermediate frequency output signal. This recorded oscillator fluctuations $\delta\omega_0$ at the intermediate frequency directly relates to the mechanical resonator nominal frequency fluctuations $\delta\omega$ because none of the loop elements induces frequency sensitive phase shift.

As shown in Fig. 6, we were able to record the frequency stability of the oscillator in closed loop and to compare it with the intrinsic resonator fluctuations. We found a minimum Allan deviation in closed loop respectively for mode 1 and mode 2 around 1.9e-7 and 1.8e-7. We observed similar results to the one obtained in open loop, suggesting that the heterodyne self-oscillating scheme does not induce any additional noise. The discrepancy in the white noise level for the low integration times could be explained by the slight change of cut-off frequency due to the PLL controller in closed loop. This result demonstrates that the heterodyne self-oscillator works in the context of multimode operations without degrading the intrinsic frequency stability of the silicon resonator therefore enabling the use of this scheme for mass sensing purposes.

IV. SIMULATED FREQUENCY SHIFTS TRACKING

Single-particle mass sensing requires a readout scheme able to track the first two modes of resonance of a doubly clamped silicon beam. Our experiment demonstrates the alternative monitoring of these modes while simulating mass adsorption on the NEMS by applying DC voltage steps on the second electrode of the silicon resonator as shown in Fig. 7. As for the first electrode, this step generates an electrostatic force whose second order component induces negative stiffness which in turn changes the resonance frequency of the system.
While tracking alternatively the self-oscillator first and second mode of operation, we simulated frequency jumps by applying DC voltages steps of 0.5V and -0.5V as described by the plot in Fig. 7. These steps result in frequency jumps (Fig. 8a) every three seconds. As the reader may notice, the intermediate frequency of both modes does not relate to the one previously mentioned in the paper. This could be explained by the simultaneous biasing of the structure with the AC bias of the first and second modes. The biasing signal tends to heat the gauges therefore changing the resonance frequency of the device. Each mode is monitored during 100 ms and it takes the monitoring setup around 15 ms to lock every time a mode switch happens. Fig. 8b is a zoom around a specific DC jumps that happened while monitoring the second mode. A 500 µs response time to a mass induced frequency jump for the oscillator was recorded. This tracking experiment confirms the potential of the heterodyne oscillator architecture to track mass induced frequency shift.

**CONCLUSION**

This work demonstrates for the first time the multimode operation of a heterodyne NEMS self-oscillator scheme built around a symmetric doubly clamped resonator. A detailed study of the open-loop response and frequency stability of the resonator was performed to determine the Barkhausen conditions that ensure the oscillations in closed loop. In the oscillator configuration, sustained oscillations on the first two modes of resonance are easily obtained without tedious tuning. The closed loop operation showed excellent frequency stability that is only limited by the intrinsic frequency stability of the resonant element itself. Finally, mass additions were simulated by electrostatically-induced frequency shifts while monitoring alternatively the first two modes of resonance. This oscillator architecture promises reduced silicon footprint and power consumption compatible with the readout of dense sensor arrays required for mass sensing applications.

**ACKNOWLEDGMENT**

The authors thank the Carnot institute (Project NEMS MS) and “Direction générale de l’armement” for financial support.

**REFERENCES**


Abstract—We targeting analyzed the UTC link delay of GPS station PTBB of PTB to TWTF of TL and expressed it as 3 groups: the uncompensated GPS common clock difference measurements, cable delay measurement, and the total delay variations of fixed GPS stations and travelling calibrator used in this METODE calibration tour. The total delay variation of fixed GPS stations PTBB and TWTF was evaluated by monitoring their long-term CCD of the same type receiver more than 350 days. For the BIPM travelling calibrator BP1C and BP0U, we used a moving cesium clock method to evaluate their instability of total delay in different antenna position. Our study was helpful for clarifying the uncertainty composition of the current PTB-TL link and could reduce it.

Keywords—METODE Calibration; uncertainty;

I. INTRODUCTION

From MJD 56595 to 56650 (30 Oct. – 24 Dec. 2013), the BIPM (Bureau international des poids et mesures) METODE (Measurement of Total Delay, [1]) calibrators StdB (composed of BIPM GNSS traveling receivers BP0U and BP1C) visited TL (Telecommunication Laboratories, Chunghwa Telecom Co., Ltd. Taiwan) to calibrate the TL’s GPS reference station, TWTF, and the UTC (Universal Coordinated Time) link from PTB (Physikalisch-Technische Bundesanstalt, Germany) to TL [2]. The BP0U and BP1C’s total electronic delays, denoted by DlyR(BP0U) and DlyR(BP1C) respectively, from the phase center of their antennae (PCA) to their 1 PPS calibration reference point had been aligned by the total electronic delays of PTBB (the GPS pivot reference station of UTC at PTB), DlyR(PTBB), at MJD 56450 and MJD 56895 [3][4], thus the DlyR(BP0U) and DlyR(BP1C) could be used to calibrate/align the total electronic delay from the PCA of TWTF to the UTC(TL) 1 PPS reference point, DlyR(TWTF), and link delay between PTBB and TWTF (denoted as DlyL(PTBB, TWTF) hereafter).

The calibration result of DlyR(TWTF) and DlyL(PTBB, TWTF) had been calculated and announced by Jiang et al [2], and gave a general total uncertainty for METODE calibration to be about 0.8–1.5 ns (1σ). This uncertainty did not include the long-term variations of all devices including antennae, cables, receivers, and other non-stationary effects such as multipath [2].

Table I. GNSS station used for TL calibration tour

<table>
<thead>
<tr>
<th>Reference station</th>
<th>Station to be calibrated</th>
<th>receiver</th>
<th>Reference point</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTBB</td>
<td>TWTF</td>
<td>Z12T</td>
<td>UTC(TL)</td>
</tr>
<tr>
<td>BIPM StdB</td>
<td>BP0U</td>
<td>GTR-50</td>
<td>Calibration Point</td>
</tr>
<tr>
<td>BIPM StdB</td>
<td>BP1C</td>
<td>PolaRx3</td>
<td>Calibration Point</td>
</tr>
</tbody>
</table>

The aim of this paper is to targeting evaluate the uncertainty of DlyL(PTBB, TWTF) including its long-term variation. In this paper, we followed the equations we derived in our previous study [5], and decomposed DlyL(PTBB, TWTF) into 3 groups: the uncompensated GPS common view [6] common clock difference (CCD) measurements at PTB and TL, the cable delay measurement, and the total delay variance of fixed GPS station PTBB/TWTF and the travelling calibrator StdB including their delay instability of antennae, cables, receiver, multipath effect, and other unknown terms.

For PTBB and TWTF, they were both fixed (antennae, cables, receiver) Ashtech Z12T Metornome stations (Table I). In this paper, we monitored the long-term behavior of Ashtech Z12T Metornome (abbreviated as Z12T hereafter) at PTB and TL more than 350 days to evaluate the non-stationary effects of PTBB and TWTF, and gave a rational total uncertainty covered their long-term behavior.

For Z12T, we also noted its internal reference point was not rigorously coherent with its 1 PPS input [7][8], so that the internal reference point between TWTF (and PTBB) and the travelling calibrator StdB might not be coherent, we also discussed the long-term incoherence between the 1 PPS and 20 MHz input of TWTF in section II.

II. THE CALIBRATION SETUP AND UNCERTAINTY OF REFERENCE POINT OF TWTF

The setup configuration of METODE TL tour was showed in Figure 1. We set the UTC(TL) point to be the positive edge slope of UTC(TL) 1 PPS reference point (1 PPS distribution amplifier port A-5) output at 1 Volt level with 50 Ω impedance. UTC(TL) was generated from micro phase stepper, AOG-110 SN-001, referenced by the hydrogen maser HM0057. The 10 MHz and 20 MHz sine wave frequency reference are generated.
from the same AOG-110 via SDI–FS020 and SDI–FS040 multipliers.

The calibration point of StdB was just the UTC(TL) 1 PPS reference point, we didn’t need to measure any cable delay in TL in this tour. The uncertainty of cable delay measurement at TL would not be introduced into the total uncertainty of this calibration.

For Z12-T, its internal reference point was latched from the zero crossing of a particular cycle of external 20 MHz signal chosen by external 1 PPS input, therefore the internal reference point of TWTF (Z12T receiver) was not directly coherent with the reference point of StdB, UTC(TL), in this tour. To estimate the incoherence between the 20 MHz and 1 PPS input of Z12T receiver, we monitored their phase difference from MJD 56700~57110, and found the peak to peak incoherence was about 0.35 ns and with standard deviation was about 0.054 ns (Fig. 2). Therefore we thought the discrepancy of the latch point of TWTF and StdB at TL was 0.054 ns (1σ).

III. THE COMPOSITION OF DlyL(PTBB, TWTF)

In METODE link calibration process, any stable receiver system could be used as the origin (pivot) reference station. In this tour, the pivot reference station was the station PTBB at PTB. Its total delay, DlyR(PTBB), was treated as a reference to align the DlyR(TWTF) and to get the link delay DlyL(PTBB, TWTF) via 2 GPS common view CCD stages at PTB and TL.

Recalling the derivation of DlyL(PTBB, TWTF), it was formed with the uncompensated GPS P3 or PPP common view CCD measurements of PTBB and StdB at time \( t_1 \) (MJD 56450~), the uncompensated GPS P3 or PPP common view CCD measurements of TWTF and StdB at time \( t_2 \) (MJD 56595~), and necessary cable delay measurements. If the total delay variation of PTBB, TWTF, and StdB could not be neglected, the link delay DlyL(PTBB, TWTF) could be expressed as [5]:

\[
DlyL(PTBB, TWTF, t) = \left( \text{REFGPS}_{\text{raw}}(PTBB, t_1) - \text{REFGPS}_{\text{raw}}(StdB, t_1) \right) + \left( \text{REFGPS}_{\text{raw}}(StdB, t_2) - \text{REFGPS}_{\text{raw}}(TWTF, t_2) \right) - \left( \text{UTC}(PTB) - \text{CalP}(PTB) \right) + \Delta DlyR(PTBB, t_1, t) - \Delta DlyR(TWTF, t_2, t) - \Delta DlyR(StdB, t_2, t_1)
\]

Where

\[
\text{REFGPS}_{\text{raw}}(PTBB/TWTF/StdB): \text{the time difference between the internal reference point of GPS station PTBB/TWTF/StdB and the uncompensated GPS time measured by PTBB/TWTF/StdB, including the delays of their antennas, antennae cable, and receiver internal delay.}
\]

\[
\text{REFGPS}_{\text{raw}}(PTBB, t_1) - \text{REFGPS}_{\text{raw}}(StdB, t_1): \text{the uncompensated GPS common-view CCD measurement of PTBB and StdB at time } t_1 \text{ (MJD 56450~).}
\]

\[
\text{REFGPS}_{\text{raw}}(TWTF, t_2) - \text{REFGPS}_{\text{raw}}(StdB, t_2): \text{the uncompensated GPS common-view CCD measurement of TWTF and StdB at time } t_2 \text{ (MJD 56595~).}
\]

\[
\text{CalP}(PTB): \text{the 1 PPS calibration point of StdB at PTB. The CalP(TL) was just the UTC(TL) 1 PPS reference point, we didn’t need to measure the time difference of UTC(TL) – CalP(TL) in this tour.}
\]

\[
\Delta DlyR(PTBB, t_1, t) = DlyR(PTBB, t) - DlyR(PTBB, t_1)
\]

\[
\Delta DlyR(TWTF, t_2, t) = DlyR(TWTF, t) - DlyR(TWTF, t_2)
\]

\[
\Delta DlyR(StdB, t_2, t_1) = DlyR(StdB, t_1) - DlyR(StdB, t_2)
\]

\[
\text{IV. THE NON-STATIONARY TERM OF DlyL(PTBB, TWTF)}
\]

The last line of (1) contained the total delay variation of 2 fixed GPS stations (PTBB and TWTF) and BIPM travelling calibration StdB (BP0U and BP1C). The BIPM StdB were only operated during the time we processed CCD calibration. We typically used the closure test to estimate the uncertainty of travelling calibrators. At MJD 56450 and MJD 56895, the StdB visited PTB and their closure total delay discrepancy were less than 0.03 ns [4]. Consider the 2 CCD processes of this tour were not operated at the same place, our previous study [5] which used a moving cesium clock method showed the total delay variance of BP1C and BP0U were less than 1.13 ns over 25 km baseline.
For PTBB and TWTF, they were both fixed (antennae, cables, receiver) Z12T stations. Fortunately, both PTB and TL owned 2 operating Z12T receivers; they were PTBB/PTBG at PTB and TWTF/TWT1 at TL. Their long-term GPS P3 common-view CCD results more than about 350 days were showed at Fig. 3. To reduce the P3 code noise and simulate the actual calibration processes, the data showed at Fig. 3 were the 1 day moving average of preliminary GPS P3 common-view results.

Since the antenna, antenna cable, and reference cable of PTBB+PTBG and TWTF+TWT1 were fixed, so that the very short baseline GPS P3 common-view CCD results could be treated as the integrated GPS P3 total delay variance contributed by PTBB+PTBG and TWTF+TWT1. Fig. 3 showed the peak to peak total delay variance of PTBB+PTBG was about 1.4 ns from MJD 56744 to MJD 57100 (with standard deviation 0.258 ns); and the peak to peak total delay variance of TWTF+TWT1 was also about 1.4 ns from MJD 56763 to MJD 57100 (with standard deviation 0.285 ns). If any Z12T receiver in each pair contributed the same non-stationary noise, we could infer that the standard deviation of the total delay variance of PTBB was about 0.182 ns over the period from MJD 56744 to MJD 57109, and TWTF was about 0.202 ns from MJD 56763 to MJD 57109.

V. CONCLUSION

In the derivation of equation (1), we found the formation of the link delay DlyL(PTBB, TWTF) was composed of uncompensated CCD measurements, cable delay measurement, and other non-stationary delay variation of GNSS stations. The uncertainty type of the GPS CCD measurements (the first and second line of (1)) was statistical uncertainty and about 0.1~0.3 ns [4]. The uncertainty of 1 PPS cable delay measurement (the third line of (1)) was mainly dominated by the systematic uncertainty of time interval counter, about 0.2 ns for this calibration [2][9]. Those 2 uncertainty budgets were determined at the calibration stages.

The non-stationary delay variation term ΔDlyR(StdB, t1, t2) was the total delay change of BIPM travelling calibrator StdB (BP1C and BP0U in this tour) from PTB to TL in this calibration period. The PTP-PTB closure test done by Jiang et al showed the change of total delay of BP1C/BP0U were about 0.03 ns over 445 days. Our preliminary moving cesium clock test showed the total delay changes of BP1C/BP0U in 2 weeks over 25 km baseline were fewer than 1.13 ns.

For ΔDlyR(PTBB, t1, t) and ΔDlyR(TWTF, t2, t), they were the total delay variations of fixed stations PTBB and TWTF at anytime t respect to their calibration time t1 and t2. After monitored the GPS P3 common view CCD of PTBB and TWTF with their local Z12T receivers, the results standard deviation was 0.182 ns (PTBB) and 0.202 ns (TWTF). Combined the uncertainty budgets listed in (1) and ignore the baseline of PTBB-TWTF link was more than 10000 km, the uncertainty of the link delay PTBB-TWTF was less than 1.218 ns including the long-term delay variations of antennae, cables, receivers, the incoherence of 20 MHz and 1 PPS at TL, and other unknown long-term effects.

From Fig. 3, we could infer that the calibration results of DlyL(PTBB, TWTF) and DlyR(TWTF) would be influenced by the time when we process the GPS CCD comparison. In an extreme case for the link delay DlyL(PTBB, TWTF): the calibration result of a tour held during MJD 56750~56790 may have up to 0.5 ns difference with the tour held in MJD 56990~57040.

In Fig. 3, we also found the total delay variations of PTBB and TWTF, or the link delay DlyL(PTBB, TWTF), were changed with time. The time deviation of GPS P3 CCD of 2 Z12Ts at PTB was about 0.126 ns at average time = 90 days, and was about 0.114 ns for 2 Z12Ts at TL at the same average time (Fig. 4). We estimated the DlyL(PTBB, TWTF) would change about 0.170 ns at 90 days after its last calibration. We guess this value would increase with the un-calibrated period, but it needs more long-term GPS common view CCD monitoring.

![Fig. 3. The GPS P3 common view CCD of PTBG-PTBB (blue dot) and TWTF-TWT1 (green dot), both peak to peak difference was about 1.4 ns](image)

![Fig. 4. The Time deviation of CCD of Z12Ts at PTB (blue dot) and TL (purple dot), green line: average time = 30 days, blue line: average time = 90 days](image)
Lack of portable hydrogen maser, we can’t execute the moving hydrogen maser time comparison test, but this study was still helpful for clarifying the uncertainty composition of the current PTB-TL link and could reduce the total uncertainty of the link delay $D\text{ly}(\text{PTBB}, \text{TWTF})$ in a reasonable limitation.

ACKNOWLEDGMENT

The authors would like to thank the colleagues of BIPM and PTB for the cooperation of the time transfer experiments and sharing the data.

REFERENCES


[12] Lewandowski W., Tisserand L., Determination of the differential time corrections for GPS time equipment located at the OP, NTSC, HKO, TL, SG, AUS, KRIS, NMIJ, and NICT, Rapport BIPM-2008/02, 27 pp


Link calibration or receiver calibration for accurate time transfer?

Z Jiang
Time Department
Bureau International des Poids et Mesures (BIPM), zjiang@bipm.org

Abstract—In almost all the studies, the differential receiver calibration and the link calibration had been discussed separately as if the two calibrations were completely independent. In fact, in the sense of the total delay for UTC time transfer, the difference between the receiver and link calibrations is not how to perform the calibration measurement but how to use the measurement data. The two calibration results are convertible to each other under certain condition. We discuss the features, advantages and disadvantages of the link and receiver calibrations, their uncertainties, and in particular, their applications in the computation of [UTC-UTC(\(k\))].

Keywords— time transfer, uncertainty, calibration, link calibration, receiver calibration

Notation

- TWSTFT: Two-Way Satellite Time and Frequency Transfer
- TWOTT: Two-Way Optical fibre Time Transfer
- Link: A time link is a clock comparison result using a particular technique, e.g., a P3 link is the clock comparison using the GPS P3 code technique. Similar, we have the link of GPS/GLONASS C/A, C1, P3, PPP, TWSTFT a TWOTT a UTC time link at present is a time link between Lab(\(k\)) and PTB
- Tour: A calibration tour is a round trip calibration campaign with start and closure measurements. It may include several laboratories and is usually not longer than 6 months
- StdB: The BIPM standard GNSS travelling station/calibrator. It consists of two GNSS systems. Each is built mainly with a receiver, an antenna, an antenna cable, connecting cables and is equipped with a PPS and a frequency distributor. It is equipped with a time interval counter (TIC) and a pre-cabled black box calibrator with unknown but constant sub- and total-delays during a calibration tour
- UTCp: The UTC(\(k\)) point at Lab(\(k\)). Here the \(k\) stands for a UTC laboratory to be calibrated, denoted as Lab(\(k\)) in the paper
- Total Delay: The total electrical delay from the antenna phase center (APC) to the UTCp including all the devices/cables that the satellite and clock signals pass through. It numerically equals to the sum of all the sub-delays (Fig. 2). The total delay uncertainty dominates directly the UTC time transfer uncertainty
- METODE: Measurement of Total Delay; the BIPM calibration scheme is composed of related methods and equipment (StdB) for the calibration of the UTC time links. It is developed in the frame of the BIPM pilot project aiming at improving the UTC time transfer calibrations

\(C_M\): The METODE total delay correction. In the present UTC network, a GNSS METODE time link correction is equal to the classic GNSS equipment calibration correction if the both take the PTB as reference

\(u_0\): Total uncertainty of the total delay correction \(C_M\)

\(u_A, u_B\): type A and type B (calibration) uncertainties (1-\(\sigma\))

CCD: Difference of two system’s data that have a common clock

DCD: Double difference of two independent measurements of clock differences, for example the DCD of TWSTFT and GPS links.

I. INTRODUCTION

Calibration is essential for accurate time transfer. The so far time transfer calibration techniques used for clock comparisons in UTC computation are as follows:

(I) Receiver or Equipment differential calibrations [1] for GNSS time transfer. The conventional \(u_0\) in Circular T is 5 ns at present [8]; Here the calibration result is of the internal delays INTDLY(P1,P2). In the coming future, the \(u_0\) will be reduced by a factor of 2 given by the ongoing new BIPM GNSS equipment guidelines [2].

(Ia) TW-Link calibration – used mainly for TWSTTT with a TW mobile ground station [3-5,21], of which the state-of-the-art of the Type B uncertainty (\(u_0\)) is 0.6 ns to 0.8 ns [4,5], the conventional \(u_0\) used in Circular T value is 1 ns [8,21]. TW-Link calibration can also be used for the GNSS time link calibration. The first such calibrated GPS UTC time link was the BEV-PTB using the TUG TW mobile station performed in 2008 and was applied in the UTC computation since then [6,7,8], of which, the \(u_0\) is 3 ns and is still the smallest one in all the UTC GNSS time links in the latest BIPM Circular T [8].

(Ib) GPS-Link calibration – similar as (Ia) but using the GPS mobile calibrator, e.g. the StdB. Unlike the technique (I), we calibrate here a TW or a GNSS time link. Many authors studied since years the GPS link calibrations. The latest developments [9-20] suggest that the \(u_0\) can be attained to 1 to 1.5 ns. This is a great achievement which suggests that link calibrations can bring an improvement in the uncertainty of time links in Circular T by a factor of about 2 or 3 vs. the 5 ns at present [8], with a positive equivalent impact on the uncertainty of [UTC-UTC(\(k\))].

The new TWSTFT Calibration Guidelines for UTC Time Links [21] approves the use of the GPS technique for the calibration of the UTC TWSTFT time links: “A TWSTFT link calibration campaign is carried out using a TWSTFT mobile station and/or a GPS travelling system that is circulated among several time laboratories contributing to UTC’’ … “It can be used as a supplement or an alternative when the TWSTFT mobile station is not applicable … Unlike differential
GPS receiver calibration using the CCD of the two GPS receivers (the travelling and the local ones), the DCD is used in this case. It is obtained by the differences between the TWSTFT and GPS time links. The local GPS receiver of Lab(k) is not involved. Based on ([9-17,20]), the combined calibration uncertainty of 1.5 ns is attainable’

An advantage of the link calibration is to be able to calibrate all types of time transfer techniques over the same baseline for UTC time transfer, e.g., a physical TWSTFT link calibration can be used to calibrate the respective GPS link (IIa), and vice versa (IIb), as discussed above.

III) Time link alignment. It is to align a, usually non-calibrated, time link to another, usually a calibrated, link. Such we transfer a link calibration from one to another. The uncertainty of an alignment is the combined uncertainty of that of the two links on question. An alignment is not a real metrological calibration although it is a routine operation of the UTC and Circular T computation at the BIPM. There are about 40 link alignments during the monthly UTC computation.

In the following section II, we will discuss in which condition the receiver and link calibration are different or identical; Section III shows the advantages and disadvantages of the link and receiver calibrations, the conversion method, their uncertainties, and in particular their application to the computation of \([UTC-UTC(k)]\). Section IV gives a numerical uncertainty analysis by a comparison of the GPS and TW link calibration results. Section V is the summary.

II. RECEIVER OR LINK CALIBRATIONS?

Very often, e.g., [1,3,5,6,7,9-11], the receiver calibration and the link calibration had been discussed separately as if the two calibrations were completely independent. Really, even the calibration results are defined differently. For the first, it is the INTDL(P1,P2) of an individual GPS geodesic receiver [1,2]; precisely, it is the sum of so called antenna and receiver internal delays. For the second, as a time link calibration, it is the total delay difference (cf. the Notation) of two remote receiver systems [13,24], cf. Figures 1 and 2.

To make sense the numerical comparison of the two calibration results, we define:

1) Both calibrations are used for UTC time transfer, i.e. to determine the UTC-UTC(k) of Circular T;
2) Both calibration are computed based on the same reference system; for the UTC time transfer, the simplest and in fact the real physically meaningful reference used in Circular T, is the pivot laboratory PTB of the UTC network;
3) To be simple, the calibration measurement is the total delay of a GPS receiver system (Fig. 2).

Given above definitions, the receiver and link calibration results are directly comparable and it is easy to prove that the difference between the receiver and link calibrations is not how to perform the calibration measurement but how to use the measurement data, e.g. the reference. The two calibration results are convertible to each other under above definitions.

Based on the theory developed in [1], the [7] investigated the mathematical detail of the relations between the absolute, differential receiver calibrations and the link calibrations. It is proven that, not only the receiver and link calibration can be converted between each other but also in a link calibrated network, the absolute calibration result can be obtained if there is at least one receiver absolutely calibrated in the network. If there are more than one receivers are absolutely calibrated, a least square network adjustment is required, of which the mathematic model is used widely in many fields, such as the precise leveling or gravimetric network adjustments, cf. e.g. [22].

Now let us analyze a concrete example (Fig. 1) of the recently carried out calibrations tour of BIPM, PTB, NIST and USNO [16,18,19].

Time transfer is to compare two remote clocks, \(C_A\) and \(C_B\). In this sense, if we take one of the clock \(C_A\) as the reference, i.e., set the total delay of the \(C_A\) to be zero, the link and receiver calibrations give the same clock comparison result (Fig. 1). Now we compare a second pair of the clocks \(C_C\) and \(C_C\). The same, we take the \(C_A\) as reference. So on so forth, we have the third, fourth, ..., \(N\)th pair of clocks to compare. The reference clock \(C_A\) makes the role of the pivot of the time transfer network as the pivot of the UTC network, PTB. It is easy to understand, it is the optical design of the calibration scheme; because the total delay correction of \(C_A\), i.e. that of PTB, is set to zero, we do not need to use the calibration value of \(Rcv(PTB)\) in the calibration and therefore its uncertainty will not at all impact the link calibration. Simply speaking, only the uncertainty of one receiver, \(Rcv(k)\), will be introduced in the link. Obviously, the impact of the uncertainty of a link to the UTC-UTC(k) is always less than that of the two calibrated receivers, \(Rcv(k)\) and \(Rcv(PTB)\).

With the help of the Fig. 1, a concrete numerical example, we can prove if taking the PTB as reference, the link and receiver calibration results are numerically identical in the sense of the total delay. This is true and unique not only for the direct links USNO-PTB and NIST-PTB but also holds for the indirect link USNO-NIST.

Assume \(c(k)\) is the differential receiver calibration correction for the three receivers: c(PTB), c(USNO) and c(NIST), see Fig. 1. \(C'(k-PTB)\) is the link calibration corrections: \(C'(PTB-PTB)\), \(C'(USNO-PTB)\) and \(C'(NIST-PTB)\). As defined above, the link calibration correction of PTB is set to zero \(C'(PTB-PTB)=C'(PTB)=0\). We have \(C'(USNO-PTB)=C'(USNO)\) and \(C'(NIST-PTB)=C'(NIST)\).

For the third, the non-UTC link USNO-NIST, we have,

\[
C'(USNO)-C'(NIST) = [c(USNO)-c(PTB)]-[c(NIST)-c(PTB)] = c(USNO)-c(NIST)
\]

This suggests that because PTB is the reference, the clock comparison result of any pair of PTB, USNO and NIST, either receiver or link calibrations, is numerically the same in the
sense of the total delay, including the time link USNO-NIST which has not been calibrated as a link.

Mathematically, the sum of the three link corrections in the triangle of PTB-NIST-USNO must be zero.

\[
C'_{(USNO)} - C'_{(PTB)} - C'_{(NIST)} = 0
\]

Fig. 1. Taking the same reference (PTB), link and receiver calibrated clock comparison results are identical.

**III. UNCERTAINTIES OF RECEIVER AND LINK CALIBRATIONS**

The theory of the uncertainty propaganda from time transfer to the UTC-UTC\((k)\) is derived from [23]. The study [27] furtherly proves that 98% of the uncertainty in UTC-UTC\((k)\) comes from the time transfer link Lab\((k)\)-PTB. The uncertainties of the UTC-UTC\((k)\) given in the Section 1 of BIPM Circular T depend almost completely on that of the UTC time link given in the Section 6 [8] but that of the receiver calibration. In this sense, we can say that the receiver calibration results are never used in the UTC time transfers.

Therefore the optical design of the calibration of the UTC time transfers should be aiming at minimizing the \(u_\text{UTC-UTC}(k)\) in UTC-UTC\((k)\)-PTB. In other words, the absolute uncertainty of a receiver or equipment or a system does not affect the uncertainty of UTC-UTC\((k)\). The absolute errors, or biases or systematics are completely cancelled when composing the UTC link Lab\((k)\)-PTB.

Before further discussion, first a question, what is the total delay and why we need it?

Fig. 2 is the plot of the setup at the UTC Lab(PL) in Poland. The total delay is physically composed of at least seven sub-delays, of which five sub-delays are measured by a time interval counter (TIC). A TIC measurement error is 0.5 ns given by the manufacture notice [25]. The combined uncertainty introduced by the TIC measurements in the total delay is therefore 0.5 ns \(\times \sqrt{5} = 1.2\) ns; this does not include yet the uncertainties in the so-called internal delays (INTDLY) which cannot be measured by a TIC. Easily to understand, it is the total delay that is really used in the time transfer and the optimal method to determine the total delay is to measure it as a whole instead of measuring every sub-delay (by a TIC), then add them and the INTDLY (L1,L2) together.

To give a numerical idea about the uncertainty of a TIC measurement, we take a rigorous experiment that was carried out under the EUROMET project 828. It was a metrological key comparison, namely, EUROMET TF.T1-K1: Comparison of time interval (cable delay) measurement [26]. A travelling cable standard equipped with three cables visited 28 EUROMET laboratories organized by the BEV during 2005-2006. The national laboratories performed the measurements according to their proper norms and technical specifications as they do to maintain their UTC facilities. Table I lists the result of the delays of the cable #3 measured by the 28 participants. The peak to peak variation is 1.8 ns. With respect to the mean values, there are 10 residuals of the measures \(\geq 0.5\) ns and 3 residuals \(> 1.0\) ns. This agrees with the estimation of the SR620 manufacture, i.e. 0.5 ns.

![Fig. 2. Total delays (Red for PL and Blue for BIPM equipment) between the GPS antenna phase centre (APC) and the UTC (PL)](image)

**TABLE I. THE DELAYS OF THE TRAVELING CABLE STANDARD MEASURED BY THE 28 UTC LABORATORIES IN THE KEY COMPARISON**

If the uncertainty of a key comparison in Europe is like this, can that of a routine delay measurement in a developing country be better? If the uncertainty of a single cable is like this, what will happen to the sum of 5 or more sub-delays? This kind of question can be continued to ask and the answers should make every scientist in metrology worry about, except
for the one that says, let us use the total delay. More sub-delays are measured, more extra uncertainties are introduced.

The classical receiver calibration requires measuring the sub-delays using the TIC and the link calibration does not. In practice, the above sub-delays are merged into three major parts: the internal delays, the antenna delay and the reference delay, namely IntDly(P1,P2), AntDly and RefDly, noting that the first two are precisely the delays of the GPS signals while the last is delay of the local clock signal.

The internal delays are obtained by subtracting all other sub-delays from the total delays. Above, we have fully proved, this subtraction may introduce up to nanoseconds of extra uncertainties due to the sub-delay measurements.

Rigorously speaking, the internal delays, IntDly(P1,P2), are that of the GPS L1/L2 signals but any laboratory PPS signals. The delays of GPS L1 and L2 are electronically different each other and the delays of GPS and lab signals are different either. The usual methods to measure them and to estimate their uncertainties are doubtful.

Deeper study shows, the definition of the INTDLY is physically not rigorous and practically unmeasurable in usual cases. It assumes that the satellite signal delays of L1/L2 in the antenna cable are equal to each other and identical to that measured by a TIC. This may produce at least 3 ns error, cf. the Table II for the cables of 9-12 m (cf. [15] for more details), knowing that an antenna cable is usually 40-50 m. Obviously, such defined, measured and obtained INTDLY(P1,P2) are erroneous. The u₀ of the INTDLY is not better than 3 ns;

<table>
<thead>
<tr>
<th>Cable</th>
<th>TIC /ns</th>
<th>TIC-P1 /ns</th>
<th>TIC-P2 /ns</th>
<th>TIC-P3 /ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 m</td>
<td>15.7</td>
<td>1.1</td>
<td>1.0</td>
<td>1.2</td>
</tr>
<tr>
<td>6 m</td>
<td>31.2</td>
<td>1.4</td>
<td>1.3</td>
<td>1.6</td>
</tr>
<tr>
<td>9 m</td>
<td>47.4</td>
<td>2.8</td>
<td>2.7</td>
<td>3.0</td>
</tr>
<tr>
<td>12 m</td>
<td>62.6</td>
<td>2.9</td>
<td>3.0</td>
<td>2.6</td>
</tr>
</tbody>
</table>

The doubts above do not exist in the link calibration. The BIPM calibrator StdB measuring the total delay is a black PPP box without being calibrated. All the sub-delays are unknown and not applied. Hens, their uncertainties do not affect the UTC time transfer and the calibrations.

Now we can answer the above question: why we need the total delay? Compared to the usual sub-delay measurements using a TIC, the METODE total delay measurement measured by the GPS calibrator obtains the required GPS signal delays of the two different frequencies (instead of the local laboratory PPS single frequency signal delay) without extra uncertainty introduced due to the sub-delay measurement uncertainties propaganda and the discrepancy between the frequencies required and measured. The uncertainty in INTDLY(P1,P2) may be tripled with respect to that of the total delay.

Below, let us discuss the uncertainties of a time link calibration taking example of the experiences of the BIPM METODE pilot project [13,14], using the StdB equipped with double systems; Here each system can perform a calibration without sharing any common part with the other. It is best to have at least two receivers of different types. This may increase the measurement discrepancies but improves the uncertainly computation as well as the robustness of the calibration result.

The total uncertainty (U_M) is corresponding to the total delay correction (CM). Unlike the traditional estimations [1,2,9-12], we use the completely different methods [24] to evaluate the U_M. Here, not any geodesic hypothesis is needed. We applied six more accurate and independent tools (Table III):

<table>
<thead>
<tr>
<th># Source of the sub-uncertainties</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Distance determination [24]</td>
<td>≤ 0.1</td>
</tr>
<tr>
<td>2 Comparing to the TIC [15,24]</td>
<td>≤ 0.5</td>
</tr>
<tr>
<td>3 Testing over the 0–100 m short baselines [24]</td>
<td>≤ 0.5</td>
</tr>
<tr>
<td>4 Comparing to the portable Cs standard [15]</td>
<td>≤ 1</td>
</tr>
<tr>
<td>5 Comparing to the TWSTFT mobile calibration station (2013-2014 campaigns, Table V)</td>
<td>≤ 1</td>
</tr>
<tr>
<td>6 Comparing to the 420 km TWOTT [17]</td>
<td>≤ 0.2</td>
</tr>
</tbody>
</table>

We compared the METODE results to that of the above six techniques and it is reasonable that the root mean squares (RMS) of the discrepancies covers the U_M, i.e., U_M ≤ RMS. The U_M of the C_M is evaluated to be composed of the sub-uncertainty given by the Table IV [24].

<table>
<thead>
<tr>
<th># Means used to evaluate the uncertainty</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PPP Measurement uncertainty (u_A) of StdB-UTC(PL)</td>
<td>0.1–0.3</td>
</tr>
<tr>
<td>2 PPP Measurement uncertainty (u_A) of UTC(k)-UTC(PTB)</td>
<td>0.1–0.3</td>
</tr>
<tr>
<td>3 TWSTFT uncertainty (u_k) of UTC(k)-UTC(PTB)</td>
<td>0.2–0.5</td>
</tr>
<tr>
<td>4 Instability and the sub-delay measurement uncertainty of the reference at Lab(k)</td>
<td>0.5–0.7</td>
</tr>
<tr>
<td>5 Instability of the traveling receivers</td>
<td>0.5–1.0</td>
</tr>
<tr>
<td>6 Others</td>
<td>0.3–0.6</td>
</tr>
<tr>
<td>Total</td>
<td>0.8–1.5</td>
</tr>
</tbody>
</table>

The U_M as estimated from the root sum square of these errors is hence (0.8–1.5) ns (1σ).

If only one GPS receiver in the calibrator is used, the instability would be factor of \sqrt{2} higher, u_A is about \sqrt{2} x (0.8–1.5) ns or 1.1–2.1 ns.

Other independent studies [9-12,15,20] with completely different method prove that the calibration uncertainty of 1.5 ns or even 1 ns is attainable.

IV. DISCUSSION

By two numerical examples we show the advantages and disadvantages of the receiver and link methods: the first is of TW and the second is of GPS.

We take the METODE experiences as an example. The later was proposed in the frame of the BIPM pilot project (2011-2014) aiming at unifying the UTC time link calibration within an uncertainty ≤ 2 ns [13], performed according to the UTC time link calibration guidelines [28]. METODE is composed of a time link calibration scheme with the calibrator denoted StdB. Since last two years, it has visited 11 UTC laboratories as shown in the Fig. 3. Here AOS includes the PL
in Poland and NICT includes NMIJ in Japan. Each calibration tour contains maximum 2 laboratories starting and closing at BIPM. The PTB has been visited twice times for verify the stability of the StdB. It is proved within 0.5 ns on average of the two receivers and vs. the BIPM fixed references. This shows its stability. In recent years, there are other calibrations where both the TW and GPS link calibration techniques were used, such as the EURAMET 1156 [20] and that of the TGVF-FOC [5].

The final result of the calibration are the internal delays and this 5.7 ns correction is forced to correct the IntDly(P1,P2) although we know clearly it was not the IntDly but the RefDly that is wrong. This explains partially why the uncertainty of IntDly is always bigger than that of the TotDly. Do not forget that the error shown in Table I is not taken into account here.

**TABLE VI. TOTAL DELAY CORRECTION VS. BIPM LAST CALIBRATION**

<table>
<thead>
<tr>
<th>Lab(k)</th>
<th>Rev/Link</th>
<th>TotDly C_m/ns</th>
<th>uB/ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMIJ</td>
<td>NM0C/GPSZ12T</td>
<td>-3.2</td>
<td>1.5</td>
</tr>
<tr>
<td>NICT</td>
<td>SepB/GPSPolAxt2</td>
<td>3.0</td>
<td>1.5</td>
</tr>
<tr>
<td>TL</td>
<td>TWTF/GPSZ12T</td>
<td>5.7</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**TABLE VII. TOTAL DELAY CONVERTED TO INTERNAL DELAYS**

<table>
<thead>
<tr>
<th>Lab(k)</th>
<th>Rev</th>
<th>TotDly(P1,P2)/ns</th>
<th>uB/ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMIJ</td>
<td></td>
<td>306.6/318.1</td>
<td>-3</td>
</tr>
<tr>
<td>NICT</td>
<td></td>
<td>210.7/216.2</td>
<td>-3</td>
</tr>
<tr>
<td>TL</td>
<td></td>
<td>304.4/312.1</td>
<td>-3</td>
</tr>
</tbody>
</table>

**V. SUMMARY**

The present GNSS receiver calibration [1,2] is figured by:

1. It is a side by side receiver setups of the travelling GPS calibrator and the local receiver with the DCD (of same data type, e.g. L1C vs. L1C and P1 vs. P1, P3 vs. P3 etc.) as the measurement;
2. The final result of the calibration are the internal delays INTDLY(L1/L2). The reference is the mean value of a few receivers (PTB, MNII and OP for the campaign 2013-2014) of which the uB is 2.5 ns [2];
3. If is applied for individual receivers and universal for the UTC GPS UTC time transfer network;
4. It is independent but may not be consistent with the UTC TW, GPS link calibrations and time transfers Lab(k)-PTB corresponding to UTCUTC(k) in Circular T. They take PTB as the calibration reference.

The BIPM GPS link calibration [13,14,28] is figured by:

1. It is a side by side link setups: the first link is between the travelling GPS calibrator and the PTB master receiver; the second link is any UTC link to be calibrated, such as the TW Lab(k)-PTB, with the DCD as the measurements without the constrain of data types;
2. The final result is the CalR of a TW or a GPS /GLONASS or a TWOTT link. The uB is 1.5 ns as given above, cf. also [14,17,24]. It can be converted to the classical INTDLY(L1,L2) but the uncertainty will be doubled;
3. It is an one-for-all calibration, applicable for individual receivers and universal for the whole UTC time transfer network using, e.g. GPS, GLONASS and TW etc.;
It is independent and consistent with the UTC TW calibrations and the UTC time transfers Lab(k)-PTB corresponding to UTC-UTC(k) in Circular T, both taking PTB as the calibration reference.

Only the measurement noise and the instability of the travelling GPS calibrator will influence the total delay uncertainty of the link calibration. While the biases, the absolute errors, including the sub-delay uncertainties etc. are either cancelled or not used in the METODE calibration. This explains why the uncertainty of a link is smaller than that of the related receiver calibrations.

As theoretically studied in [23] and numerically investigated in [27], at present [8], 98% of the uncertainty in UTC-UTC(k) comes from that of the time links Lab(k)-PTB. The receiver calibration uncertainty impact that of the UTC-UTC(k) through the time link Lab(k)-PTB.

In a typical link calibration, e.g. that of TW, if we replace the TW link by a GNSS or a TWOTT link, it becomes a GNSS or TWOTT link calibration. This calibration becomes a classic differential receiver calibration if the calibrations include the UTC network pivot (PTB), whose absolute calibration error is assumed to be zero. The link total delay calibration correction Ck can be converted to the Internal Delay, INTDLY(L1/L2). This, however, may introduce extra uncertainties.

Acknowledgement

The author thanks his colleagues for their supports to this study:


VI. REFERENCE


[2] BIPM GNSS equipment calibration guideline (Draft) 2015


[21] CCTF WG on TWSTFT and the BIPM (2015) TWSTFT Calibration Guideline for UTC Time Links (draft v2.1), TWSTFT PS meeting during IFCF-EFTF2015, April 2015, Denver, CO. USA


[28] BIPM TM228, BIPM guideline for UTC time link calibration V2.2 draft 2/2014
The performance evaluation of the BD one-way time service

Wei LI  
National Time Service Center/ Key Lab  
of Time-frequency Standard of the  
Chinese Academy of Sciences  
Xi’an, China  
kim_weili@ntsc.ac.cn

Jihai Zhang  
National Time Service Center/ Key Lab  
of Time-frequency Standard of the  
Chinese Academy of Sciences  
Xi’an, China  
zhangntsc@126.com

We GUANG  
National Time Service Center/ Key Lab  
of Time-frequency Standard of the  
Chinese Academy of Sciences  
Xi’an, China  
guangwei@ntsc.ac.cn

Yongliang XU  
National Time Service Center/ Key La-  
boratory of Precision Navigation and  
Timing Technology of the Chinese  
Academy of Sciences, Xi’an, China  
xuyl@ntsc.ac.cn

Zhe Gao  
National Time Service Center/ Key Lab  
of Time-frequency Standard of the  
Chinese Academy of Sciences  
Xi’an, China  
kim_weili@ntsc.ac.cn

Yajing Wei  
National Time Service Center/ Key Lab  
of Time-frequency Standard of the  
Chinese Academy of Sciences  
Xi’an, China  
kim_weili@ntsc.ac.cn

abstract: The capability of one-way time service as an important index of satellite navigation system reflects the ability that a satellite navigation system broadcasts the system time to the clients. In this paper, the performance evaluation method of BD one-way time service is designed standing in the clients’ position. The UTC(NTSC)-BDT result via Space signal reception method is obtained in NTSC, the performance of BDT is evaluated reference to UTC (NTSC). The results of BD CV is chose as the reference to evaluate the precision of BD one-way time service. Calculating the root mean square error of residual, The uncertainty of one-way time service is 3.01 ns, the result shows that the precision of BD one-way time service have higher level.

I. INTRODUCTION

Time service is an important function of satellite navigation system[1]. Navigation satellite broadcasts system time by navigation message and the user gets system time through space signal. In this paper, the performance evaluation method of BD one-way time service is designed standing in the clients’ position. The UTC(NTSC)-BDT result via Space signal reception method is obtained in NTSC, the performance of BDT is evaluated reference to UTC (NTSC). The results of BD CV is chose as the reference to evaluate the precision of BD one-way time service. Calculating the root mean square error of residual, The uncertainty of one-way time service is 3.01 ns, the result shows that the precision of BD one-way time service have higher level.

II. THE BASIC PRINCIPLE OF TESTING AND EVALUATION

A. Basic principle

The basic principle of testing and evaluation is shown as Figure.1.

In the National Time Service Center, using space signal receiving method (one-way timing service) to get the results of UTC(NTSC)-BDT to evaluate the accuracy and stability of the one-way timing service. [2].

B. Calculate the result of one-way time difference

To the satellite i and the local time A of the BD receiver, pseudo-range between them can be expressed as formula (1).

\[ \rho_{i,A} = P_{i,A} + d_{\text{trop}} + d_{\text{ion}} + d_{\text{Sagnac}} + c \Delta t_{i,A} \]  

In this formula,  
\[ \rho_{i,A} \]  
is the pseudo-range observations of the frequency  
\[ L_i \]  
;  
\[ P_{i,A} \]  
is the geometric distance of the satellite to the receiver antenna;  
\[ d_{\text{trop}} \]  
is the tropospheric delay correction;  
\[ d_{\text{ion}} \]  
is the ionospheric delay correction;  
\[ d_{\text{Sagnac}} \]  
is the Sagnac effect correction ;  
\[ c \]  
is the speed of light in vacuum;
\( \Delta t_{i,A} \) is the time difference between satellite i and the local time A, the unknow parameters.

The process calculating the time difference between local time and BDT is shown as Figure.2.

\[ \Delta t = \alpha_i + \alpha_j(t - t_i) \]

Geometric distance calculation

Ionospheric delay correction using dual frequency

Tropospheric delay correction using Hopfield model

The time difference of satellite I and BDT calculation

\[ \Delta t = \alpha_i + \alpha_j(t - t_i) \]

In the formula (1), \( \alpha_i \) can be obtained directly from the original observation file. The receiver antenna center position can be precisely determined. According to the satellite ephemeris and station coordinates, the geometric distance \( P_{i,A} \) between satellite and receiver antenna can be obtained. According to navigation message and satellite ephemeris, a variety of delay correction can be gotten. Finally, though \( \Delta t_{satellite,i} = \text{Clock}_{satellite} - T_A \) we can calculate clock error between satellite clock i and the local time A of the receiver. Then, according to the parameters of the satellite clock model broadcasted, calculating the \( \Delta t_{BDT,A} \) (forecast) = \( BDT_{satellite} - T_A \), this is meaning of time difference between system time and local time obtained by satellite i and the satellite clock parameters model.

C. The remote time comparison links

The remote time comparison methods between UTC(NTSC) and BDT include GPS CV, BD CV and TWSTFT. The comparison result UTC (NTSC) - BDT of GPS CV, BD CV and TWSTFT three links is shown in Figure.3, from August 19, 2013 to September 23, 2013. The residual among the links is shown in Figure.4. The statistical and analysis is shown in table 1.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>BDCV-TW</th>
<th>GPSCV-TW</th>
<th>GPSCV-BDCV</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>0.79 ns</td>
<td>1.83 ns</td>
<td>1.46 ns</td>
</tr>
</tbody>
</table>

III. CALCULATION INDEX METHOD

A. Accuracy and stability of the BD one-way time service

Frequency accuracy is the consistency of the output frequency to standard frequency, it can show correctness of frequency output. The frequency stability is expression that the ability of frequency standards to generate same time or same frequency in a certain period of time, and it describes the fluctuation frequency of output frequency.

The frequency accuracy calculation is shown as formula (2).

\[ A = \frac{\Delta t}{T} = \frac{\bar{x}(t_f) - \bar{x}(t_i)}{t_f - t_i} = \frac{\bar{x}(t + \tau) - \bar{x}(t)}{\tau} \] (2)

In this formula (2), \( \bar{x}(t) \) is time difference data, \( \tau \) is the sampling time, because of the short-term frequency stability of frequency standard is relatively
poor, will have great influence on the frequency accuracy measurement. The accuracy of calculation of the frequency should be selected for a long time as far as possible, in general $\tau \geq 1d$.

The frequency stability is calculated using Allen variance, the formula is,

$$
\sigma_s(\tau) = \frac{1}{\tau} \sqrt{\frac{1}{2(N-2)} \sum_{i=1}^{N-2} (\bar{x}_{i+2} - 2\bar{x}_{i+1} + \bar{x}_i)^2}
$$

$\bar{x}(t)$ is time difference data, $\tau$ is the sampling time, $N$ is the sampling number.

**B. Precision of the BD one-way time service**

BDT based on space signal, we get reference value of local time and BDT though remote time comparison, acquisition all the time comparison data in the test period, make the two make difference, calculate of RMS, evaluation precision of the BD one-way time service,

$$
RMS = \sqrt{\frac{\sum_{i=1}^{n}(x_i - X)^2}{n}}
$$

In this formula, $x_i$ is BDT of each time based on the space signal, $n$ is the total number of measurement data, $X$ is BDT reference value.

**IV. RESULT**

Select a total of more than 5 months of UTC (NTSC) -BDT monitoring results, from June 27, 2013 to November 10, 2013. The reference to the Beidou satellite common view data as the monitoring results of the evaluation, draw the comparison chart of two kinds of data as shown in Figure 5, the deviation between them as shown in Figure 6.

**V. CONCLUSION**

The BD one way time service evaluation is researched in this paper, and in the case of BDT, the evaluation system is established for experiment, the result shows that the precision of BD one-way time service have higher level..

BDS is in the construction stage now, its one way time service capability will improve with the further development of system.

**REFERENCES**

[3] Chinese satellite navigation system management office, the BDS space signal interface control documents [M], 2013, 35
Techniques of antenna cable delay measurement for GPS time transfer

Daniele Rovera∗, Michel Abgrall∗, Pierre Uhrich∗ and Marco Siccardi†
∗LNE-SYRTE, Observatoire de Paris - LNE - CNRS - UPMC, France
Email: daniele.rovera@obspm.fr
†SKK Electronics, Cuneo, Italia

Abstract—We compare the measurement results of antenna cable delay obtained from six different techniques, including some that allow the measurement of the cable installed in-situ. Among these techniques, the reflection of a pulse from the open end of the cable is described and compared with the results obtained with all the other techniques. We obtain a sub-ns consistency between all techniques for three cables of different types, the best technique, based on a Vector Network Analyzer, providing an uncertainty we estimate within 100 ps.

I. INTRODUCTION

In principle, the global delay of a GPS station used for time transfer can be evaluated at the epoch of a relative calibration campaign [1], disregarding the individual delay contributions as the one due to the antenna cable. A relative calibration is performed with a traveling receiver, which needs to be installed at remote station sites, and the calibration results are the global GNSS station delays of visited laboratories. However, subsequent modifications to the remote stations hardware would invalidate the calibration, which might require a repetition of the campaign, and this can prove to be very expensive in either time, money, or logistic. In order to mitigate the issue, the measurement of the antenna cable delay at the time of the campaign might allow for the potential replacement of the antenna cable.

In this paper we compare the measurement results of antenna cable delay obtained by six different techniques, including some that allow the measurement of the cable when installed in-situ. Among these techniques, the reflection of a pulse from the open end of the cable is described and compared with the results obtained with other methods. We used three antenna cables of different types: Cable 1 is a RG58 like, Cable 2 is an Eupen Hyflex, and Cable 3 is an Andrews Heliax†. We provide actual measurement results and we propose an estimation of each related uncertainty.

II. GNSS STATION CALIBRATION

The installation of a high performance GNSS receiver for time transfer requires nowadays several elements. The more important ones are the active antenna, the antenna cable, the GNSS receiver, and the cables that connect it to the local reference signals. Other devices which might belong to the installation are lightening protectors and similar ancillary elements as coaxial adapters. In a typical setup, the antenna is either on the roof of the building or outside in an open field, while the receiver is confined within the laboratory in a more controlled environment. The antenna cable, which connects the antenna to the receiver, is usually 10 m to 50 m long. It introduces significant cable losses, which are compensated by the gain of the active antenna. Fig. 1 shows the attenuation of the three cables under test as a function of the signal frequency.

The calibration is needed either to achieve accurate time transfer with other timing stations using common view of the same GNSS satellite set, or to compare the local time with the time of the GNSS constellation of interest. In a station devoted to time transfer only, it is preferable to perform a "system calibration" instead of an element calibration. We mean for "system calibration" the comparison of the local station to a traveling station, used as a pivot in the frame of a relative calibration campaign which includes a set of remote laboratories. The calibration technique requires to connect the local and the traveling stations in common clock setup, usually expressed by means of a set of 10 MHz and 1 PPS reference signals, and to compare the measurement files produced by the receivers. A less stringent requirement is to have the antennas reasonably co-located. One campaign general requirement would be to operate at different sites the traveling receiver at the same temperature and power levels of the reference signals to improve the common mode cancellation of delays in the pivot. By using this link calibration technique the delays introduced by each element of the local station under calibration are not evaluated, simply because they are not taken into account in the computation of the global delay, whose

†Disclaimer: Product names and model numbers of the equipment are included for reference only. No endorsement or critique is implied.
uncertainty is mainly limited by the stability of the pivot and of the local stations.

\section{Motivation}

Calibration campaign of GNSS receiver station are expensive and not performed so often. It might be interesting and convenient to know the delay of the antenna cable in order to replace it without invalidating the current calibration, and having to wait for the next campaign. During a calibration campaign, if the antenna delay is known, it can be used to evaluate the calibration parameters by attributing to the antenna delay a meaningful value. However, in the case of an antenna cable change, the uncertainty budgeted must be at least reexamined.

Another case for which the knowledge of the GNSS antenna cable is useful can be to solve some issues which are peculiar to the installation site. In this frame, it might not be possible to install the antenna cable of the traveling station, which has then to be connected to a previously installed spare cable, which delay is in principle unknown. The problem is similar to the previous one, except that now the burden is on the balance delays of the traveling station. It must be noticed also that in this case both cables are available at the time of calibration, and so some uncertainty contributors to individual delay might be in common mode, especially if cables are of similar length and type.

The knowledge of the delay of the GNSS antenna cable is mandatory when performing an absolute calibration of the delay of the GNSS receiver installation, where the delay of all the elements of a GNSS station is evaluated off-line and independently.

We expect the system calibration to yield better results than the absolute calibration since it allows to cancel many more sources of uncertainty, in particular the one in common mode between the two stations.

\section{Cost and technical issues}

The cable delay can be measured in a laboratory prior to its installation in order to yield a “calibrated cable”. At the same time many other characteristics of the cable, as attenuation and reflection coefficient can be retrieved. These supplementary information can tell if the cable is in a good condition and, at least, within specifications. This solution is preferable for some means, because one (or more) expensive and accurate machine, read Vector Network Analyzer (VNA), must be available to perform the tasks, especially the ancillary ones.

In other cases the cable installation came first, and its delay must be measured on the field at the time of calibration of the receiver. A VNA might not be available on site, due to prohibitive costs or to logistic issues. The problem might be solved, allowing for less performing but useful measurements, by using far less expensive devices as Time Interval Counter (TIC), which should anyways be available on site to measure the 1 PPS delay of the reference signals.

\section{Transmission and reflection techniques}

The pulse method is based on transmission line theory. When a pulse enters in a cable, if the cable has the same characteristic impedance as the signal generator, no reflection is produced and the voltage at the input connector rises from 0 to half of the available voltage at the output of the generator. When the pulse reaches the open end of the cable it will be reflected back from the impedance mismatch and the voltage at the input of the cable will raise to the maximum available voltage of the generator. By recording the rising edge of the pulses it is then possible to estimate the delay introduced by the cable.

Almost always, the delay measurement of a cable that is already installed on a site can be done by reflection only (RX), while the transmission techniques (TX) are more accurate but can be performed in a laboratory environment only. As an example, if we try to do a TX delay measurement of an already installed cable in transmission with a VNA, we need to calibrate the VNA with a very long test cable in order to reach the antenna emplacement on the roof. It is feasible, but with some complication.

\section{Cable losses and delay stability}

Last but not least, the delay of a coaxial cable changes with frequency because of losses. In principle, delays should be measured at each of the frequencies of the GNSS signals of interest. This is feasible with measurement techniques based on a VNA which allow for such features, but sometime it is not enough.

Because of velocity dispersion, the delay in coaxial cables decreases for increasing of the frequency, but the outcome is almost negligible in the difference between the GPS frequencies L1 and L2. A completely different behavior shows up when the cables are not in adequate condition, especially when one or both connectors are not properly fitted. In this case a huge delay oscillation might show up, depending on the frequency of measurement, with a periodicity related to the overall cable delay. Using some filtering for smoothing reduces the oscillation, but we cannot assume in principle that the GNSS receiver has a similar band-pass profile. Here again, a calibration with a traveling receiver will solve the issue, but leave the burden on the antenna cable delay stability.

In the following, we detail some measurement techniques, and we show the results we obtain by applying them on actual cables.

\section{Measurement techniques}

We performed several kinds of measurements over a set of three cables with N-male connectors, and which have been extended by a N-female to N-female adapter. All cable delay measurements are eventually lead back to this nominal configuration, which allows for direct cable insertion in TX setup.

We first provide here some general remarks on the instrumentation. The pulse distributor generates at its outputs 2.5 V pulses over a 50Ω load. The TIC in TX and RX 1 PPS setup is a Stanford Research SR620, but TX measurements have been also achieved with a high performance STX, which is a laboratory developed event timer [2]. The Vector Network Analyzer is an HP8753C with option -010 to perform Time Domain Reflectometry (TDR). We used VNA acquisition on
For the 1 PPS TX measurement, the hardware setup is made of a pulse source and distributor, and a T adapter. A 1 PPS pulse distributor output is connected by an arbitrary cable to the start input of the TIC, terminated on 50 Ω. Another pulse distributor output is connected by a second arbitrary cable to one arm of the T adapter, which foot is connected to the TIC stop input, terminated on high impedance. The free arm of the T adapter is the setup test port: one side of the cable under test is connected to this port of the T adapter, while the other side stays unconnected.

In Fig. 2 are plotted the cable delays versus $V_{\text{TH}}$ corrected for the reference delay. The sensitivity to $V_{\text{TH}}$ is due to a shape change of the 1 PPS pulses along the cable under measurement: when the test cable is included the rising edge is different from the one of the reference measurement. We decide to use the cable delay for $V = 0.5 \text{ V}$, since it is almost not influenced by the pulse distortion. This last effect might indeed lead to a delay overestimation. Fig. 2 puts in evidence the bias depending on $V_{\text{TH}}$ introduced by this technique. The bias is due to the attenuation of high frequency components and to the cable dispersion. It came out that it is better to use a low $V_{\text{TH}}$ in order to minimize the bias. We estimate that the uncertainty of this technique, when used to estimate the delay of an antenna cable, will never be lower than 500 ps, but can be much worse than this in case of cables with high attenuation.

### B. The 1 PPS RX technique

For the 1 PPS RX reflection measurement, the hardware setup is made of a pulse source and distributor, and a T adapter. A 1 PPS pulse distributor output is connected by an arbitrary cable to the start input of the TIC, terminated on 50 Ω. Another pulse distributor output is connected by a second arbitrary cable to one arm of the T adapter, which foot is connected to the TIC stop input, terminated on high impedance. The free arm of the T adapter is the setup test port: one side of the cable under test is connected to this port of the T adapter, while the other side stays unconnected.

Fig. 3 shows each delay vector measured by changing $V_{\text{TH}}$ from 0.1 V to 5.0 V using 0.1 V steps. The intrinsic delay vector is measured first on the measurement system without any cable connected to the test port. Then the cables under test are connected one by one to the setup test port, and the relative delay vector is measured. Data processing is done by using the threshold value as an index, and subtracting the system delay from each cable delay measurement. The difference is then halved to account for the reflection at the open end. Uncertainty elements of this measurement are the same as in the TX transmission, with some of the effects magnified since

![Fig. 2. Cable delay measured with the 1 PPS TX method with two different TICs. The delays reported here are the differences between the TIC measured value as a function of the threshold level and the value measured with the VNA.](image1)

![Fig. 3. Signals recorded with the 1 PPS RX method. The red curve is the scan of the rise front of the PPS when no cable is connected to the second arm of the T connector. The blue and lilac curve are the scan of the PPS rise front when Cable 1 and Cable 2 are respectively connected to the second arm of the T.](image2)
in this setup the attenuation is doubled. The cable delays versus $V_{TH}$ measured with the SR620 are plotted in Fig. 4 after data post-processing. We arbitrarily decided to use the value for $V_{TH} = 2.8$ V, that is the minimum value for which there is a plateau of the reflected signal, and for which the effect of the pulse distortion is minimized. We estimate that the uncertainty of such delay measurements will be similar to the 1 PPS TX technique in despite of the fact that here there is no evident bias.

C. The $\lambda/2$ resonator RX technique

The RF resonance method is carried out by searching the resonant frequency at which the cable under test, with one end connected to a short-circuit, exhibit a voltage minimum at the other end, the excitation port [3]. The hardware setup is made of a tunable RF source, a voltage zero detector, and an RF T adapter. The output of the tunable RF source is connected by an arbitrary cable to one arm of a T adapter, which foot is connected to the input of the zero detector, set to high impedance. One end of the cable under test is connected to the free arm of the T adapter, while the other end is closed on a short-circuit adapter. Starting from 1 MHz, the source frequency is tuned until a minimum is found, and its frequency is recorded. The procedure is then repeated for a significant number of minima, until it is no longer possible to detect where the minimum is appearing.

The cable attenuation and losses are not only responsible for the velocity dispersion, but also for a deformation of the expected cable impedance. This deformation can be seen on a Smith chart as a displacement of the rotation point from the center and away from the real axis. The data processing consists in calculating the cable delay at each resonant frequency, interpolating or extrapolating the value at the frequency of interest, and potentially correcting for the reference system delays. Uncertainty elements in this measurement typology are again the absence of direct measurement at the frequency of interest in the case of extrapolation, and the estimation of the delay from the center of the T adapter to the test port, as well as the delay introduced by the short-circuit at the end of the test cable. Unwanted reflections along the cable can exhibit an opposite sign on reflected and transmitted signals. In Fig. 5 are shown the delays in ns of the cables under test against the frequency in MHz. For practical reason, we obtain most of the data from an alternative but equivalent method fully based on VNA, where the measurement consists in looking at the Smith chart of the $S_{11}$ reflection of the test cable with a short-circuit at the end. For reference, in the plot are shown measurement on Cable 3 with a TX VNA measurement, and with the $\lambda/2$ resonator. The two methods yield equivalent results, with differences smaller than 30 ps.

We estimate that this technique provides a lower uncertainty than the 1 PPS RX, and this is supported by the fact that there is no evident bias on obtained results. Considering the uncertainty due the flatness of the signal minima, especially at high frequencies, and the uncertainty on the delays of the T adapter and of the short-circuit we estimate for this techniques an uncertainty of 300 ps.

D. The VNA TX Technique

A VNA allows for accurate an phase measurement of the transmitted signal at any GNSS carrier frequency. The proposed technique consists in connecting the cable to the VNA...
The availability of a GPS signal simulator allows measuring the antenna cable delay in transmission, by using two GPS receivers and a GPS common-views (CV) software. Our experimental setup includes, on the GPS signal paths, a GPS simulator, a power splitter for GNSS signals, two GPS receivers, and three coaxial cables, each of them including one or more adapter. One arbitrary cable is used to connect the GPS simulator to the input port of the splitter, while the other cables are used to connect the splitter output ports to both GPS receivers. Note that the short cable connecting the OPM8 receiver to the power splitter allows for the insertion of the test cable on the signal path. Both GPS receivers and the simulator are referenced to the local UTC(k) time scale, as shown in Fig. 7. The test consists in collecting GPS data and processing the collected files with a proprietary CV software, in order to get the P1- and P2-code offsets between OPM8 and OPM7. The reference delay is first evaluated without the test cable, then the cables under test are inserted on the OPM8 arm, one by one. Each measurement takes at least one day in order to reduce the noise.

This technique should be the best suited among all since on one hand the signal source is in common mode and on the other hand the cable delay is measured by using actual GPS signals. The largest uncertainty source should be the power sensitivity of the receiver main units. In practice the noise and the stability of the system must be taken into account. The results obtained show that in case of Cable 2, which exhibits mismatching problems in one connector, gives non consistent results because the difference of the delays for L1 and L2 is about 180 ps. This difference is not justified by the cable characteristics, but is probably due to multiple reflections that occur in different ways for code carried by L1 and code carried by L2. We expect for this technique an uncertainty of 200 ps for well mounted cables.
IV. EXPERIMENTAL RESULTS

The delay of an antenna cable is not a constant. It varies with intrinsic parameters like the signal frequency, as well as with environment conditions. It would be interesting to measure an antenna cable delay for any GNSS signal coupled to any given GNSS receiver, but we focus on the delay for GPS signal frequencies only. All techniques are nevertheless not equivalent. In the 1 PPS techniques, the signal bandwidth is not strictly representative of the GNSS signal, and the method suffers of some signal distortion in the cable due to signal attenuation. In the $\lambda/2$ resonator technique the bandwidth problem is partially solved by a frequency extrapolation, since cable losses limit the resonance at the frequencies of interest. When using the VNA, the diagnostic capability of the measurements is improved, and it is possible to detect potential defects of the test cable in use. We obtain a global consistency over all the measurement techniques of about 0.58 ns for Cable 1, about 0.95 ns for faulty Cable 2 and about 0.09 ns for Cable 3 peak to peak. However, Cable 3 delays where only obtained when using techniques based on VNA measurements. If we restrict the consistency estimation to these techniques only for the other cables too, we obtain about 0.27 ns for Cable 1 and about 0.23 ns for Cable 2. Note that Cable 3 is appearing as the one exhibiting the lowest discrepancy of all three cables. Finally, we consider that the observed peak to peak offsets in Table 1 are in line with our estimated uncertainties for all the measurement techniques.

### TABLE I. CABLE DELAYS FOR VARIOUS MEASUREMENT TECHNIQUES.

<table>
<thead>
<tr>
<th>Method</th>
<th>Mode</th>
<th>Cable 1</th>
<th>Cable 2</th>
<th>Cable 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PPS SR620</td>
<td>TX</td>
<td>158.65</td>
<td>106.25</td>
<td>-</td>
</tr>
<tr>
<td>1 PPS STX</td>
<td>TX</td>
<td>158.39</td>
<td>106.08</td>
<td>-</td>
</tr>
<tr>
<td>1 PPS SR620</td>
<td>RX</td>
<td>158.65</td>
<td>106.84</td>
<td>-</td>
</tr>
<tr>
<td>$\lambda/2$ Resonator L1</td>
<td>RX</td>
<td>158.22</td>
<td>106.11</td>
<td>193.57</td>
</tr>
<tr>
<td>$\lambda/2$ Resonator L2</td>
<td>RX</td>
<td>158.24</td>
<td>106.12</td>
<td>193.58</td>
</tr>
<tr>
<td>VNA L1</td>
<td>TX</td>
<td>158.07</td>
<td>105.90</td>
<td>193.51</td>
</tr>
<tr>
<td>VNA L2</td>
<td>TX</td>
<td>158.07</td>
<td>105.89</td>
<td>193.51</td>
</tr>
<tr>
<td>VNA TDR</td>
<td>TX</td>
<td>158.27</td>
<td>106.08</td>
<td>193.60</td>
</tr>
<tr>
<td>VNA TDR</td>
<td>RX</td>
<td>158.34</td>
<td>106.07</td>
<td>193.60</td>
</tr>
<tr>
<td>Sim L1</td>
<td>TX</td>
<td>158.116</td>
<td>106.079</td>
<td>-</td>
</tr>
<tr>
<td>Sim L2</td>
<td>TX</td>
<td>158.126</td>
<td>105.991</td>
<td>-</td>
</tr>
</tbody>
</table>

V. CONCLUSION

Several measurement techniques of GNSS antenna cable delays have been implemented experimentally on three different cables. The delays produced by all the techniques are consistent at the sub-ns level, even when low cost solution like the 1 PPS measurement with a TIC are in use. Transmission methods are appearing preferable to reflection, but the latter are more easy to implement on an antenna cable already implemented on site. The use of a VNA reduces the discrepancy between the techniques to below 0.3 ns for all three cables under test. A good way to classify such measurement techniques is to consider their ability to achieve measurements using signals similar to the ones broadcast by the GNSS satellites: the more representative they are, the more money they cost. When measuring the delay of poor quality cables, the method based on frequency changes are able to spot any defect, by showing how difficult it is then to define what really is the object of the measurement. The GNSS signal satellite simulator should be the technique that yields the best results. But it remains still subject to the performances of the equipment in use, which might not match all GNSS receivers on site. Clearly, it is highly recommended to use several different techniques for such antenna cable measurements whenever possible.

REFERENCES

Relative Calibration of Galileo Receivers within the Time Validation Facility (TVF)

R. Piriz, D. Rodriguez, P. Roldán
GMV
Madrid, Spain
rpiriz@gmv.com

A. Mudrak
ESA/ESTEC
Noordwijk, The Netherlands

A. Bauch, J. Leute
Physikalisch-Technische Bundesanstalt
38116 Braunschweig, Germany

P. Pánek, A. Kuna
UFE
Prague, Czech Republic

Abstract—GMV is the prime contractor for the Time and Geodetic Validation Facility (TGVF) in the Galileo FOC phase (Full Operational Capability), a contract of the European Space Agency (ESA). Within the TGVF, the Time Validation Facility (TVF) is the subsystem in charge of steering Galileo System Time (GST) to UTC, among other duties. The TVF is operated at GMV headquarters near Madrid, Spain. Calibrated Galileo receivers are needed in the frame of TVF activities to, among other tasks, assess the UTC-GST offset broadcast in the Galileo navigation message. Absolute receiver calibration is a complex activity involving the availability of a signal simulator and of a calibrated reference antenna. An alternative data-based method to evaluate the Galileo receiver delay in E5 signals relative to GPS is also possible. The key feature of such method is the cancellation of the GPS and Galileo ionospheric delays when combining pseudoranges from two satellites with a close position in the sky.

A software tool called *gecal* has been developed by GMV within the Galileo TVF, implementing such method and processing RINEX 3 observation files. Satellite positions are read from a SP3 orbit file. The tool allows the rapid E5 calibration of a new or existing Galileo receiver. This paper describes the tool and the calibration of three Galileo receivers used in the TVF. Galileo Signal-In-Space validation results obtained using such receivers are also presented.

Keywords—Galileo, GNSS, receiver calibration, UTC dissemination, ground infrastructure.

I. INTRODUCTION

The Time Validation Facility (TVF) is the TGVF subsystem in charge of determining and assessing timing-related performances of the Galileo system. In particular the TVF supports the validation of the FOC Galileo timing infrastructure, acts as a preliminary Galileo Time Service Provider (TSP) steering Galileo System Time (GST) to UTC, and coordinates the national timing laboratories participating in the operations. Initially, GMV inherited part of the TVF in use during the IOV phase [1]. For the FOC phase, the TVF has been re-coded and migrated to a new hardware and software infrastructure hosted and operated at GMV headquarters in Tres Cantos near Madrid, Spain. The new TVF started operations in March 2014.

Operational GPS time-transfer is based on pseudorange measurements, normally in two frequencies (L1 and L2) in order to remove the ionospheric delay. Pseudoranges are affected by all the receiver delays from the antenna to the clock reference point. These delays must be calibrated and corrected for in order to refer the pseudoranges to the actual laboratory clock or time reference. The P1-P2 pseudorange combination obtained from the P code signals transmitted on both frequencies is normally used for GPS, and thus P1 and P2 receiver calibrations are needed.

The absolute calibration of a GNSS receiver is a delicate, time consuming, and costly activity. Normally every timing laboratory has in operation at least one well-calibrated GPS timing receiver. The relative calibration of an additional co-located GPS receiver connected to the same antenna (or an antenna nearby) and to the same clock or time reference is relatively straightforward by direct differencing of simultaneous pseudoranges to the same satellite (all effects cancel out except the different receiver delays).

Galileo time-transfer is also possible, with the same calibration requirements described for GPS, but in this case for the Galileo frequencies E1 and E5, where E5 can be E5a, E5b, or E5AltBOC. If a dual GPS+Galileo GNSS receiver is used we can take advantage of the fact that the GPS receiver delays are known or can be easily calculated relative to a co-located calibrated receiver. Another fact to take into account is that since the GPS L1 and the Galileo E1 frequencies are the same (1575.42 MHz) we can consider that the receiver delays in GPS L1 and Galileo E1 are very close to each other, although this depends largely on the receiver model.

II. METHOD OVERVIEW

During the TVF In-Orbit-Validation phase [1] a method to evaluate the Galileo receiver delay in E5 was proposed and developed by the Observatoire Royal de Belgique (ORB) [2], assuming that the GPS delays in L1 and L2 are known and...
that the receiver delay in Galileo E1 is known or is the same as the GPS L1 delay. Differential Code Biases (DCBs) for GPS and Galileo satellites are also assumed to be available. The method is based on pseudorange data processing exclusively, without any need of instrumental intervention at the receiver site. The method description and formulation is fully described in [2] and only the key facts and main equation are summarized and adapted here.

The key feature of the proposed method is the cancellation of the GPS and Galileo ionospheric delays when combining pseudoranges from two satellites with a close position in the sky. If a GPS satellite and a Galileo satellite have a small angular distance between them as seen from the receiver, we can assume that their signals propagate along approximately the same path through the ionosphere, and then the slant Total Electron Content (TEC) along the GPS and Galileo signals are approximately the same.

After re-arranging the basic raw pseudorange equations for GPS and Galileo in two separate frequencies the following equation for the Galileo E5 receiver delay ($DSE_r$) is obtained:

$$DSE_r = \left( (C1G - C2G)K - (C1E - C5E) \right) + \frac{(D1G' - D2G')K - (D1E' - D5E')}{{GPS}_{DCB}} - \frac{(D1G_r - D2G_r)K + D1E}{{Galileo}_{DCB}}$$

(1)

On the right-hand side of (1), the first part (observations) is the combination of GPS (G) and Galileo (E) pseudoranges in two frequencies. This can be obtained from processing receiver observation files in RINEX format. $K$ is a constant related to the GPS and Galileo frequency combination, its value is around 1.25 for the Galileo frequency E5a.

The second part (${GPS}_{DCB}$) is the so-called Differential Code Bias (DCB) for the GPS satellite in the P1-P2 combination. GPS DCBs can be obtained from the GPS navigation message.

The third part (${Galileo}_{DCB}$) is the Differential Code Bias for the Galileo satellite. Galileo DCBs have been calibrated by the satellite manufacturer.

The fourth part (${GPS}_{receiver\_delays}$) is the combination of GPS receiver delays in L1 and L2. They are assumed to be known in a calibrated timing receiver operating in a national metrological laboratory, either because the receiver is already calibrated for the GPS chains or because it can be easily calibrated relative to a co-located receiver.

The fifth and final part ($D1E_r$) is the Galileo receiver delay in E1. Normally it can be assumed that this delay is the same as GPS P1 delay but this depends on the actual receiver model.

Finally, notice that the total noise of the pseudorange combination in (1) is, assuming equal and independent noise levels in GPS and Galileo pseudoranges:

$$\varepsilon_t = \sqrt{K^2 + K^2 + 1 + 1} \cdot \varepsilon = \sqrt{2} \cdot K^2 + 2 \cdot \varepsilon \approx 2.24 \cdot \varepsilon$$

(for the case E5a where $K=1.23$). As a consequence, we are amplifying the raw pseudorange noise by a factor of 2.24. This is important in order to understand the precision of the resulting estimated delay.

### III. Satellite Code Biases

For GPS satellites there are two possible sources of DCBs. One source is the GPS navigation message, that provides the so-called TGD (Timing Group Delay) for each satellite. The broadcast TGD refers to the P1-P2 pseudorange combination and its relationship to the P1-P2 DCB is:

$$DCB = D1G' - D2G' = (1 - \gamma) \cdot TGD$$

where $D1G'$ the GPS satellite delay in P1 code, $D2G'$ is the GPS satellite delay in P2 code, and $\gamma$ is a constant relating the two GPS frequencies. The other source of GPS satellite DCBs are the values estimated by the Center for Orbit Determination in Europe (CODE). However it must be noted that CODE DCBs are not absolute values, they are estimated imposing a zero-mean condition for all the GPS satellites, therefore they cannot be used for receiver calibration purposes. It is then necessary to use DCBs from the GPS navigation message. Nevertheless CODE DCBs can be used for validation purposes in order to evaluate the accuracy/uncertainty of the DCBs from the navigation message. We have observed that typically the agreement between the two sources, after removing the average difference, is quite good with maximum discrepancies of the order of half ns and a global agreement of the order of 0.2 ns (RMS).

Regarding Galileo DCBs, the values estimated by ESA have been provided upon request for the first four In-Orbit-Validation (IOV) Galileo satellites. They are absolute values calibrated on ground before launch.

### IV. Software Tool

A software tool called gecal has been developed by GMV within the Galileo TVF, implementing the ORB method and processing GPS+Galileo RINEX 3 files. Satellite positions are read from a SP3 orbit file. GPS and Galileo satellite DCBs are given in an input file in CODE DCB format.

The tool allows to select the RINEX identifiers of the GPS and Galileo pseudorange pairs to be processed, and also the GNSS antenna phase center coordinates. Two configuration parameters are then used to narrow or widen the search of Galileo/GPS satellite pairs with a common ionospheric path: one parameter for the minimum elevation cutoff, and the other one for the maximum angular distance between the pair of satellites (as seen from the receiver location). Relaxing the two parameters results normally in more satellite pairs found, but also in a noisier and less precise E5 delay estimation. Normally we use 60° as elevation cutoff and 5° as angular separation.
An additional parameter is used as threshold for invalid data rejection, this is useful to discard pseudorange combinations with invalid values, for example in the case some pseudorange in the RINEX file is zero. Finally, a “multiplier of standard deviation” parameter is useful to detect outliers in consecutive pseudorange combination values of the same Galileo/GPS satellite pair (satellite “pass”). An outlier is detected if the difference between the value and the pass average is larger than the pass standard deviation multiplied by this parameter.

To validate the software implementation, we have processed in gecal RINEX data from BRUX station located at ORB. The GPS and Galileo calibration values for this Septentrio PolaRx4 receiver that have been provided by ORB are shown in TABLE I.

We have processed seven days of RINEX data using the gecal tool and taking the provided GPS P1 and P2 delays as input, and we have obtained the E5a and E5b delays shown in TABLE II. and TABLE III., respectively.

The value highlighted in bold (Delay(ns) column) is the Galileo receiver delay in E5 (\(D_{\text{5E}}\)) that we are intending to calibrate, for the different passes (satellite combinations). In normal circumstances the values should be quite similar for the different passes, with variations of the order of ±1 ns. As a final value one can select the average value of all passes, and calculate the standard deviation of all values, as quality check.

As can be seen, the agreement of the gecal estimated values with the E5a and E5b values provided by ORB (see TABLE I.) is 0.4 and 0.2 ns, respectively.

V. THE RECEIVERS

As practical cases of Galileo calibration using gecal, two GPS+/Galileo receivers operated by PTB had to be calibrated in the frame of the Galileo TVF activities: a Septentrio PolaRx4 (GPTB station) and a DICOM GTR51 (PT10 station). The PolaRx4 time-transfer receiver installed at the second Galileo Precise Timing Facility (PTF) located in Oberpfaffenhofen, Germany (PTF2), has also been calibrated using gecal.

The PolaRx4 is at the core of the new Galileo Experimental Sensor Station (GESS) installed at PTB, called GPTB. The GESS network is managed under the TGVF contract. The network consists of fourteen GPS+Galileo receivers worldwide distributed. During the FOC phase all the GESS stations have been upgraded with new PolaRx4 receivers for improved measurement quality. The GPTB station at PTB is new in the FOC phase. Its installation was completed at the end of 2014. Fig. 1 shows the GPTB antenna and the rack hosting the receiver and ancillary equipment. GPTB is the only station in the GESS network that has been fully calibrated for GPS and Galileo signals.

![Fig. 1. The new GPTB station at PTB.](image)

For Septentrio receivers, the manufacturer claims that the receiver GPS P1 and Galileo E1 delays will be close to each other because the carrier frequency is the same, but they will not be identical because their modulation differs. The difference is expected to be at the sub-nanosecond level. This is confirmed by [3] where DLR measured a difference of 0.9 ns between the two delays, using absolute calibration techniques. Thus, for the value of the Galileo E1 receiver delay needed as input for the algorithm we have used the value of the GPS P1 delay. Regarding the GPS calibration of GPTB, its P1 and P2 delays have been computed relative to the PTBB calibrated reference station, which is part of the IGS network,
owned and operated by the German Bundesamt für Kartographie und Geodäsie (BKG). GPTB and PTBB are connected to the same time reference UTC(PTB) with a known time offset between them. Therefore if we correct the GPTB and PTBB pseudoranges by their respective receiver delays (including antenna cable, time reference, and internal delays) and compare them for the same satellites we should obtain very similar values if the calibrations are correct (satellite, propagation, and local effects cancel out except for the different receiver delays). The small geometrical effect due to the slightly different antenna location has been taken into account.

Regarding the GTR51 receiver at PTB (PT10 station), no claim is done by the manufacturer about the similarity of GPS P1 and Galileo E1 delays. Fortunately, it is possible to compare the Galileo E1 pseudoranges with the one of the co-located GPTB station, since both PT10 and GPTB are connected to UTC(PTB). By this method we obtain a Galileo E1 delay for the GTR51 that is 3.2 ns away from its GPS P1 delay. This seems to confirm that not for all receivers the Galileo E1 delay can be considered the same as the GPS P1 delays.

Once the Galileo calibration of PT10 and GPTB using gecal is completed, a final validation is done comparing all GPS and Galileo pseudorange types from the two stations, averaging for all satellites at high elevation (higher that 60°) and for the same days processed in the Galileo calibration. The obtained differences were around 0.3 ns for all pseudorange types, namely GPS P1, P2, and Galileo E1, E5a, and E5b. This means that the calibrations of the two receivers are consistent.

Regarding the PolaRx4 receiver at PTF2, for the value of the Galileo E1 receiver delay needed as input for the algorithm we have used as usual for this type of receiver the same value as the GPS P1 delay. The calibration of the GPS chain of the receiver was carried out by Observatoire de Paris (OP-SYRTE) during a calibration campaign carried out in 2014 in the frame of the TVF activities [4], and involving also the receiver of the first PTF at Fucino, Italy (PTF1), and the receivers at the five UTC(k) laboratories participating in the TVF [5].

VI. PRACTICAL APPLICATION

The UTC-GST offset is disseminated in the Galileo navigation message through the Galileo Signal-In-Space (GALSIS). UTC-GST is typically needed by Galileo users in order to time-stamp their position obtained from the Galileo PVT (Position/Velocity/Time) solution in UTC or local time. Timing applications are using this offset to evaluate the offset of the local timescale versus UTC. The GST-UTC offset is provided to the GMS by the TVF. The disseminated offset value is validated independently by the TVF. The operational value of the broadcast offset is obtained by the TVF through navigation messages from the Galileo Experimental Sensor Station (GESS) network. The PT10 station is used for the validation as described hereafter. PTB daily sends Galileo CGGTTS files from PT10 to the TVF.

Galileo CGGTTS files from PT10 give the station clock offset with respect to GST for each Galileo satellite. The satellite average at each epoch provides the station clock offset versus GST. Since the station has been calibrated, the station clock being UTC(PTB), we obtain UTC(PTB)-GST(GALSIS). We then apply the UTC-GST offset from the GESS navigation messages in order to obtain UTC(GALSIS)-UTC(PTB). An example is shown in Fig. 2.

Another application of calibrated Galileo timing receivers is the usage of Galileo in GNSS time transfer, instead of or in addition to GPS, which is the de-facto standard. Fig. 3 shows an example of Galileo Common-View time-transfer between UTC(PTB) and UTC(ORB) in dual-frequency. PTB is located in Germany and ORB is located in Belgium. As can be seen from Fig. 3, Galileo results are only slightly less accurate than GPS, even with only 4 Galileo satellites in orbit. For comparison, the rapid UTC(PTB)-UTC(ORB) difference as published by BIPM is shown (green dots). Due to the reduced number of satellites, Galileo time-transfer can currently only be used for short-baselines (Common-View), but in the future it is expected to be used also in inter-continental links (All-in-View), with quality similar to GPS or better.
VII. CONCLUSIONS

A software tool called gecal for Galileo receiver calibration relative to GPS has been developed within the TVF project. The tool implements a method previously proposed by ORB and allows the rapid calibration of a new or existing Galileo receiver by processing RINEX 3 measurement files and SP3 orbit files. Three receivers have been successfully calibrated using the described method and tool in the context of TVF operations.

ACKNOWLEDGMENT

The TGVF contract is being carried out under a programme of and funded by the European Union, and managed by ESA. The views expressed herein can in no way be taken to reflect the official opinion of the European Union and/or ESA. We are grateful to P. Defraigne from ORB for having provided freely information and measurement data related to the Galileo receiver calibration method relative to GPS. The calibration of the GTR51 receiver at PTB (PT10 station) was kindly supported by UFE.

REFERENCES

A New Modem for Two Way Satellite Time and Frequency Transfer

*Metrology and Calibration Laboratory, Beijing Institute of Radio Metrology & Measurement, Beijing, China
+Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO, U.S.A.
E-mail: Zhangsk@126.com

Abstract— A new time transfer modem for two-way satellite time and frequency transfer (TWSTFT) has been developed recently at Beijing Institute of Radio Metrology and Measurement (BIRMM). The Direct Sequence Spread Spectrum (DSSS) and Binary Phase-Shift Keying (BPSK) modulations are used to generate a PRN signal. A FFT fast parallel algorithm is applied to achieve fast acquisition of the PRN modulated receiving signal. A 2nd order FLL assisted 3rd order DLL is designed to keep both of the performance of loop dynamic stress and carrier phase tracking accuracy, and a 2nd order DLL is used to track and measure the code phase. A short baseline TWSTFT experiment was done with two 1.2 m VSAT earth stations and a commercial geosynchronous orbit communication satellite to evaluate the modem’s performance. The result shows very low noise with the standard deviation (1 σ) equal to 0.13 ns at a 2.5 MChip/s code rate.

Keywords—two-way satellite time and frequency transfer (TWSTFT); modem; acquisition; phase lock loop; delay lock loop

I. INTRODUCTION

Two-Way Satellite Time and Frequency Transfer (TWSTFT) is a precise time and frequency comparison technique, which is widely used in time metrology [1-2], time synchronization in the applications of satellite navigation, radio ranging and measurement [3], etc. Its performance is better than that of GPS common-view [4-5]. The TWSTFT result has no data boundary discontinuity and we can obtain the results more directly and quickly compared to GPS carrier-phase time transfer [6-7]. Nowadays, many time metrology laboratories in Europe, United States and Asia established TWSTFT links. Time transfer modem is the most crucial instrument in TWSTFT system. The modem’s transmitter unit (Tx) modulates a timing signal (in some case, the one pulse per second (1 PPS) signal) from a reference clock onto the pseudorandom noise (PRN) codes and outputs the Tx signal at intermediate frequency (IF). The IF signal is then up-converted to the radio frequency (RF), amplified, and transmitted to the satellite. The received RF signal is again amplified, down-converted to IF, and demodulated by the modem’s receiver unit (Rx). The time interval between the Tx 1 PPS and Rx 1 PPS is measured on each site. Based on the reciprocity of the TWSTFT signals’ bidirectional paths, the time offset of the two clocks are obtained by exchanging and differencing the time interval measurements at the two sites [8].

Due to the limit of Rx channels in a modem, most TWSTFT operations have to switch between links. That is, a station can only do TWSTFT with up to three stations simultaneously, and the station has to change the Tx and Rx settings for doing TWSTFT with another group of stations. It usually takes from about 40 s to 1 min to change the settings and to lock on a set of new Rx signals. This link switch scheme makes TWSTFT operations less efficient. The time spent on the link switch can be used to extend the measurement period and that may improve the link performance.

To increase the efficiency of TWSTFT operation, a new time transfer modem for TWSTFT has been developed recently at Beijing Institute of Radio Metrology and Measurement (BIRMM). In this instrument, we generate the Tx IF signal by direct sequence spread spectrum (DSSS) and binary phase-shift keying (BPSK) modulation. In order to retain the performance of loop dynamic stress and carrier phase tracking accuracy, a 2nd order frequency lock loop (FLL) assisted 3rd order phase lock loop (PLL) are designed, and a 2nd order delay lock loop (DLL) is used to track and measure the code phase. The Rx of modem is accomplished by a field-programmable gate array (FPGA) and digital signal processing (DSP) based all-digital structure. Using this structure, we can increase the processing ability of the Rx PRN signals. Thus, the Rx channel can be extended easily to about eight or even more without any extensions on the hardware. A fast Fourier transform (FFT) algorithm is achieved for signal acquisition which significantly decreases the time needed for Rx signal search and lock.

In the first phase to evaluate the performance of the modem, a short baseline experiment was done. Two rubidium (Rb) clocks are used as the time and frequency references for two BIRMM modems. A TWSTFT link is established by two 1.2 m dish earth stations using the ChinaSat N10 satellite. The modems measured the time difference of Tx 1 PPS – Rx 1 PPS, where the Tx 1 PPS is derived from the local Rb clock and the Rx 1 PPS is derived from the remote Rb clock plus the delay of the signal path. Meanwhile, a commercial time interval counter (TIC) is used to directly measure the two Tx 1 PPS signals. Then we difference the TWSTFT and TIC measurements, which show quite small instability with standard deviation (1 σ) equal to 0.13 ns at a 2.5 MChip/s code rate.

II. DESIGN OF THE MODEM

A. Structure of the Modem

This paper is partly financed by China Scholarship Council (CSC).
The modem is mainly composed of a modulator unit, an
demodulator unit, an embedded Rb clock unit, a TIC unit, a
frequency distribution amplifier (FDA) unit, a frequency
synthesizer unit, a pulse distribution amplifier (PDA) unit, a 12
times frequency multiplier unit, a Global Positioning System
(GPS) receiver and its antenna, etc. Fig. 1 shows the block
diagram of the modem structure and its inner connections. The
Rb clock is the frequency reference of all the units in the
modem, which is distributed by a FDA. The modulator unit
generates the PRN codes and data to be transmitted, and then
modulates them on a 70 MHz carrier. In case of demodulating,
the 70 MHz Rx signal is initially preprocessed by an automatic
gain controller (AGC) to make the signal voltage level stable.
Then the stabilized signal is digitalized by an analog to digital
(AD) converter in the demodulator unit. After that the code in
the digital signal is acquired, which will be used for the next
tracking process. Signal tracking is the most critical part for the
modem. It’s divided into two loops called the carrier loop and
code loop. The carrier loop tracks the carrier phase and
frequency. The code loop follows the code phase variation.

B. Signal Modulation

Signal modulation is used to generate the PRN direct
sequence spread spectrum signal and to modulate the Tx 1 PPS
and the Tx – Rx data measured by the modem using BPSK.
The logic block is shown in Fig. 2. The embedded computer
sends the data and commands through a serial communication
(COM) port. Then data are decoded and buffed before they are
formed to a whole frame and coded with cyclic redundancy
check (CRC). Each time a 1 PPS signal arrives, a frame is
generated. The frame consists of a header, timestamp, delay
measured by demodulator unit and CRC segment. Each frame
is comprised of 500 bits. After that, the frame is modulated to a
sequence of PRN codes and a transient digital carrier at 15
MHz. Fig. 3 provides the spectrum graph measured by a
spectrum analyzer. Then, it is up-converted and transformed to
a 70 MHz analog signal.

C. FFT Fast Acquisition

Signal acquisition is the first step to start the measurement.
A commonly used method called moving correlator is utilized
to slide the local code phase and to correlate them with the
received code. It will take some time to find the correlation by
this serial search method, especially when the PRN code is
quite long. FFT based acquisition is an all-digital fast algorithm
to acquire the coded signal. Fig. 4 illustrates the flow chart of
the algorithm. The local codes are calculated, transformed to
the frequency domain by FFT, conjugation processed, and then
stored in memory for use in the correlation process. Each time
the acquisition function starts, the Rx digital signal sequence
collected flows to a data buffer and is multiplied by an
estimated local digital carrier to remove the carrier. After
carrier removal, the data sequence is filtered by an
interpolation and decimation filter. Then it is transformed to
the frequency domain by FFT processing. The spectrum data
sequence of Rx signal is then multiplied by the local sequence,
and transformed to the time domain by inverse FFT (IFFT)
process. From the amplitude of the IFFT output, it can be
identified if the correlation is found. If there is no peak large
enough to be identified, the estimated carrier frequency will be
adjusted until a peak appears. In the BIRMM modem, the
frequency adjustment step is 500Hz, and the carrier searching
range is ±10 kHz. This allows the signal to be acquired quickly
even if there is a quite large frequency offset, which is a
common case in many TWSTFT links.
Fig. 4. Flow chart of FFT acquisition algorithm.

Fig. 5 shows a simulation plot of the FFT acquisition. When the code is matching and the estimated carrier frequency is close enough to the received intermediate frequency, a large peak will come out. If the peak is larger than a predetermined threshold, it means that the correlation of Rx signal is found. The corresponding value of the peak on the carrier frequency axis and code phase axis indicates the estimated carrier frequency and code phase, respectively.

D. Carrier and Code Tracking

A carrier tracking loop steers the local reference frequency to follow the phase and frequency of the carrier. A code tracking loop measures the code phase. They are the most delicate parts which affect the modem performance directly. A narrow band 3rd order PLL is designed to track and measure the carrier phase or frequency precisely. However, sometimes, it may lose lock due to interference, an oscillator transient jump, or other dynamic stress. Therefore a 2nd order FLL is also included to assist tracking the carrier. If a large frequency jump occurs, the FLL will soon find the frequency offset and pull the PLL to a new balance.

![Diagram of the 2nd FLL assisted 3rd PLL designed in the BIRMM modem.](image)

A code tracking loop aligns the local code with the restored code signal and measures the phase. Code tracking is fulfilled by a 2nd order DLL in the modem. The code phase discriminator is achieved by a normalized non-coherent early minus late envelope algorithm which is not sensitive to amplitude variation. Figure 7 shows the loop filter structure of the DLL. It is a first order filter with a differential equation as in eq. (1), where \( e(k) \) is the digital sample of code phase error, \( y(k) \) is the output of the filter, and \( \omega_n \) is the natural radian frequency of the DLL that directly affects the loop’s bandwidth.

\[
y(k) = y(k-1) + \frac{(\sqrt{2} \omega_n + \omega_n^2 T)}{K} e(k) - \frac{\sqrt{2} \omega_n}{K} e(k-1) \quad (1)
\]

Fig. 7. Diagram of the 2nd DLL filter structure used in the modem.

III. EXPERIMENTS AND DISCUSSIONS

In order to evaluate the performance of the BIRMM modem, a short baseline experiment was done. The devices used and the cable connections of the experimental system are illustrated as shown in Fig. 8. In the system, two 1.2 m diameter dishes and two Aancom transceivers are used as the earth station (ES). Two BIRMM modems with Rb clocks as references in their interior are connected to the earth stations. Then, the short baseline TWSTFT link was established using the Chinasat N10 commercial communication satellite. The modem makes the Tx – Rx measurements every second. The interval of the two Tx 1 PPS derived by the two internal Rb clocks is obtained from the difference of the two TWSTFT measurements. At the same time, a TIC is used to measure the same interval. The nominal resolution of the TIC measurements is 50 ps. Fig. 9 shows the experimental system on a building roof separated about only several meters. Thus, the difference of time of arrival (TOA), ionosphere delay, Sagnac effect, and some other nonsymmetrical effects in TWSTFT are eliminated.

![Short baseline experimental system diagram.](image)
Theoretical, the results from the TWSTFT and the TIC measurements should be absolutely the same because they are measuring the same physical signals. In practice, the device delay and cable delay make the measurements quite different. The variation of the double differences between the TWSTFT and TIC measurements indicates the performance of the BIRMM modem, because the TIC measurement is simple and should be more stable. Fig. 10 shows the results of the double differences. Two different code rates are tested in the experiments. In case of the 2.5 MChip/s code rate, the double differences have an average value of 4 ns, and have instability with standard deviation (1 $\sigma$) equal to 0.13 ns. When code rate changed to 1 MChip/s, the mean value is almost the same. But the jitter becomes larger with the standard deviation (1 $\sigma$) equal to 0.26 ns.

Currently, the modem collects data at a rate of one point per second, and there isn’t any smoothing process. In fact, those data can be smoothed such as with N seconds. In this circumstance, the data instability will decrease with a factor of $\frac{1}{\sqrt{N}}$. For example, if the data is smoothed with a average time of 100 s, the data jitter of 2.5 MChip/s code will be about 13 ps.

Besides, the FFT acquisition algorithm decreases the link initializing time greatly. Basically, the total time required for acquisition and lock together is far less than 1 s. But because a data exchange is needed in the TWSTFT link, it takes about 5 s to get the first TWSTFT time difference value in the actual test.

ACKNOWLEDGMENT

Special thanks are given to Victor Zhang, Tom Parker and Mike Lombardi at the Time and Frequency Division of NIST for their reviews of this manuscript.

REFERENCES

SASO Time Scale and Measurement Capability

Khalid S. AlDawood
National Measurement and Calibration Center (NMCC)
Saudi Standards, Metrology and Quality Organization (SASO)
Riyadh, Kingdom of Saudi Arabia
k.dawood@saso.gov.sa

SASO Time and Frequency Laboratory is responsible for national time scale generation and dissemination on Saudi Arabia. Recently at SASO NMCC the Time and Frequency Laboratory was developed using primary 5 Cs atomic clocks, 2 GNSS receivers and high technology modern equipments. The national time scale was generated with an uncertainty better than $2 \times 10^{-14}$ and disseminated through the internet for industrial applications using an NTP time dissemination system with an uncertainty less than 50 ms. The fully automatic time and frequency calibration system that was developed is traceable to the national time scale and has the capability for frequency generation and measurement in the DC – 50 GHz range including signal analysis and phase noise measurements. Currently SASO is a member of BIPM Atomic Time Club in order to contribute to the Universal Coordinated Time (UTC) time scale and the realization of the international traceability of the SASO time scale.

EASE OF USE
SASO time scale generated by using 5 Cs atomic clocks with high performance tubes (5071) and calibrated 2 multichannel GNSS receivers (TTS-4). The block diagram of SASO time keeping system is given in Fig. 1. Time signals from the GPS, GLONASS and GALILEO systems are received using 2 GNSS antennas which are installed on the metrology (NMCC) building. The time signals from the antenna are received by a TTS-4 model multichannel GNSS receiver using a 35 m low noise cables. The reference one pulse per second (PPS) time signal from the reference atomic clock is also sent to GNSS receivers and these receivers measures the time difference between satellite clocks and reference SASO Cs atomic clocks in accordance with the BIPM satellite tracing schedule. All cables delay is measured by counter and 50 GHz oscilloscope with an uncertainty less than 100 ps. Cable delay measurement results indicated on the block diagram (Fig.1).

---

Fig.1: Block diagram of SASO time and frequency system

After first steering of Cs clocks time difference between reference Cs clock and satellite clocks is measured by using TTS-4 CNSS receiver in accordance with the BIPM satellite tracing schedule and every day this time difference information is automatically send to BIPM. In additionally time difference between Cs atomic clocks is measured by computer controlled counter trough switch box and results also every day send to BIPM. By this way SASO contributed to generation of UTC and rapid UTC. Result of time difference between UTC and UTC (SASO) regularly published on Circular T and rapid Circular T at BIPM web site. In accordance Circular T information and time difference information between reference Cs clock and other 4 Cs clocks prepared time difference information between UTC and all 5 Cs clocks is presented in Fig. 2. As shown in this graphics reverence Cs clocks (UTC (SASO)) last 100 days deviated approximately 2 ns/day, which correspond to the frequency accuracy about $2 \times 10^{-14}$. All of SASO clocks accuracy was calculated as $<5 \times 10^{-14}$. 

---

978-1-4799-8866-2/15/$31.00$ ©2015 IEEE

254
In SASO, national time scale (UTC(SASO)) was compared with UME and MTC time scale by using GPS common view method (Fig.3) with an uncertainty less than 5 ns.

For statistical analyzing of SASO and UME time scale comparison, standard deviation of data between time difference and its linear fit was calculated and standard deviation was found about 2.3 ns. In additionally for uncertainty evaluation 2 Cs clock in SASO, same time interval compared trough counter and 2 GNSS receiver using GPS common view method. This difference also found less than 5 ns. Currently SASO time scale generated with type A uncertainty 0.7 ns and with type B uncertainty 7.3 ns.
The NTP Time Dissemination System was developed for the distribution of time generated from a Cs atomic clock among local area networks (LANs), wide area networks (WANs), and the internet/intranet by using a network time protocol (NTP) at stratum - 1 level. The Time Dissemination System includes 3 NTP Stratum-1 Servers, a Time Coder and a UPS system. With this system, time dissemination is realized with an uncertainty better than 5 ms for LAN and better than 50 ms for WAN.

At SASO NMCC time and frequency laboratory fully computer controlled automatic time and frequency calibration system in the DC – 50 GHz range including signal analysis and phase noise measurements was developed and verified by manual measurement and used for atomic clock, signal generator, spectrum analyzer, counter and timer calibrations. Time interval measurement with <10 ps uncertainty and high frequency oscilloscope calibration system was developed by using fs laser and 50 GHz oscilloscope system. In additionally, comparison SASO and UME phase noise measurement system using low noise oscillators in progress. Using the time and frequency calibration and measurement system, it is possible to calibrate Cs atomic clocks, Rb atomic frequency standards, quartz frequency standards, signal generators, spectrum analyzers, counters, tachometers, chronometers and timers. The calibration and measurement capabilities of the system are as follows:

- **Frequency Range :**
  - DC – 50 GHz
- **Frequency Measurement Uncertainty (k=2)**
  - DC – 2 GHz (1 x 10\(^{-11}\) x f(Hz))
  - 2 GHz – 50 GHz (1 Hz)
- **Time Interval Measurement :**
  - 0.5 ns – 10\(^{10}\) s
- **Time Interval Measurement Uncertainty (k=2)**
  - 100 ps
- **Amplitude Range :**
  - -100 dBm to +20 dBm
- **Amplitude Measurement Uncertainty (k=2)**
  - 2 dB (-100 dBm to -10 dBm, 30 Hz – 26.5 GHz)
- **Modulation Parameter Measurement**
  - AM ( %0 to %100), FM (1 kHz to 400 kHz), PM (0.1 rad to 10 rad)
- **Phase Noise Measurement**
  - Carrier Frequency Range : 100 kHz - 26.5 GHz
  - Offset Range : 1 Hz - 100 MHz
  - Noise Floor : <-165 dBc/Hz

- 0.2 dB (-10 dBm to +20 dBm, 50 MHz to 40 GHz)
Stability analysis of the French timescale UTC(OP)

M. Abgrall, S. Bize, B. Chupin, J. Guéna, Ph. Laurent, P. Rosenbusch, P. Uhrich, G. D. Rovera
LNE-SYRTE, Observatoire de Paris - LNE - CNRS - UPMC, France
Email: michel.abgrall@obspm.fr

Abstract—This paper presents the current results obtained with the new version of the French timescale UTC(OP) in operation since more than two years now. The time scale is based on an hydrogen maser steered by one of the SYRTE atomic fountains. Thanks to this technique, UTC(OP) is one of the best real time realization of UTC. A statistical analysis of different UTC - UTC(k) comparisons is presented.

I. INTRODUCTION

This paper presents the performances of UTC(OP) generated at LNE-SYRTE, Observatoire de Paris (OP). UTC(OP) is the French real time realization of UTC (Universal Coordinated Time) calculated by the BIPM (Bureau International des Poids et Mesures), and it is also at the basis of the legal time disseminated in France. It is mandatory for this time scale to be as accurate, as stable and as reliable as possible for multiple reasons. First, it is an operational time scale used as a pivot for French contributions to international time scales calculated by the BIPM: French industrial clock data included in Echelle Atomique Libre (EAL) computation and SYRTE frequency standard calibrations contributing to the steering of International Atomic Time (TAI). It is also a time reference provided daily to French laboratories. Second, its performances have an impact on our scientific activities such as testing the performances of the new time transfer techniques and contributing to the space missions GALILEO and ACES (Atomic Clock Ensemble in Space).

We present briefly the current implementation of UTC(OP) before comparing its performances to other UTC(k).

II. UTC(OP) IMPLEMENTATION

The operation of UTC(OP) has been presented in [1], [2]. It takes benefits from the ensemble of clocks and oscillators operated at LNE-SYRTE. This ensemble includes three cold atom fountains [3], [4]: the two primary frequency standards FO1 and FOM and the dual fountain FO2 operated simultaneously with cesium and rubidium atoms. The accuracy of the fountains ranges from $\sim 2 \times 6 \times 10^{-16}$. Practically, the fountains measure the frequency of the ultra stable reference distributed to the different experiments of the laboratory. This reference is produced using an ultra low noise cryogenic sapphire oscillator (CSO), that is phase locked to a hydrogen maser with a time constant of about 1000 s. This allows to take benefits from the short term stability of the CSO (about $2 \times 10^{-15}$ for an averaging time of 1s) and from the low frequency drift of the hydrogen maser (of the order of $10^{-16}/d$).

Thanks to the reliability of the fountains that operate almost continuously, we have been able to implement an automatic software that daily provides calibrations of the maser frequency. This software includes the correction of the frequency shifts based on the automatic monitoring of the fountain parameters, the filtering of possible perturbations on the reference signals generation and distribution, the averaging over periods of 0.1 days, in order to reduce the amount of data, and a final 5σ filtering of possible remaining outliers. This calculation does not provide the final data that are further scrutinized before being sent to the BIPM to participate to the steering of TAI. Nevertheless this automatic process is sufficient to estimate the current frequency of the maser within an uncertainty of $10^{-15}$ or better. This estimation is performed using a linear fit over the past 20 days averaged data. This period has been chosen to be robust against possible interruptions due to periodic liquid helium refill of the CSO or to planned maintenance of the fountain. The calculation is performed for the three fountains, but currently only one is used to calculate the steering. This has been chosen in order to leave more flexibility on the fountain operation that can be interrupted for a planned maintenance or be dedicated to a specific experiment. This process provides the main part of the frequency steering that is applied to a microphase stepper fed by one of the 5 MHz output of the maser and updated daily.

A second small frequency steering is added to this daily calibration is order to remain close to UTC. This coefficient is updated monthly using the data published in the last available Circular T by the BIPM. It is composed of two terms. The first part compensates for the slope of UTC(OP) - UTC. The second part is estimated to reduce by a factor of 2 the departure between UTC(OP) and UTC within the following month. This factor 1/2 has been chosen in order to ensure the stability of the phase lock loop.

With this system implemented since October 2012, UTC(OP) is one of the best real time realization of UTC. It is an autonomous time reference over 1-2 month duration periods, between two updates of the steering towards UTC, relying on LNE-SYRTE facilities only. This notion of autonomy is sometimes forgotten in some laboratories that use other time references available in real time, like GPS time, and a short time constant phase lock loop to generate their UTC(k).

III. RESULTS ANALYSIS

Fig. 1 presents the comparison to UTC of UTC(OP), of UTC(PTB), that is the pivot of the time transfer links between laboratories contributing to international time scales, and of UTC(USNO), that provides the highest number of commercial clock data included in EAL computation by the BIPM. The three time scales are realized using a similar technique based on the output signal of an hydrogen maser steered using atomic fountain calibrations [5], [6], and present equivalent performances.
During the considered period, UTC(OP) (red curve) remained well below 10 ns close to UTC. It was even below 5 ns if we except a short period around MJD 56670. The two other UTC(k) presented also a strong slope around this period. As can be seen in the insert of Fig. 1, this is an improvement of about one order of magnitude compared to the previous system that used a commercial caesium clock manually steered towards UTC. During the past year, the departure between UTC(OP) and UTC remained below 3 ns, reaching closely the uncertainty of the time transfer link calibrations.

Fig. 2 gives the overlapping Allan deviation of the time scale comparisons plotted in Fig. 1. The obtained stabilities are comparable. They vary between $1.2 \times 10^{-15}$ at 5 d observation times and $2 - 3 \times 10^{-16}$ for averaging periods of $2 \times 10^7$ s. We observe bumps for intermediate averaging periods that are partially due to the phase lock loop time constant of the time scales towards UTC.

Fig. 3 presents the time deviation (TDEV) of the same time scale differences. They vary between 200-300 ps at 5 d averaging periods and 1-2 ns for observation periods of $2 \times 10^7$ s.

We have also computed the overlapping Allan deviation of the phase differences, as could be done on data of an arbitrary quantity. This analysis was done with the idea that, contrary to the TDEV, the overlapping Allan deviation is not removing the sensitivity to a slope in the data set. The results are plotted in Fig. 4. The deviations at 5 d averaging periods are a bit higher than for the TDEV, at about 400 ps. They increase proportionally to the observation time ($\propto \tau^{1/2}$) until averaging periods of 1-2 months. This could be interpreted as linear drift of the phase due to the frequency steering of the time scales towards UTC, for periods shorter than the time constant of the loop. We observe also that over this period, the time differences overlapping deviations are closer to each other for the three time scales comparisons than for the TDEV. On the long term, for observation periods of $2 \times 10^7$ s, the deviations reach about 1-2 ns as for the TDEV, indicating the compensation of the phase drift thanks to the steerings towards UTC.

We believe that this type of analysis could also be interesting to access the performances of the time transfer techniques.

**IV. CONCLUSION**

We have described the UTC(OP) performances since October 2012. Over that period, UTC(OP) is one of the best realisations of UTC with performances similar to UTC(PTB) and to UTC(USNO). It is the case despite the fact that other UTC(k) are being improved this last year as can be seen in the BIPM Circular T. UTC(OP) remained a few ns close to UTC, which approaches the uncertainties of the time transfer calibrations.

We are currently working on the improvement of the system hardware by setting up the new generation of microphase stepper [7] at the input of a switch allowing a hot swapping between nominal and backup timescales with a negligible impact.

**REFERENCES**


Fig. 3. TDEV of the time scale comparisons plotted in Fig. 1

Fig. 4. Overlapping Allan deviations using the time scale comparisons plotted in Fig. 1 as the data set.

UTC(PTB) as a fountain-clock based time scale, Metrologia, 49, no. 3, 180-188 (2012).


Preliminary step for a UTC(IT) steering algorithm based on the ITCsF2 primary frequency standard measurements

*Physics Metrology Division, INRIM, Turin, Italy
+ Electronics Department, Politecnico di Torino, Turin, Italy

Abstract—A preliminary test to base the generation of the Italian standard time UTC(IT) also on a primary Cesium fountain has been carried out and the results are reported also in comparison to a steering strategy based only on UTC or rapid UTC data. In addition, the results of the first months of a hardware implementation of the fountain steering algorithm currently in operation in INRIM is also reported. This work represents the first step towards the realization of UTC(IT) with a complete set of primary and commercial frequency standards.

Keywords—Time scale, steering algorithm, primary frequency standard, Cesium fountain;

I. INTRODUCTION

The generation of the Italian reference time scale UTC(IT) is currently based on the signal of a Hydrogen maser which is continuously monitored versus commercial Cesium frequency standards and other Hydrogen masers [1]. Our plan is to base the realization of UTC(IT) on a larger set of frequency standards with a triple generation chain, also based on the frequency measures referenced to a Cesium fountain primary frequency standard, to improve the stability, accuracy, and reliability of the UTC(IT) time scale.

As a first step we have developed, implemented, and tested a new steering algorithm based on the measurements of the ITCsF2 Cesium Fountain developed and operated in INRIM [2,3].

The algorithm implementation is described and the results obtained with almost one year of experimental data are reported and compared with similar results that could have been obtained by steering the same H maser by using only BIPM products, namely UTC and rapid UTC.

In addition, since January 2015, an experimental time scale, named Exp TA(IT), is physically realized by applying the fountain steering to a H maser. The results of the steered time scale are reported and compared with the previous estimates.

The performances on the test period are finally analyzed and discussed and the benefits of using fountain measures in the steering algorithm are pointed out.

II. THE NEW STEERING ALGORITHMS

The current UTC(IT) generation is based on a 5 MHz signal provided by an Active Hydrogen Maser fed to a phase micro stepper (named AOG), that compensate the H maser frequency offset and drift. The steering parameters are obtained from the BIPM evaluation of UTC and rapid UTC and automatically provided to the AOG [1].

In order to improve robustness, reliability, accuracy, and stability of the Italian Time Scale, a new time scale generation system is under development.

The realization will always be based on steered H masers (one master plus a redundant one), but the steering values will be evaluated by three parallel independent steering chains. The first one is based on BIPM rapid UTC data, the second one will be based on an ensemble time scale using INRIM six commercial Cesium standards and four Hydrogen masers, while the third one will use the measures of INRIM Cesium fountain ITCsF2. The ITCsF2 also contributes to the TAI realization by the BIPM.

All input measurement data will be automatically provided. The first step in the analysis will be a pre-processing [4] stage aimed at detecting and filtering possible anomalies on the measurements. Then the steering parameters of the same H maser will be evaluated in parallel by the 3 chains.

Finally, the outputs consistency will be checked by the post-processing block and then a combiner will determine the frequency correction to be provided to the AOGs, to steer the frequency of the Hydrogen Masers and, hence, to generate the future time scale UTC(IT).

The concept of the new steering algorithms is illustrated in the Fig 1.
Fig. 1. Scheme of the new UTC(IT) steering algorithms

III. EXPERIMENTAL DATA WITH ITCsF2 FOUNTAIN

ITCsF2 is the new INRIM Nitrogen cooled Cs Fountain primary frequency standard, that is fully operative since 2013. Its Type B uncertainty is $1.7 \times 10^{-16}$, while its short term stability in the high density regime is $2 \times 10^{-13} \text{e}^{-1/2}$ [2]

ITCsF2 has been used to calibrate TAI providing during 2014 nine frequency evaluations for a total measurement time of 165 days. ITCsF2, together with its twin system NIST-F2, also operating at cryogenic temperature, were used to measure the Black body radiation shift, comparing their frequency with that of the room temperature PFS NIST-F1 [3].

ITCsF2 has operatively measured the frequency of a H maser, namely HM3, for a period of almost one year, from mid February 2014 to March 2015. These primary measures are used in this test to evaluate which could have been the steering of the H maser compared with the steering that could have been obtained basing on UTC or rapid UTC data.

Since HM3 is the H maser currently realizing UTC(IT), the results of the real UTC(IT), based on the current steering strategy, are also compared.

The frequency rates of the HM3 reported in Fig 2 have been measured with respect to 3 different references:

- ITCsF2 measurements (blue)
- UTC (red)
- rapid UTC (green)

The plot represents also the BIPM estimated rates (yellow) as reported in the BIPM web page

Fig. 2. The measures used for the steering test

IV. THE FREQUENCY STEERING EVALUATIONS

The steering algorithm based on ITCsF2 data uses every day a certain number of past daily measurements, typically 30 and compute a frequency correction to be applied to the HM3 through the estimation of a linear fit based on the weighted least squares method.

In case there are few available measures or when the last update is not available, the steering estimation is not updated and the previous frequency steering is maintained, only updated for the effect of the frequency drift previously estimated. This is the reason why in case of missing ITCsF2 measures, the frequency steering could not be optimal. With the final triple chain, in those cases, the steering would be evaluated by another chain, but here we concentrate on the steering evaluated only by a single method.

In Fig 3 the frequency steering corrections computed through ITCsF2 measurements are compared with the correction computed for the same HM3 in the same period (February 2014 – March 2015) with the current steering strategy applied to generate the official UTC(IT) scale.

Fig. 3. Steering corrections estimated from the ITCsF2 measures and the correction estimated by the current UTC(IT) algorithm

To analyze the performances of the new algorithm based on the ITCsF2 measures, we compare such steering results with the ones that we could have obtained by steering the same HM3 using BIPM data. In particular we evaluated the possible steering estimates based only on monthly UTC or weekly rapid UTC measures.

The approach is the same used for the steering through the fountain: every day a frequency correction is estimated by a linear fit on a certain number of previous data. To this aim the following measures have been obtained and used in the fit computation:

$$\text{UTC-HM3} = (\text{UTC-UTC(IT)})_{\text{BIPM}} + (\text{UTC(IT)-HM3})$$

$$\text{UTCr-HM3} = (\text{UTCr-UTC(IT)})_{\text{BIPM}} + (\text{UTC(IT)-HM3})$$

where $(\text{UTC-UTC(IT)})_{\text{BIPM}}$ comes from the BIPM publications and UTC(IT)-HM3 is measured inside the INRIM lab.

In such cases the data latency plays a crucial role as it is one month for UTC or one week for rapid UTC. Consequently
we estimated the new steering parameters once a month and once a week respectively, but considering also that monthly data are available on the 10th of successive month, while weekly data are available on successive Wednesdays.

Fig 4 represents the steering estimates in this case. The steering estimated with UTC monthly data (red) are updated every month, on the 10th of successive month, and they are kept constant (in frequency and drift, therefore giving a straight line) for one month. The steering correction based on rapid UTC are updated every week (green), while the correction estimated with ITCsF2 are updated daily (pale blue), unless measures are not available. It appears that the use of “old” data does not allow to quickly follow possible changes in the clock linear trend, but on the whole all the steering methods are following the trend of the HM3.

To evaluate the effect of such a steering, we estimate the steered rate that could have been obtained by subtracting the steering correction from the H maser rate. The results is a sort of residual rate and it is represented in Fig 5.

From the residual rate we can easily infer by integration what would be the residual time offset gained by such a steered time scale and this is represented in Fig 6 reporting the time scale that would have been obtained by steering the HM3 using ITCsF2 (blue), UTC (red), or rapid UTC (green) steering techniques.

It appears that the steering obtained with ITCsF2 is very accurate in the first period, but then it degrades when the measured are not available (from MJD 56850 till 57000) and the steering is based on old measures. When fresh measures are available starting from MJD57000, the steering is immediately effective. Similarly the steering based on UTC (red points) suffers from being based on old measures not promptly following the HM3 behavior. The steering based on rapid UTC is not the most stable as it suffers important variations from week to week, but on the whole the time accuracy is very good.

For comparison we report in the same plot the behavior of UTC-UTC(IT) (red crosses), official time scale data.

The obtained results give us a good confidence on the steering strategies and their implementation with an automatic software and they show that the planned triple chain would help by using other techniques when the primary steering based on ITCsF2 is not available.

V. EXPERIMENTAL TA(IT)

Since mid January 2015, the steering algorithm based on ITCsF2 measures has been automatically applied to the signal of HM3 to generate a physical steered time scale named TA(IT). The experimental set up for the generation of the official UTC(IT), the experimental TA(IT), and a back up UTC(IT) is represented in the following block diagram, Fig 7, while the devices realizing the time scales are reported in Fig 8.
The behavior of the experimental TA(IT) with respect to UTC(IT), as measured through a Time Interval Counter in the laboratory, and versus rapid UTC and UTC is reported in Fig 10. The apparent “jumps” in the estimate versus rapid UTC visible on MJD 57054 and 57089 are due to the monthly update of rapid UTC. This is in fact visible also in Fig 11 where the differences UTCr-UTC(k) for some time scales are reported and the apparent jumps appear on the same date for all the labs.

We can now check the behavior of the experimental TA(IT) with respect to the evaluation presented before, for the common period of Jan-March 2015. Results are in Fig 10. The blue line represented the time scale that we estimated could have been obtained by steering the HM3 with the ITCsF2. The blue dots in the last period represents was has been experimentally obtained with the Exp TA(IT) physically realizing a steered time scale based on ITCsF2. For comparison purposes this last results are also reported with the addition of an arbitrary constant to show the excellent agreement with the theoretically evaluation we explained before.
VI. CONCLUSION

We have discussed and reported experimental tests on the possibility to steer a H maser based on measures obtained with reference to ITCsF2, UTC, and rapid UTC. These results are a sort of offline evaluation of what could be the time scale behavior if the estimated steering were applied.

In the last months we have also set up a physical realization of an experimental time scale based on a H maser steered by the measures on the Cesium fountain ITCsF2 and we could confirm that the offline estimation are sound as in excellent agreement with what can be experimentally obtained.

Moreover the positive rate of UTC-TA(IT) of about $1 \times 10^{-15}$ observed in February 2015 is in agreement with the rate of IT-CsF2 versus TAI (opposite sign) appearing in Circular T n 326 as

\[
\begin{align*}
\text{IT-CsF2} & \quad 57029 \text{ 57054} \quad d = -1.11(0.5) \times 10^{-15} \\
\text{IT-CsF2} & \quad 57054 \text{ 57074} \quad d = -0.76(0.72) \times 10^{-15}
\end{align*}
\]

This residual rate will be compensated for in the steering algorithm.

We are therefore quite confident on the algorithms and on the software implementation of the steering strategy based on ITCsF2 and we will continue the work to set up a complete redundant steering evaluation chain to realize the national time scale UTC(IT) with a more robust and accurate steering strategy using all the frequency standard available in INRIM.

REFERENCES

Acquisition Method of Loran-C Signal Based on Matched Filter

Yuanyuan Gao∗†, Yu Hua∗, Yuanhong Cao‡, Haifeng Jiang∗
∗National Time Service Center, Chinese Academy of Sciences, Xi’an, China
†University of the Chinese Academy of Science, Beijing, China
‡Sichuan Spaceon Time & Frequency Tech. Co., Ltd, Chengdu, Sichuan, China
E-mail: gaoyy@ntsc.ac.cn

Abstract—A novel acquisition method of the Long-range navigation (Loran)-C signal based on matched filter is proposed to improve the acquisition performance of Loran-C signal under heavy noise environment. Higher signal to noise ratio (SNR) could be obtained through matched filtering the received signals. Then outputs are correlated with standard pulse signal and the correlation peaks are judged. Theoretical analysis and simulation experiments show that the proposed method can eliminate the noise effectively, the anti-noise performance is superior to -20dB, improving about 10 decibel (dB) compared with the traditional methods, and higher acquisition processing gain is obtained, while its implementation is simple. It has important significance for designing digital Loran-C receiver.

Keywords—Loran-C; matched filter; anti-noise performance

I. INTRODUCTION

Long-range navigation (Loran)-C is an independent terrestrial radio navigation system, which mainly relies on ground-wave propagation, its transmission distance is long, the power of transmitted signal is high, and its anti-interference ability is strong, can be compensated for some defects of satellite navigation system [1], [2]. Therefore, Loran-C system is considered as an enhancement or backup for global navigation satellite system (GNSS) to reduce the Position Navigation Timing (PNT) service risk of navigation system [3], [4]. Because the traditional Loran-C system can’t satisfy all requirement of the satellite navigation system backup, many countries have to Loran-C system with modern transformation and upgrade, which is of great significance to improve the reliability and the integrity of national navigation timing system [5].

Loran-C system achieves the capture of Loran-C signal mainly through the detection and recognition of phase encoding, while phase code detection needs a high signal to noise ratio (SNR) condition. In the case of using the traditional capture method, amplitude modulated signal of Loran-C system is more sensitive to noise, which makes the Loran-C receiver can’t work under heavy noise environment. Therefore, capturing useful weak Loran-C signal effectively with low SNR, is the first solved problem for Loran-C navigation digital receiver.

II. LORAN-C SIGNAL

Loran-C signal is a 100 kilohertz (kHz) modulated Gaussian bell-shape pulse, which has a steep rise characteristic through strict controlling the pulse front edge. The current waveform is given by [6], [7]

\[ i(t) = \begin{cases} 
0, & t < \tau_d \\
A(t - \tau_d)^2 \exp[-2(t - \tau_d)/\delta] \times \sin(0.2\pi t + P_c), & \tau_d \leq t \leq 65 + \tau_d 
\end{cases} \] (1)

Where \( A \) is a normalizing constant related to the magnitude of the peak antenna current in amperes; \( t \) is the time in microseconds (μs); \( \tau_d \) is the envelope-to-cycle difference (ECD) in μs, which is the starting point of the envelope; and \( P_c \) is the phase-code parameter in radians, which is \( 0 \) for positive phase code and \( \pi \) for negative phase code.

Loran-C station periodically emits a group of pulses, where each standard pulse has the duration of approximately 200μs. The pulse group repetition interval (GRI), which varies from 50 to 100 milliseconds (ms), can be used to distinguish different emission signal stations. Each Loran-C master station in a given chain emits a group of nine pulses, the former eight pulse interval is 1ms, eighth and ninth pulse interval is 2ms, Ninth pulse is used to identify the master station, and slave stations only emit the similar former eight pulses. Signals from the master and slaver stations employ different phase code as in Table I [7]-[9], two groups of phase codes (i.e. GRI-A and GRI-B), which represent the complete phase code pattern, are called Phase Code Interval (PCI) of Loran-C signal. The Loran-C single pulse signal and a group of pulses signal emitted by master station are shown in Fig. 1.

<p>| TABLE I. PHASE CODE OF LORAN-C PULSE |
|-------------------------------|-------------------------------|</p>
<table>
<thead>
<tr>
<th>GRI</th>
<th>Master station</th>
<th>Slave station</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>++---++---+</td>
<td>++++++---+</td>
</tr>
<tr>
<td>B</td>
<td>+---+++++---</td>
<td>-+---++++---</td>
</tr>
</tbody>
</table>

978-1-4799-8866-2/15/$31.00 ©2015 IEEE 265
A. The Principle of Matched Filter

The matched filter is used to eliminate the noise in channel of digital receivers, which is a kind of important method to improve the value of SNR. According to the principle of information transmission, when the received digital base-band signal is through matched filtering, the noise can be effectively filtered out and the SNR can be achieved the best [10]. Transfer function of matched filter is complex conjugate frequency spectrum of the input signal. According to its frequency characteristic analysis, at a given time, all of the input signal frequency components are with the same phase after matched filtering, which makes the algebraic sum of the amplitude of frequency component reach the maximum, while the input noise is Gaussian white noise, and has the greater differences with the Loran-C signal in spectrum, thus, can be suppressed, and output SNR of matched filtering can be greatly improved [11]. The transmission principle of matched filter is shown in Fig. 2.

As given in (2), $S_f(\omega)$ is the spectrum of Loran-C signal $s_f(t)$, the matched filter transfer function is defined as follows

$$H(\omega) = S_f(\omega) e^{-j\omega t_0}$$

(2)

The input signal $x(t)$ of matched filter is composed of Loran-C signal $s_f(t)$ and noise $N_f(t)$, as in

$$x(t) = s_f(t) + N_f(t)$$

(3)

Where noise is assumed to be white noise, its power spectral density $P_n(\omega)$ is equal to $n_0/2$, the output signal $y(t)$ of matched filter is also composed of two parts, as in

$$y(t) = s_f(t) + N_o(t)$$

(4)

$s_f(t)$ is calculated as follow

$$s_f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\omega) H(\omega) e^{j\omega t} d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} |S_f(\omega)|^2 e^{j\omega(t-t_0)} d\omega$$

(5)

The average power of output noise $N_o$ is obtained as follow

$$N_o = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 \left(n_0/2\right) d\omega$$

$$= \frac{n_0}{4\pi} \int_{-\infty}^{\infty} |S_f(\omega)|^2 d\omega = \frac{n_0 E}{2}$$

(6)

Where $E$ represents the total energy of Loran-C signal, and $E$ is calculated by $E = \int_{-\infty}^{\infty} |S_f(\omega)|^2 d\omega$. $s_f(t)$can be estimated by integral summing as follow

$$s_f(t) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} |S_f(\omega)|^2 e^{j\omega(t-t_0)} \Delta\omega$$

(7)

According to the vector representation of a complex sinusoidal signal, equation (7) can be regarded as the sum of infinite vectors

$$s_f(t) = \sum_{k=-\infty}^{\infty} a_k e^{j\phi_k(t)}$$

(8)

Where $a_k$ and $\phi_k(t)$ are obtained by $a_k = \frac{\Delta\omega}{2\pi} |S_f(\omega)|^2$, $\phi_k(t) = \omega(t-t_0)$. For $t \neq t_0$, the orientation of each vector is inconsistent, for $t = t_0$, the orientation of each vector is consistent, which makes the sum of vector modulus reach the maximum value, at this point, the instantaneous power of output signal also reach the maximum value. According to (6), the average power of output noise $N_o$ is given, so the SNR of output can be realized the maximization.

B. The Design of Matched Filter

The transmission characteristics of the matched filter can also be represented by the impulse response:
\[ h(t) = s(t_0 - t) \]  

(9)

The impulse response of matched filter is the image signal of Loran-C signal \( s(-t) \) by time translation \( t_0 \). Therefore, Loran-C digital matched filter can be achieved in Finite Impulse Response (FIR) filter structure, which consists of a shift register, multiplier and accumulator, with the standard Loran-C pulse signal as the tap coefficient of FIR filter for filtering the received signal. Obviously, it is easy to implement, and its computational complexity is determined by the order of matched filter.

The order selection of digital matched filter is critical, which directly affect the filtering result. If the order of filter is too small, the acquisition precision will be reduced; if the order of filter is too high, the computational complexity will be increased, which can delay capture time and augment the difficulty of project implementation. A group of matched filter orders under different input SNR conditions are compared, through simulation analysis, the appropriate filter order is chosen. In this experiment, input signal is simulated as follow: the delay of Loran-C signal is 21 sampling points, the sampling frequency \( f_s \) is 400kHz and the carrier frequency \( f_c \) is 100kHz. By the matched filtering, detected delay points of 8 pulses signal are accumulated, and the average results are shown in table II.

Table II illustrates: along with the increase of the filter order, the delay time error introduced by matched filter decreases and the acquisition precision increases continuously. When the filter order is close to 64th order (i.e. 16 carrier cycles of Loran-C pulse signal), under the noise free condition, the delay time error is less than \( 1/f_s \), but as the filter order continues to increase, the delay time error is restricted by the sampling frequency, and will not be further reduced significantly. Therefore, considering from the perspective of implementation, the order of matched filter may be selected as the 16 carrier cycles. To further improve the acquisition precision, it is needed to improve the sampling frequency for fulfilling the specification requirements.

### TABLE II. THE EFFECT OF MATCHED FILTER ORDER ON ACQUISITION PRECISION

<table>
<thead>
<tr>
<th>The order of filter</th>
<th>Detected Delay Points After Filtering</th>
<th>Delay Time Error (in μs) Introduced By Matched Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Noise free</td>
<td>6dB</td>
</tr>
<tr>
<td>8th order (2f/( f_c ))</td>
<td>41</td>
<td>41.50</td>
</tr>
<tr>
<td>16th order (4f/( f_c ))</td>
<td>37</td>
<td>37.50</td>
</tr>
<tr>
<td>32th order (8f/( f_c ))</td>
<td>29</td>
<td>29.75</td>
</tr>
<tr>
<td>50th order</td>
<td>23</td>
<td>23.25</td>
</tr>
<tr>
<td>64th order (16f/( f_c ))</td>
<td>21</td>
<td>21.25</td>
</tr>
<tr>
<td>80th order</td>
<td>21</td>
<td>21.25</td>
</tr>
<tr>
<td>128th order (32f/( f_c ))</td>
<td>21</td>
<td>21.00</td>
</tr>
</tbody>
</table>

C. The Design of Acquisition Method

Acquisition method is proposed by combining matched filtering and cross-correlation of traditional capture method [12], [13], so as to improve the ability to capture the signal. Firstly, out-band interference of Loran-C received signal is filtered out by the band pass filter (BPF). Secondly, in-band noise is further filtered by matched filter, thereby the SNR of correlated input is improved. Thirdly, signal is sliding correlated with the standard pulse, which is generated by the frequency divider and the encoder. If the correlation value is positive, the phase of Loran-C received pulse is the same as that of the standard pulse, and if the correlation value is negative, the phase of Loran-C received pulse is opposite to that of the standard pulse. When the Loran-C signal is aligned with the standard signal, their correlation value can reach the maximum. Then, signals are accumulated of many repeated cycles. Finally, a detection threshold is set up to judge the result of accumulation. Consequently, Loran-C signal is captured. Parameters of the standard pulse, such as the interval of pulse and phase encoding, which is used in the cross-correlation, are identical with that of Loran-C signal, which is acquisition target. The flowchart of acquisition method based on matched filter is shown in Fig. 3.

![Flowchart of acquisition method based on matched filter.](image)

IV. ANALYSIS AND TEST OF METHOD PERFORMANCE

A. Analysis and Test on Anti-noise Performance of Matched Filter

The differences between signal and noise in the frequency spectrum distribution are utilized to suppress noise, which make the matched filter achieve the purpose of anti-jamming.
Generally speaking, the SNR of instantaneous output power is $2\Delta F_n T$ times larger than the SNR of average input power in matched filter (where $\Delta F_n$ is the noise bandwidth of the matched filter, $T$ is the effective duration of the input signal) [11]. For Loran-C signal, 99% of its energy is concentrated in the 90kHz~110kHz [14], so $\Delta F_n$ is 20kHz and the effective duration of input signal is about 100μs. When the Loran-C signal is matched filtered, the SNR of instantaneous output power is about 16 times larger than that of the average input power, namely, SNR is improved about 12dB by matched filtering. Simulation comparison is made to verify the effect of matched filtering on the anti-noise performance of Loran-C signal with different input SNR, the results of comparison are shown in Fig. 4.

Simulation results show that the SNR of Loran-C output signal by matched filtering is significantly improved, and can be improved about 12dB, which is consistent with the calculated value.

The anti-noise performance of matched filter is also drawn a comparison with that of other methods: in linear digital averaging method, the SNR can be improved about 10 dB as the number of pulse accumulation averaging is about 80 times[15],[16]; the anti-noise performance of Bi-spectrum Estimation for Loran-C signal with Gaussian white noise is about -12dB[17]; the SNR of Loran-C can be increased about 10dB by the Adaptive Line Enhancer [14],[18]. The anti-noise performance of matched filtering processing, therefore, is slightly higher than that of the above methods, while its implementation is simple, and its computation is less.

B. Analysis and Test on anti-noise performance of Acquisition Method

Matched filtering and cross-correlation are combined to capture Loran-C signal, which is on the basis of using the anti-noise performance of matched filter, and further exploiting on the differences between signal and noise in statistical characteristics to suppress the noise of Loran-C signal under low SNR condition, so as to enable the acquisition performance of Loran-C signal has been enhanced greatly. United State Coast Guard (USCG) published the minimum performance regulations of Loran-C receiver, where input SNR of antenna is greater than or equal to -10dB [19]. In order to verify the anti-noise performance of acquisition method based on matched filter, under the same conditions, the characteristics of correlation peak are analyzed and compared with that of traditional capture method, where the Loran-C input signal with SNR=-5dB,-10dB,-15dB,-20dB is simulated respectively, and the simulation results are shown in Fig. 5 and Fig. 6.

Fig. 5 and Fig. 6 show that; for $SNR>-10dB$, the ideal correlation peaks are obtained in the traditional capture method and acquisition method based on matched filter, both can be realized accurate peak detection, Loran-C signal capture, phase encoding detection and recognition of the master and slaver stations, while the characteristics of the correlation peak in matched filter method are superior to that in the traditional capture method; But with the loss of SNR, for $SNR \leq -10dB$, the anti-noise performance of matched filter method is slightly inferior.
the characteristics of correlation peak in the traditional capture method are not obvious enough, cannot be realized accurate peak detection; for $SNR=-20dB$, the characteristics of correlation peak in the matched filter method are still good, can be realized peak detection and Loran-C signal capture; At the same time, under the same input condition, higher acquisition processing gain is obtained in matched filtering acquisition method.

In conclusion, anti-noise performance of acquisition method based on matched filter is superior to -20dB, improving about 10dB compared with the traditional methods and the acquisition method based on delay correction[5], which is also higher than the regulations published by USCG. Higher acquisition processing gain is obtained to improve the capture capabilities of weak Loran-C signals under heavy noise environment.

V. CONCLUSION

In this paper an acquisition method based on matched filter is proposed for Loran-C signal, which is based on the study of the characteristics of Loran-C signal and digital matched filter, and the differences between signal and noise in the spectral distribution and statistical properties are employed. Acquisition accuracy and anti-noise performance of this method are verified by theoretical analysis and simulation. The result shows that anti-noise performance of this method is better than that of traditional capture method and its implementation is simple, so as to effectively solve the capture of Loran-C signal under strong noise background. It has important reference value for the development of Loran-C digital receiver.

REFERENCES


Verification of Time Telegrams in Long Wave Radio Systems

Schneider, Matthias; Ruland, Christoph
Institute for Data Communication Systems
University of Siegen
57068 Siegen, Germany
{matthias.schneider, christoph.ruland}@uni-siegen.de

Abstract—This paper describes a method to verify received time telegrams distributed by Long Wave Radio systems for the example of the radio ripple control technology.

In the given approach, the time between two time telegrams is continuously measured and compared with the time difference calculated from the time information contained in the telegrams. In other words, physical and logical information is compared. The physical time difference is directly calculated using the carrier frequency of the transmission system (for example: DCF49 transmission system uses a carrier frequency of 129.1 kHz with FSK modulation). A counter is clocked using the received carrier frequency of the system. The physical time difference between the transmissions of two time telegrams can be derived from the number of carrier cycles. The logical time difference is given by the content of the time telegrams.

The comparison of physical and logical time differences is continuously verified to detect time jumps, which may appear during the transmission of time telegrams. Therefore, manipulated or delayed time telegrams can be accurately identified. This method can be applied without changing the time distribution protocol and it can be adapted to other time distribution services.

Verification, radio ripple control, DCF77, DCF49, WWWB, MSF, BPC

I. INTRODUCTION

Long wave radio systems are widely used to distribute the current date and time information, because wireless transmission guarantees timeliness. Worldwide, there are many different long wave time services (e.g. DCF77, WWVB, MSF, JJY40/60, BPC etc.). Besides time services, radio ripple control services exist in long wave radio systems as well. In Germany this data service is offered by the EFR GmbH (Europäische Funk-Rundsteuerung). Radio ripple control services are mainly used to distribute broadcast control information in the energy industry. Thus, many thousand devices can be controlled simultaneously, with one radio ripple data telegram. In addition to the distribution of data messages, the system time is also synchronized in the communication network.

While time signal services continuously distribute the current time, the time will be sent out only at certain times in radio ripple control systems, as specified in the standards [1], [2] (for Germany). The provision of timeliness and reliability of time information is very important for many devices, which are controlled by time stamped telematic telegrams or time triggered events. Therefore it plays an essential role for public and private safety. The correct behavior of the devices in each system is only possible, if a synchronized and correct system time is available for all devices. The distribution of time telegrams is used to synchronize the system clocks of the receivers and their real-time clocks (RTC). The received time messages, which contain the actual system time, cannot be verified by the Long Wave Radio Receiver, because in the current protocols, only linear checksums and parity check bits are implemented.

II. RADIO RIPPLE CONTROL TECHNOLOGY

In Germany a radio ripple control service is offered by the EFR GmbH. In contrast to pure time services the radio ripple control service works packet-oriented. This service broadcasts on the long wave transmitter DCF49 (129.1 kHz) in Mainflingen, DCF39 (139 kHz) in Burg and in addition with the station HGA22 (135.6 kHz) in Lakhghy (Hungary). The long wave radio systems works with FSK modulated transmit data. The frequency shift is \pm 170 Hz and the data transfer rate is 200 bit/s. TABLE I. shows the (frame) structure of the protocol [1] [2].

A complete radio telegram consists of a maximum of 24 characters. Time telegrams are sent to the user address 0000 hex at regular intervals. Currently, during normal operation, a time frame is sent every 10 seconds. In general the maximum time difference between two time messages must not exceed 30 minutes. The format of the 16 character long time telegram (7 character time information) is defined in the standards [1], [2] and [4]. TABLE II. shows the information section of a time telegram excluding the start bit, parity bit and stop bit. In this time telegram the 07 April 1993 15:30:42 CEST is coded as an example.
TABLE I. RADIO RIPPLE CONTROL DATA TELEGRAM

<table>
<thead>
<tr>
<th>Start-bit</th>
<th>Data bits</th>
<th>Parity-bit</th>
<th>Stop-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Start symbol 68 bit</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>Length</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>Repeat length</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>Repeat Start symbol 68 bit</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>Telegram number</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>User address Part 1</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>User address Part 2</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>User defined information section</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2-15 Bytes</td>
<td>x</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>Checksum</td>
<td>X</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>Stop symbol 16 bit</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The EFR GmbH sends since 1995 radio ripple control telegrams in the long wave frequency band and controls currently over a million recipients.

III. RISKS

Time telegrams can be manipulated or generated by special attacks, in the current systems. For instance, it is possible to transmit a forged time telegram, thus changing the system clocks of device groups or individual devices depending on the location of the attacker. By manipulating the receiver's system clock the control behavior of the device can be changed. That means in the simplest case all control commands stored in the receiver are executed at different times. Such a temporal shift of the control commands could lead to malfunctions in the device. E.g., if a wind turbine is switched to full power during a storm.

In practice, a distinction is made between passive and active attacks against a communication system, where the former example constitutes an active attack.

A. Passive attacks

Passive attacks threaten the confidentiality of communications. This means that only the information from the eavesdropped data is used. The data will not be altered or tampered. In the long-wave radio services discussed here, this attack can be performed with each receiver or a data sniffer. A simple example for a data sniffer in a radio ripple control system is presented in [3].

B. Active attacks

Whenever data is manipulated or the reception is intentionally disturbed in a communication system in some way, this is called an active attack. Active attacks are for example, Denial of Service (DoS) attacks, man in the middle (MITM) attacks and replay attacks.

1) Denial of Service attacks (DoS)

DoS attacks on the system cannot easily be prevented in a radio transmission system. To perform a DoS attack, it is sufficient to install a jamming transmitter in the area of a receiver which transmits at the carrier frequency (e.g. 77.5 kHz for DCF77 or 129.1 kHz for DCF49).

The maximum interference range is dependent on the available transmission power of the jammer and is not in focus of this paper.

2) Man in the middle attacks (MITM)

With a MITM attack, the attacker has the ability to send the receiver false control information or incorrect time telegrams with a long wave jammer. For this, the attacker only needs to know the structure of the different transmission protocols. This information is freely available on the internet for all time signal services. With a MITM attack, the attacker intercepts the user data for the receiver and replaces it with his own manipulated information- or time-telegrams.

3) Replay attack

A replay attack is a specific type of a MITM attack. During this attack an attacker sniffs data telegrams in the transmission system and plays them back to the broadcast medium at a later instant. With such an attack, the system clock can be reset to an earlier time for one or multiple receivers, because the history of the correctly received time telegrams is not generally checked in all tested receivers.

Time telegrams can be manipulated or generated by "man-in-the-middle" attacks, because the transmission of time telegrams is not secured, as described in the introduction. Therefore, it is possible to manipulate the system clocks of device groups or individual devices depending on the location of the attacker. By manipulating the receiver's system clock the control behavior of the device can be changed as described at the beginning of this section.

Further attack scenarios on Broadcast Data Systems in the long wave frequency band have been published in [3].

IV. DIGITAL SIGNATURES

In the context of information security, authenticity is often achieved using digital signatures. For this purpose, the sender signs the user data with its secret key SK. A receiver is able to verify the signed messages by using the senders public key PK. For both operations, a suitable signature or verification algorithm is used. The following Fig. 1 shows the sequence of signing and verification process.
A digital signature is based on an asymmetric cryptosystem and a hash function $H$. The transmitter generates a unique hash value $h$ of the message $m$ with a specified length (e.g. 192 bit, 256 bit, etc.).

The sender calculates the signatures $s$ using the hash value of the message $(m)$ and a signing algorithm SIG. The signature algorithm SIG takes as input the message hash together with the private key of the sender and outputs the message signature $s$. This signature is sent together with the message to the receiver.

The receiver, computes $h'$ from the received message $m'$ using the same hash algorithm $H$ just as in the transmitter. Together with the sender’s public key $PK$, the hash value $h'$ and the received signature $s'$ the signature is verified using the algorithm VER.

A digitally signed time telegram consists of the time information (the message $m$) and the signature $s$. The receiver can thus verify the authenticity of the signed telegram. However, a data sniffer can monitor and record the telegram including the valid signature. An attacker can resend these telegrams delayed back into the transmission system. This shows that the use of digitally signed data does not protect against replay attacks. Furthermore, it must be noted that the data volume increases, because the signature must be transmitted in addition to the time frames.

The integration of digital signatures in the data transmission protocols of pure time signal transmitters, such as the DCF77 time signal transmitter in Germany is not possible. The current protocol structure cannot accommodate the added signature bits.

V. VERIFICATION OF TIME TELEGRAMS IN A LONG WAVE RADIO SYSTEM

Currently, a receiver cannot verify the time telegrams that are transmitted from a time signal service or a radio ripple control service. The use of digital signatures or encryption of the data does not protect against replay attacks, as described in the previous sections. Only if there is an optional back channel available in the receiver (WLAN, UMTS, GPRS etc.), the received time telegrams can be verified with a handshake protocol between receiver and transmitter. A receiver might send a request to verify the time telegram and gets a confirmation from the transmitter. If the message is valid, the internal RTC of the receiver can be readjusted if necessary.

Normally there is no back channel available in a long wave receiver, thus a different approach must be found for the verification of time telegrams. This approach shall be downward compatible. The function of existing receivers has to be ensured.

In order to perform the verification of time telegrams at the receiver, it is not enough to check only the received user data. There must be extracted additional information from the transmission system in order to verify the telegrams unambiguously. From the long-wave radio channel, the following additional information can be extracted on closer inspection:

- The field strength of the received signal.
- The carrier frequency of the radio channel.
- The channel characteristics between longwave transmitter and receiver.

The received field strength and the channel characteristics of the receiver are dependent on environmental factors and structural changes to the environment of the recipient. Because the jamming transmitter of an attacker is geographically placed in the vicinity of the receiver, the received signal for the jammer is different from the normal reception of long wave transmitter. In order to evaluate these quantities each receiver must be individually calibrated in the field. This would be costly.

Another possibility is the evaluation of the carrier frequency. The carrier frequency is very precise in all long-wave radio services. E.g., the carrier frequency of the transmitter DCF77 can also be used as a standard frequency (77.5 kHz).

Normally, data telegrams in digital transmission systems are digital signed to ensure their authenticity. If time telegrams have to be verified, it is important that the time information contained in the telegrams is correct.

In this paper it will be shown that time telegrams can be verified without the use of digital signatures. The verification method presented here is intended to be used for long-wave time services such as the radio ripple control service.

In this approach, the carrier frequency of the long wave radio service is evaluated for the verification of the messages, because it is very precise as described above.

The time between two time telegrams is continuously measured and compared with the time difference calculated from the time information contained in the telegrams. In other words, physical and logical information is compared. Fig. 2 shows a summary of the verification process.

After the receiver is assembled and installed, the user verifies the first time string manually. This is necessary for a defined start of the verification procedure.
Fig. 2. Flow chart of the verification

If the system is configured the physical time difference between the last verified time telegram \( t_{tc} \) and the currently receiving time telegram \( t_{tt} \) is measured. The physical time difference is directly calculated using the carrier frequency \( f_c \) of the transmission system (for example: The DCF-99 transmission system uses a carrier frequency of 129.1 kHz with FSK modulation). For this purpose, the number of carrier cycles measured between two time messages is counted by a digital counter which is increased at each positive zero-crossing of the carrier signal. The counter \( cd \) is started after the last verified time telegram \( t_{tc} \) and stopped after the correct reception of the current time telegram \( t_{tt} \). The counter value of \( cd \) can be converted into \( td_{cd} \) with the equation (1).

\[
td_{cd} = \frac{cd}{f_c} \quad [s] \tag{1}
\]

In parallel, the logical difference between the two time telegrams (\( \Delta t \)) is calculated from the content of the time telegrams with equation (2).

\[
\Delta t = t_{tc} - t_{tt} \tag{2}
\]

The values in the time telegrams must increase with each received time telegram, because time is running continuously.

If \( \Delta t \) is less than or equal to zero, an attack on the system is detected. When the difference \( \Delta t \) between \( t_{tc} \) and \( t_{tt} \) is zero, then the same time telegram was received twice. If \( \Delta t \) is negative, an attacker tried to adjust the time to the past.

In order to verify the time telegram \( t_{tt} \), the logical time difference \( \Delta t \) must be compared with the measured time difference \( td_{cd} \). Formally, this relationship is described in (3).

\[
\Delta t_{cm} = \Delta t - td_{cd} \tag{3}
\]

If \( \Delta t_{cm} \) is zero, then the time string was successfully verified. Otherwise, if \( \Delta t_{cm} \) is nonzero, the verification fails. To narrow down the error in a failed verification, for each verification the current value of the internal real-time-clock (RTC) is saved. If the verification of the current time telegram \( t_{tc} \) fails, the difference value \( \Delta rtc \) of \( rtc_v \) and \( rtc_p \) will be calculated analogously to the calculation of \( \Delta t_{tt} \) (4).

\[
\Delta rtc = rtc_v - rtc_p \tag{4}
\]

\( rtc_v \) is the time value of the internal RTC after receiving the last verified time string and analogous \( rtc_p \) is the value of the internal real time clock after the reception of the current time telegram \( tt_c \). Now the logical difference value \( \Delta rtc \) is compared to \( td_{cd} \). This relationship is described in equation (5).

\[
\Delta tcr = \Delta rtc - td_{cd} \tag{5}
\]

If the value of \( \Delta tcr \) is equal to zero, an attack is detected by the recipient, because the internal RTC is synchronized to the physical time difference \( td_{cd} \). In this case, the receivers system time is not adjusted, since the internal RTC runs synchronously to \( td_{cd} \). The attack must be stored in the log file of the receiver, and the receiver can continue working. If \( \Delta tcr \) is nonzero, then there may be several different sources of error:

- MITM- or replay attack.
- Insufficient accuracy of the internal RTC.
- Insufficient accuracy of carrier frequency for clocking the distance counter \( cd \).
- Receiver interference in the long wave frequency band.
  - Atmospheric disturbances (e.g. as snow, rain, thunderstorms).
  - Man-made noise (e.g. interference from switching power supplies).
- Voltage dips.
- etc.

In this case, the receiver must generate an error message and save it in the log files.

With this approach, a logical chain to verify time messages is formed in the receiver. The time difference is always calculated from an already verified time telegram \( t_{tt} \) to a time telegram, which is to be verified \( t_{tc} \). Fig. 3 shows some example scenarios.

The time frames \( tt_{1} \) and \( tt_{2} \) are sent without a break in the example in Fig. 3. For the verification of the time string \( tt_{2} \) in the receiver, the following steps are necessary:

- \( td_{cd} \) is calculated with the count \( cd_{2} \) from equation (1).
- The logical time difference \( \Delta t \) from the contents of the received time messages \( tt_{1} \) and \( tt_{2} \) is calculated using equation (2).
- The validity of the time telegram \( tt_{2} \) is tested by equation (3).
If $\Delta \text{cm}$ is equal to zero, the verification is OK and the time of the internal RTC can be synchronized with the received time telegram. Otherwise, if $\Delta \text{cm}$ is not equal to zero then the verification fails. Now $\Delta \text{cm}$ has to be calculated as described above.

Fig. 3 also shows the handling of transmission errors. If an incorrect time string is received ($tt_5$), it will have no effect on the verification. The receivers time string is rejected as defective and the counter will continue until the next correctly received time telegram ($tt_5$). Now, the verification procedure is performed as described. The verification process of $tt_5$ starts at the last verified time telegram ($tt_4$).

As a third scenario, a replay attack is shown in Figure 2. The time telegram $tt_5$ is recorded by an attacker and sent out with a delay to the receiver. At the receiver the delayed time telegram is received as $tt_4'$. The recipient detects that $\Delta \text{cm}$ is non-zero and the described actions are performed.

VI. IMPLEMENTATION AND TEST

The verification of the time telegrams was tested in the radio ripple control system in Germany. For this purpose a receiver has been developed consisting of the RF front-end with FSK demodulator, a narrow-band band-pass filter for detecting the carrier frequency, an FPGA and a PC.

RF frontend and FSK demodulator are from a radio ripple control receiver and the filter for the carrier frequency was constructed from discrete components.

The logics for frame detection and the counter $cd$ for the carrier frequency are realized in an FPGA, which transmits the data via a serial interface to the PC for evaluation. With a simple long-wave transmitter, as for example described in [3], attacks on the transmission system can be initiated. The operating range for the attacks depends on the transmission power. TABLE III. shows the output of the verification software V01-012-OUT. The table shows from left to right the counter-value $cd$, the physical time difference $td_{\text{cd}}$, the logical time difference $\Delta tt$, the received date and the received time. At the upper part of the table can be recognized that all time frames are received at intervals of 10 seconds in this measurement.

Between the logical ($\Delta tt$) and physical time measurement ($td_{\text{cd}}$) there is a time difference of 767 up to 778ms because in the current software version only the time between two telegrams is measured. In the new version of the software, the time difference $td_{\text{cd}}$ will be measured from the end of the verified data string $tt_4$ to the end of the current time telegram. For the use of this software version an offset must be considered for $td_{\text{cd}}$ to evaluate the difference times. The highlighted section of TABLE III. shows a MITM attack. During the attack, the time was adjusted by 49 minutes and 31 seconds to the past. After receiving the first manipulated time telegram, $\Delta \text{cm}$ is less to zero. After the detection of the manipulated time telegram, a failure analysis must be carried out as described in Sec. VI.

VII. FUTURE WORK

The paper presents the approach for the verification of time messages in long wave radio systems using the example of the radio ripple control technique. The future aim is to implement and test this approach for different time signal transmitters. Hereby should be shown that the method can be translated into universal for time signal transmitter. Furthermore, the software for the evaluation of the time messages needs to be adjusted and must support different modulation methods like FSK, ASK, and special combinations.
REFERENCES


Two-Way Coherent Frequency Transfer in a Commercial DWDM Communication Network in Sweden

Sven-Christian Ebenhag, Martin Zelan, Per Olof Hedekvist, Magnus Karlsson* and Börje Josefsson+
Department of Measurement Technologies, SP Technical Research Institute of Sweden, Borås, Sweden
* Microtechnology and Nanoscience, Photonics Laboratory, Chalmers University of Technology, Göteborg, Sweden
+ Swedish University Computer Network, Stockholm, Sweden
E-mail: sven-christian.ebenhag@sp.se

Abstract— An experimental fiber link is being established between SP Technical Research Institute of Sweden in Borås and Chalmers University of Gothenburg in Sweden. The one way fiber length is about 60 km and implemented in SUNET (Swedish University Network). The aim of the project is to evaluate the signal quality when sending a stable optical frequency utilizing a wavelength in a DWDM (Dense Wavelength Division Multiplexing) system fiber pair. The experiment uses a channel in the DWDM with the wavelength of 1542.14 nm. This wavelength is within the C band and is therefore compatible with common Erbium doped amplifiers in this network. Another aim of the system is to be ultra-stable which corresponds to a stability of $1 \times 10^{-13}$ for $\tau = 1 \text{ s}$ as well as providing the ability to distribute monitored ultra-stable frequency with a future traceability to UTC (SP) (National realization of Universal Time Coordinated within Sweden) to multiple users within the network.

Measurements of an optical frequency transfer using a fiber-link based on unidirectional light signals in parallel fibers have shown promising results in a free-running setup and in a lab environment. The fractional frequency stability, analyzed as the Overlapping Allan deviation, is approximately $3 \times 10^{-13}$ at $\tau = 10 \text{ s}$ and almost $1 \times 10^{-14}$ at $10^5 \text{ s}$.

Keywords— Frequency transfer, Optical fiber network, Optical fiber, DWDM.

I. INTRODUCTION

An experimental fiber link is being established between SP Technical Research Institute of Sweden (SP) in Borås and Chalmers University of Technology (Chalmers) in Gothenburg, both in south west of Sweden, see Fig 1. The one way fiber length is about 60 km and implemented in SUNET (Swedish University Network). The network connection is DWDM-based (Dense Wavelength Division Multiplexing) and connects the network routers in a central node with the client network, where each channel can be configured with terminal equipment based on user needs, such as Ethernet or POS (Packet-Over-SONET / SDH) technology at different bitrates.

SP is the National Metrology Institute (NMI) responsible for the national frequency and time as well as the development and distribution of these to users within Sweden. Through participation in EMRP-project NEAT-FT, there is also development of techniques useful on a European scale [1, 2]. With the large amount of fiber networks in Sweden, there is a huge benefit of constructing a system that can co-exist with the standard telecom infrastructure. There is also a desire to optimize the possible stability of such connections [3, 4] and to compare these with the performance of bidirectional transmission in single fiber [5-11]. As part of these efforts, the time lab at SP has recently acquired an optical frequency comb and an ultra-stable laser. The performance of the laser is unique in Sweden and there are several users that would benefit from access to this type of equipment in in their areas of research. With both SP and Chalmers connected to the Swedish university network (SUNET), there is a potentially huge value for Chalmers to connect to this resource in their development of optical combs and optical oscillators. Furthermore the possibility to use this fiber based technique to compare optical frequency standards at different locations will be evaluated [12-13].

Fig. 1. The figure shows a map of the south parts of Sweden and the fiber connection between SP, Borås and Chalmers, Gothenburg. The distance between the locations is approximately 60 km.
This paper presents the methods of the fiber link and the ongoing work of establishing the final connection. However, since it will be working in a commercial DWDM which cannot be interrupted unless planned in advance, and not for more than a couple of hours, it is extremely important that everything is working immediately at installation. Therefore the connection setup needs to be verified and evaluated in a controlled lab environment. The results of these evaluations are also presented here.

II. METHOD DESCRIPTION

The aim of the project is to evaluate the signal quality when sending a stable optical coherent frequency utilizing a single, predefined, wavelength in a DWDM system fiber pair. The experiment uses the channel with the wavelength of 1542.14 nm, corresponding to a frequency of $1.944 \times 10^{14}$ Hz, which also is the wavelength of our ultra-stable laser. This wavelength is within the optical communications C band and is therefore compatible with common Erbium doped amplifiers in this network. Another aim of the system is to achieve a stability better than $10^{-13}$ for $\tau = 1$ sec (Overlapping Allan Variance), as well as providing the ability to distribute monitored ultra-stable frequency with a future traceability to UTC (SP) (National realization of Universal Time Coordinated within Sweden) to multiple users within the network.

One major difference in the final setup, in comparison to most frequency transfers that are presented [5-11] is that this experiment is using a fiber pair instead of one single fiber. This means that this setup will introduce an asymmetry that most likely will affect the fractional frequency stability of the connection.

A. Frequency transfer setup between SP and Chalmers

The final setup between the SP and Chalmers will be implemented in an active fiber network, see Fig. 2. The field experiment uses a fiber pair, i.e. one fiber for transmission from east to west and another, parallel fiber from west to east. Each end of the setup utilizes dedicated fibers, but the main part of the path is on the assigned channel, where the light from the ultra-stable laser is launched into the transmission fiber using a DWDM-multiplexer and at the far end it is extracted using a DWDM demultiplexer. To match the balance with the other channels in the fiber, the power level of the light in the stable channel must be adjusted to 9 dBm at each input port.

Any reconnections requires that the commercial data traffic in the DWDM-system is interrupted, which affects the other users. Hence any work on the fiber requiring reconnections must be minimized and the equipment and method must therefore be tested and verified in advanced by proof of concept experiments in lab environments.

B. Evaluation of the frequency transfer setup in lab environment

The lab evaluation system, as illustrated in Fig. 3 is setup using a CW laser with a bandwidth of less than 1 Hz. This laser light is divided into two paths through a passive optical 50/50 splitter, of which one connects to the fiber path that are to be evaluated. The other path is used as a reference in a heterodyne mixing with the returning signal in order to be evaluated and measure the stability. To be able to measure the linewidth of CW light an AOM (Acusto Optic Modulator) is connected to the path of evaluation which shifts the light 27 MHz.

The output of the AOM is the connected to a fiber spool with 24 km of ribbon fiber and in order to create an authentic connection, an EDFA (Erbium Doped Fiber Amplifier) is used before a drop-off of light at the receiving location. An optical passive optical 10/90 splitter is used for the drop-off and passes 90 % of the amplified light back through another 24 km ribbon fiber on a spool to the initial transmitting location. The split-off 10 % of the light is intended to be the CW light to be used at the remote site. At the initial location the polarization of the returning light is adjusted manually and combined with the light from the initial splitter light through an optical 50/50 combiner before detection in a P-I-N receiver with 1 GHz electrical bandwidth.
This figure presents the CW laser source connected into two paths through an optical splitter of with one is used to transmit light in 24 km of ribbon fiber to the remote location for use. At that location the amplified light is dropped-off and returned through another 24 km of ribbon fiber. The polarization of the returning light is adjusted and combined with the initial light before measurements are made with a 1 GHz P-I-N detector.

At present the AOM is not steered which means that the connection is free running. The next step is to actively steer the frequency by feedback to the AOM.

III. RESULTS

The results for this evaluation are divided into three subsections, A to C, starting with presenting the results for the complete frequency transfer setup (A). The second subsection (B) presents results from measuring the AOM and the final subsection (C) presents results from evaluating the EDFA.

All the measurements are performed in controlled lab environment. Outliers are removed from the raw data.

A. Complete connection

This subsection presents the results from measuring the fractional frequency stability of the whole free-running connection, i.e. 48 km of ribbon fiber, one AOM, passive optical splitters and combiners, polarization controllers, and one EDFA.

The red curve in Fig.4 starting at approximately 3x10^{-13} at $\tau = 10$ s presents the free running connection during four days of measurements.

The connection performs as expected according to results from free running connections published by other research groups [11].

B. Measurements of the AOM setup

This subsection presents the results from measuring the fractional frequency stability on the AOM to evaluate what its effect is on the connection.

The black curve in Fig. 4 starting at approximately 8x10^{-15} at $\tau = 10$ s presents measurements of the AOM for four days and shows that its contribution to the stability of the complete connection at the moment is negligible for averaging times less than 4x10^9 s. The data is valid for the free running setup and when a control signal is added to the AOM setup, the influence may be higher.

C. Measurements of the EDFA

This subsection presents fractional frequency stability measurements of the EDFA for evaluation of its impact on the connection. In order to do that the ribbon fiber spools are replaced by passive optical attenuators that are presumed not to affect the measurements.

The blue curve in Fig. 4 starting at approximately 2x10^{-14} at $\tau = 10$ s presents measurements of the EDFA for four days and shows that its contribution to the stability of the connection is negligible for averaging times less than 4x10^9 s.

IV. CONCLUSION

Measurements of an optical frequency transfer using a fiber link based on unidirectional light signals in parallel fibers have shown promising results in a free-running setup and in a lab environment. The fractional frequency stability, analyzed as the overlapping Allan deviation, is approximately 3x10^{-13} at $\tau = 10$ s and almost 1x10^{-14} at 10^5 s.

The successful results from the lab experiments will enable the continuation in connecting an optical frequency transfer in a commercial fiber network.

ACKNOWLEDGEMENTS

This project is carried out within SP’s commitment as a National Metrology Institute and is supported by The Swedish Post and Telecom Authority (PTS), the and Swedish Governmental Agency for Innovation Systems (Vinnova) and...
the European Metrological Research Program EMRP under SIB-02 NEAT-FT. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.

V. REFERENCES


Frequency distribution in delay-stabilized optical DWDM network over the distance of 3000 km

Śliwczyński Ł., Krehlik P., Lipiński M.
Department of Electronics
AGH University of Science and Technology
Krakow, Poland
sliwczyn@agh.edu.pl

Turza K., Binczewski A.
Poznań Supercomputing and Networking Center
Poznań, Poland
kturza@man.poznan.pl

Abstract—In the paper we are presenting the results of the experiments we performed with sending the frequency signals (10 MHz) to the remote location exploiting the optical dense wavelength division multiplexed telecommunication network. To stabilize the phase of the frequency signal we applied the approach with the electronic stabilization of the propagation delay. We measured the residual instability resulting from the fact that in a telecommunication network the signals in the forward and backward direction do not share the same fiber and are transmitted through different pieces of equipment when passing through reconfigurable optical add drop multiplexers or optical amplifiers. Our experiments show that results may depend substantially on the route of the link. For all tested links, however, the stability was better than the stability of the signal generated by commercial 5071A cesium standard. In case of one link even the stability better than stability of H-maser was observed for averaging times longer than 1000 s.

Keywords—frequency transfer; fiber optics; delay stabilization; fiber network; optical transport network

I. INTRODUCTION

Thanks to the wide spread of optical fibers and their good transmission properties the opportunities exist to exploit them for long-distance frequency and/or time transfer. Such systems may offer superior performance compared to traditional links exploiting satellite transfer techniques or may supplement them creating backup or reserve means to supply the time/frequency signals to the end users.

The possible strategies that may be used for implementing fiber-based frequency/time transfer systems are shown in Fig. 1. The best strategy exploits a dark-fiber approach (Fig. 1a), where an effective cancellation of the phase fluctuations of the signals propagating in two opposite directions is possible thanks to the highest possible symmetry of the optical path [1]. This approach proved to be very efficient for transferring an optical carrier [2], radio frequencies [3], [4], or for joint distribution of time (pulse per second) and frequency signals [5]. The dark fiber, however, may not always be available, especially in long-distance links, spanning hundreds or even thousands of kilometers. Implementing bidirectional transfer scheme within a standard fiber telecommunication network carrying a live traffic, although in principle possible [6], [7], faces substantial technical problems because it requires bypassing each active network node with optical add-drop multiplexers and bidirectional optical amplifiers (see Fig. 1b). Installation of these additional components from the network operator point of view is an interference in the network infrastructure creating possible problems with the network reliability and safety.

In this situation an attractive option would be to exploit the capabilities of modern purely photonics optical dense wavelength division multiplex (DWDM) telecommunication networks equipped with reconfigurable optical add-drop multiplexers (ROADM), reconfigurable optical switches (called colorless channel module - CCM), erbium-doped and Raman optical amplifiers, optical chromatic dispersion compensator.

This work was supported in parts by EMRP (SIB-02 NEAT-FT project), NCBiR (PBS1/A3/13/2012 project) and Faculty of Computer Science, Electronics and Telecommunications, AGH University of Science and Technology.
modules (DCM), etc. In such a network it is possible to configure a fully-transparent optical paths (that do not require any optical to electrical conversion for routing at the network nodes) that, on the basis of so-called alien-lambda, may carry user’s optical signals in both directions (see Fig. 1c).

For the dissemination of the time/frequency signals the substantial difference exists between the two previously mentioned scenarios and the currently proposed one because of its inherent asymmetry. In the DWDM network the signals in both directions are transmitted using a pair of optical fibers and the internal structure of ROADMs and CCMs is not the same for both transmission directions (in one a wavelength selective switch is used, whereas in the opposite one it is simple passive optical splitter/coupler). This asymmetry affects the distribution of the time/frequency signals in two ways. Firstly, the calibration of the time transfer in a DWDM network will require using of some reference, calibrated time transfer system to determine the propagation delay between the local and the remote modules, and secondly, the stability of both time and frequency transfer may be expected to suffer.

The main purpose of our work was to determine experimentally the practical limits on the stability of the distribution of the radio frequencies (10 MHz) in the long-distance links exploiting DWDM optical telecommunication network and using an active stabilization of the propagation delay. Basing on our previous experience we expected quite good compensation of the fluctuations of the propagation delay of the optical fibers sharing the same optical cable as they share very similar thermal and mechanical conditions. The potential deterioration of the stability that may result from the asymmetry of the network components (ROADMs, CCMs, DCMs, optical amplifiers etc.) was, however, unknown.

II. Measurement Setup

The block diagram showing the main components of our measurement setup is presented in Fig. 2. The Local Module of the transmission system with electronic delay lines used to stabilize the propagation delay [8] was connected with a pair of fibers to the ROADM being a part of Polish Optical Internet Network (PIONIER). The second ROADM in the same location was connected to the Remote Module of the transmission system. This way we were able to configure various length optical loops by remote reconfiguration of ROADMs. The Local Module was driven using stable 10 MHz oven controlled crystal oscillator (OCXO). To assess the stability of the frequency transfer we used Quartzlock A7MX frequency and phase difference comparator operating as a phase fluctuations multiplier (with the gain of $10^2$ or $10^3$, depending on its settings), connected between the output of the Remote Module and the reference output of 10 MHz generator. (In some experiments we also used a passive hydrogen maser as a driving clock, but we observed no significant difference of the results.) The outputs of the comparator was brought to the time interval counter (TIC), that was used as a recorder, allowing to log the phase fluctuation data onto the computer.

The map showing the routes of the links we used during the tests of the stability of the frequency transfer is presented in Fig. 3. In the most of our tests the fiber network from the node located in Poznań Supercomputing and Networking Center (PSNC) – loops marked (2), (3) and (4). In one case we had the access to the fiber network from the node located in Kraków. The lengths of the links during the tests varied from 263 km to almost 3000 km.

III. Measurement Results

The first our experiments were performed on the relatively short distance of about 263 km (loop (1) in Fig. 3). The results were very optimistic – the Allan deviation dropped almost linearly to the value around $3 \times 10^{-17}$ for the averaging time of $10^2$ s. This result is very close to the noise floor of the transmission system alone, measured with the short patchcord replacing the fiber link. The open-loop fluctuations started to saturate at the level of $5 \times 10^{-13}$ after 100-200 s (see Fig. 4a). Comparison with the reference levels marked in Fig. 4 with dashed lines shows that the stability of the link outperforms substantially the stability of commercial high performance 5071A cesium clock for all averaging times, and for times longer than about 1000 s even the stability of hydrogen maser is worse.

Good results encouraged us to go for longer distances. To do this we moved the setup to Poznan and installed the equipment in the main server room at PCSS. We started from around 3000 km long link (link (4) in Fig. 3) and this time the results...
were substantially worse (see Fig. 4b, blue curve) comparing to our previous measurements. The Allan deviation plot reveals a broad bump showing two maxima around 700 s and 2500 s, that corresponds with some cycles of around 20 min and 2 hours that may be noted in the plot of time fluctuations (see Fig. 5b). No similar behavior was however observed in case of the first link we investigated – see Fig. 5a. The measurement performed on shorter routes showed that the fluctuations are a kind of general feature of investigated links, however their intensity decrease with decreasing the length (see red and gold curves in Fig. 4a).

The short-duration, 20 min cycles may be caused by the influence of the air-conditioning systems working in the server rooms - records of internal temperature show fluctuations with a peak-to-peak value of 2°C and very similar periodicity. These fluctuations affects not only the ROADM and CCMs (internal asymmetry of these devices is in the order of 30-40 ns according to the manufacturers specifications), but the chromatic dispersion compensators modules (DCM) as well. DCMs include relatively long spans of compensating fibers that are rather loosely thermally coupled so substantial asymmetry may be associated with them.

However we failed to correlate the large fluctuations with the periodicity of around two hours with any kind of activity. We learned that in the network we used the auto leveling process is performed each two hours, this however does not explain observed behavior - after turning this procedure off the situation remained unchanged. In addition, on the “silent” link number (1) the same number of ROADM were involved as on the link number (2), that showed clear two-hours bump. It is also highly improbable that the procedure that lasts for some seconds may result in observed kind of fluctuations.

IV. Conclusion

In our experiments we tried to perform systematic investigation of the effects limiting the stability of the stabilized fiber optic links exploiting optical DWDM network for long distance frequency transfer. We observed that the performance that may be obtained depends substantially on the segment of the net.
work that is used. In one, relatively short span including 4 ROADMs we obtained results as good as may be obtained using a dark fiber. In other cases the results were much worse and disturbed by some kind of periodic processes with the periodicities of about 20 minutes and two hours. The probable reason of the faster process are the fluctuations of the temperature in the server rooms whereas we found not any satisfactory explanation of the second process. It seems to depend rather on some particular segment of the network, not necessarily on the length of the fiber route. To further investigate the performance of the network we planned to perform the measurements in the longer link starting in our first location in Krakow. It however appeared to be impossible at this moment because of the upgrade of the PIONIER network. We hope to continue the experiments after finishing this upgrade process.

The general result of our experiments is that for all tested spans the obtained stability outperformed the stability of commercial 5071A cesium clocks, even at the distance of about 3000 km. The “silent” link number (1) even outperformed the stability of a H-maser for averaging times longer than about 1000 s. It thus seems reasonable to use DWDM optical telecommunication networks to deliver frequency signals to the end users that do not require the highest stability levels.

REFERENCES

The Research Progress of Two Way Time Synchronization with Fiber Based on Spread Spectrum Signal

Zhu Xiangwei, Gong Hang and Sun Guangfu
Department of Electronic Science and Engineering
National University of Defense Technology
Changsha, China
zhuxiangwe@nudt.edu.cn

Liang Kun
Time Keeping Laboratory, Division of Time and Frequency Metrology
National Institute of Metrology, NIM
Beijing, China
liangk@nim.ac.cn

Abstract—High-precision time synchronization is a basic element in some areas of aeronautical engineering, such as satellite navigation and deep space exploration. It is more accurate and stable to use the optical fibers while performing time frequency transfer than using other media such as GNSS common view and two-way satellite time frequency transfer. The frequency transfer is the main focus of the current research in optical fiber time frequency transfer. However, there is little study on time transfer, and the accuracy of time transfer is commonly in nanosecond level. In this paper, a two way time transfer method based on spread spectrum ranging is studied. For the proposed method, the accuracy is improved and extra links for data check are not needed any more. The designing schemes and the implementation progress of the engineering prototype are presented. The experimental results indicate that the time synchronization uncertainty is less than 30ps by use of the proposed method when the optical fiber transmission distance is within 2km.

Keywords—time synchronization; optical fiber; spread spectrum; two way time transfer

I. INTRODUCTION

In order to provide technical support to satellite navigation, deep space exploration and international atomic time (IAT), it’s necessary to develop high precise time synchronization technique. Usually, optical fiber time transfer is more accurate and stable than GNSS common view and two-way satellites time frequency transfer. Optical fiber time frequency transfer techniques can be divided into two categories according to the transmission mode, i.e., one-way frequency transfer and two-way time frequency transfer.

In one-way frequency transfer, slave stations receive the time reference signals sent by the master station to realize the time synchronization between the slave stations and master station. This method is easy to implement, however, it is difficult to reach high-precision, especially in sub-nanosecond level, since the signals may be fluctuated a lot by environmental impacts.

In the optical two-way time frequency transfer methods, time signals (such as 1PPS, B-code) are sent and received by both the master station and the slave stations. Signals from both directions are coupled in a single optical fiber by multiplexing techniques. The time-delay can be canceled out effectively since signals are sent and received by both stations with the same path. Furthermore, the time frequency standard can be sent from master station to slave stations with combination of one-way time frequency transfer methods under circumstances the slave stations dose not has a time frequency standard.

Currently, Optical two-way time frequency methods are widely used in remote time alignment owe to its high precision[2][3]. There are mainly four optical time frequency transfer methods with high stability, i.e., the two-way time frequency transfer method, the optical mechanical temperature compensation method, the electron conjugation phase compensation method and the Doppler cancellation compensation method. The comparisons of these four methods are shown in Table I[4]-[8].

The performances of widely used high-precision optical fiber time frequency transfer techniques are compared and summarized, and this article mainly focused on the analysis of the principle and errors of the two-way optical fiber time frequency transfer method. In light of the problems that the 1PPS phase coding and recovery of traditional two-way optical fiber may affect the accuracy, and extra two-way transfer for data check makes the system more complex.

To overcome such shortages, referring to the two-way satellite time frequency transfer (TWSTFT)[9], an optical fiber time frequency transfer method based on pseudo code ranging is proposed[10]. Due to the two way signal transfer with the same optical fiber and frequency, most of the link error is canceled out. Then, the picosecond level time delay measurement based on spread spectrum ranging is used to realize the absolute time synchronization in picosecond level. In this paper, a novel two way time transfer method based on spread spectrum and pseudo code ranging will be introduced. The designing schemes, implementation progress and experiment performance of the engineering prototype will be introduced.

National Natural Science Foundation of China (61403413)
TABLE I. COMPARISONS OF SEVERAL OPTICAL TIME FREQUENCY TRANSFER METHODS

<table>
<thead>
<tr>
<th>Transfer Methods</th>
<th>Optical Two-way Time-Frequency Transfer Method</th>
<th>Optical Mechanical Temperature Compensation Method</th>
<th>Electron Conjugate Phase Compensation Method</th>
<th>Doppler Cancellation Compensation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle</td>
<td>two-way transfer</td>
<td>feedback loop</td>
<td>feedback loop</td>
<td>feedback loop</td>
</tr>
<tr>
<td>Appliance</td>
<td>Electronical measurements</td>
<td>Optical measurements</td>
<td>Electronical measurements</td>
<td>Optical measurements</td>
</tr>
<tr>
<td>Signal transmission</td>
<td>Time and frequency signals are transmitted simultaneously</td>
<td>Time and frequency signals are transmitted simultaneously</td>
<td>Time and frequency signals are transmitted simultaneously</td>
<td>Only frequency signals are transmitted</td>
</tr>
<tr>
<td>System integration</td>
<td>Simple structure</td>
<td>Complex structure</td>
<td>Complex structure</td>
<td>Complex structure</td>
</tr>
<tr>
<td>Design applications</td>
<td>Measurements and phase modulation are independent of each other</td>
<td>Measurements and phase modulation are coupled together</td>
<td>Measurements and phase modulation are coupled together</td>
<td>Measurements and phase modulation are coupled together</td>
</tr>
</tbody>
</table>

II. THE PRINCIPLE

The presented optical time synchronization method based on spread spectrum will be discussed in detail, including the principle and implementation, which is the basis for the development of time transfer terminal.

A. The principle of the method

The principle of the optical time transfer method based on pseudo code ranging are shown in Fig. 1, the terminal contains a transmitter, a receiver and a data processing module. The local transmitter can be divided into three parts according to its structure: spread spectrum module (SSM), BPSK modulation and transmit link. The local receiver can be divided into three parts according to its structure: BPSK demodulation, spread code demodulation and receiver link. The data processing module contains several functions such as the generating local data code, recovering the remote pseudorange, computations related to pseudo code ranging and data processing of two-way time time synchronization. Local pseudo code and pseudorange are transmitted to the remote node; then the two-way time-difference of local and remote pseudorange are computed taking the advantage of the symmetry of the transfer link; Finally, 1PPS phase adjustment are performed using the previous computation results, and the remote time synchronization is achieved.

![Fig. 1. The principle of the optical time transfer based on pseudo code ranging](image)

The principles of two-way time transfer measurements, 1PPS signal transfer and data communication are described respectively as follows.

- The principle of two-way time transfer measurements

The pseudo code ranging system is used for two-way time transfer measurements, unified local spread spectrum code are used for all the terminals of the master stations and distribution nodes, ranging signals from all the terminals are modulated by the spread spectrum code, the initial phase of the signals are determined by the local 1PPS of each terminal, the intermediate frequency signals are transmitted between the the master station and distribution nodes.

- The principle of 1PPS signal transmission

There is no direct connected optical fiber link for transmitting 1PPS signal, local 1PPS is generated by 10MHZ during the fault recovery of themaster station, then the initial phase of spread spectrum code of the signal are controlled by 1PPS, thus the phase difference between local 1PPS and 1PPS of distribution nodes is obtained by two-way pseudo code ranging, and the local 1PPS is adjusted according to the phase difference, in this way, the transmission of 1PPS signal between themaster station and distribution nodes is realized.

- The principle of data transmission

The data-out of the optical time synchronization module includes the following aspect: the forward delay $\Delta t_S$ measured by distribution nodes needs to be transmitted to the master station, the backward delay $\Delta t_M$ measured by the master station needs to be transmitted to the distribution nodes, furthermore, the time information corresponding to each 1PPS also needs to be transmitted between the master station and distribution nodes. The data above must be transmitted synchronously with signals since the data has a one-to-one correspondence with time. Therefore, the data above are modulated into the spread spectrum code and transmitted with the ranging signals, the ranging signals and the communication data are transmitted simultaneously by use of the spread spectrum signal.
B. Implementation process

The principle of the optical fiber time transfer method based on pseudo code ranging is discussed above, the implementation process of this method is described in Fig. 2:

1. Generate local ranging code and transmit
2. Calculate forward time difference between local and remote node by pseudo code ranging and send it to remote
3. Recover remote ranging code transmitted by fiber
4. Calculate absolute time difference between local and remote node
5. Demodulate remote data code transmitted from optical fiber and obtain the reverse time difference
6. Perform phase alignment to local 1PPS and enable local phase adjustment when time difference over threshold

Fig. 2. The implement process of the optical time transfer based on pseudo code ranging

The details are presented as follows:

a) The local ranging pseudo code is generated by use of the local 1PPS at the transmitter, and then the pseudo code is sent to the remote;

b) The ranging pseudo code from the remote side is recovered at the receiver;

c) Perform the correlation operation between the local and remote pseudo code, compute the forward time difference between the local 1PPS and remote 1PPS, and send the forward time difference to the remote.

d) Perform the demodulation of remote data code and obtain the backward time difference of local 1PPS which were measured on the remote.

e) Compute the absolute time difference between the local 1PPS and remote 1PPS by use of the forward time difference and backward time difference.

f) Take the phase alignment control over the local 1PPS, the phase adjustment of local 1PPS should be enabled as soon as the absolute time difference is beyond the threshold.

III. ACCURACY ANALYSIS

The time frequency signal may be affected by many factors during its transmission procedure in the optical fibers, although there is little signal attenuation due to the fiber, the synchronizing precision is significantly affected by the change of external environment. Therefore, the error analysis of the optical fiber two-way transfer method is taken in this section.

A. The accuracy of pseudorange measurements

Without the consideration of multipath and other sources of interferences, the sources of measurement error of code loop mainly includes the phase jitters due to thermal noise and the dynamic stress error\(^{[11]}\). In fact, only the static pseudorange measurements are considered in application scenes corresponding to this paper, thus only the error analysis aimed at the code phase jitters due to thermal noise is taken.

Suppose the variance of the error of phase measurements due to thermal noise is denoted by \(\sigma_{\text{DLL}}\), consider the C/A code under BFSK modulation with early-late code tracking loop, the value of \(\sigma_{\text{DLL}}\) can be modulated as follows\(^{[12]}\):

\[
\sigma_{\text{DLL}} = \begin{cases} 
    \frac{B_t}{2} \frac{c}{C/N_0} \left( \frac{1}{2} - D \frac{C}{N_0} \right) & D < \frac{1}{2} \frac{C}{N_0} \\
    \frac{1}{2} \frac{B_t}{C/N_0} + \frac{1}{2} \frac{D}{C/N_0} & D = \frac{1}{2} \frac{C}{N_0} \\
    \frac{1}{2} \frac{B_t}{C/N_0} \left( \frac{1}{2} - D \frac{C}{N_0} \right) & D > \frac{1}{2} \frac{C}{N_0}
\end{cases}
\]

In equation (1), \(B_t\) is the bandwidth of RF, \(T_c\) is the code width of pseudo code, \(D\) is the distance between correlate devices. The above formula seems to be too complicated, but qualitatively, given a narrow \(D\), or a narrow \(B_t\), or a stronger \(C/N_0\), or a longer correlation integration time \(T_{coh}\) then we get a smaller \(\sigma_{\text{DLL}}\).

B. The asymmetry of the two-way optical fiber link

For the transmission of optical two-way time frequency signals, usually the method of bidirectional transmission in single fiber is used, i.e., the optical signals back and forth are transmitted in a single fiber. Thus, the asymmetry of the transfer link is corrected physically, and the error caused by path can be canceled out in principle. Dispersion characteristics are the primary cause of the asymmetry of the optical fiber link, normally shown by the following two aspects:

1) The transfer link asymmetry caused by group delay. Asymmetry of this type is proportional to the distance, and bias is likely to be introduced to the system. The bias can be estimated by use of the typical dispersion coefficient around the working wavelength of the fiber, the delay difference of two different wavelengths can be computed approximately as follows:

\[
\tau_d = D(\lambda) \Delta \lambda \cdot L
\]

Here \(D(\lambda)\) is the dispersion coefficient around the working wavelength, \(\Delta \lambda\) is the interval of the two wavelength, and \(L\) is the length of the optical fiber. Currently, for an optical fiber with length of 100km, the asymmetry error caused by group delay is within 2ps.

2) The asymmetry of transfer link caused by Polarization Mode Dispersion (PMD). Two polarization modes which are mutually orthogonal will be generated when the mode beam enters the optical fiber. Normally, these two polarization modes transmit in the optical fiber with different velocities. PMD cannot be suppressed effectively by two-way transfer since PMD may change with the direction of the mode beam. However, the PMD value is usually very small, we take Corning Corporation’s G655 (non-zero dispersion shift fiber) as an example, the PMD is about 0.04ps/km. One finds that the asymmetric error caused by PMD is about 4ps for the optical fiber with length of 100km after some computation.
C. The error from optical devices

Errors from optical devices contain the error introduced by the instability of wavelength and the error introduced by the delay jitters of the photon electricity conversion, and so on.

1) The error caused by the instability of wavelength is shown by the center wavelength shift when the operating current and operating temperature changes. Currently, the commercial spectrometer’s measurement accuracy reaches 2pm, according to which, the synchronizing error caused by wavelength instability is about 3.32ps in case that the length of optical fiber is 100km.

2) The error can be caused by the photon electricity conversion because the delay of photon/electricity or electricity/photon conversion jitters in a certain range. There two methods to suppress the error jitters as much as possible: to improve the stability of the delay of electricity/photon conversion; and to increase the code rate, the jitters is halved as soon as the code rate is doubled. For example, if the code rate has increased to be $2.5 \times 10^9$bit/s, then a code element with width $1 \times 10^{-3}$ decreases to be 0.4ps, the peak value of delay jitters of electricity/photon conversion decreases to 38ps, and its mean value is 4ps at this time, and the error caused by instability of photon/electricity and electricity/photon conversion is about 5.6ps.

D. Phase adjustment error

High precision phase adjustment needs to be performed while synchronizing the time of the distribute nodes by use of the results of bidirectional alignment. The synchronizing accuracy is usually affected by the phase adjustment error. Common used phase adjustment methods contain delay line method, digital PLL and DDS method, the accuracy of these methods varies from several ps to hundreds of ps. Currently, the phase adjust method combined with DSS and PLL reaches an accuracy of 1~2ps[13][14].

IV. EXPERIMENTS

The method proposed in this paper will be validated by experiments, and the test platform is shown as in Fig. 3. In this figure, the local and remote time frequency (T&F) terminals have the same time and frequency standards, both of them are from the same hydrogen maser. The frequency stability of hydrogen maser is less than $5 \times 10^{-13}$/s, thus the bidirectional alignment clock difference is zero theoretically, thus only the time delay fluctuation of the time transfer link should be taken into consideration. The pseudo code ranging results of local time frequency terminal is defined as forward delay, and the remote time frequency terminal is defined as reverse delay. Then the tow-way alignment results are computed by the formula (forward delay - reverse delay)/2. To verify the reliability of the results, the experiment is divided into multiple groups and test repeatedly in 25h. The two optical fibers used for signal transmission has the length of 2km, the test platform is set in the room, and the temperature and humidity fluctuate with the change of external environment.

In the experiment, the test platform did not use the bidirectional transmission in single fiber for absence of optical device with wavelength division multiplexing. The optical back and forward signals transmit in two different optical fibers and both of them has length of 2km. Therefore, the symmetry of transfer link is not implemented physically, and the accuracy of the two-way time synchronization method may be affected in some degree.

![Fig. 3. The experiment for the optical time transfer based on spread spectrum](image)

![Fig. 4. The developed optical time frequency transfer terminal](image)

The test results from multiple groups are compared to analyze the accuracy. According the fluctuation condition, the delay of pseudo ranging can be divided into 4 typical scenarios: the round trip delay is stable and unchanging (S1), the round trip delay changes irregularly (S2), the round trip delay changes secularly (S3), the round trip fluctuate wildly in short term (S4). The experiment results of two-way time synchronization of these scenarios are listed as follows:

![Fig. 5. The test results of forward and reverse delay (S1)](image)

![Fig. 6. The accuracy of two-way time synchronization (S1)](image)
phenomenon indicate that the delay of the optical fibers is not
tenfold delay (S2). In Fig. 7, corresponds to the second group, the
forward and reverse delay fluctuate wildly in several time
intervals; In Fig. 9, corresponds to the third group, the forward
and reverse delay are diminishing slowly. In Fig. 11,
corresponds to the fourth group, the forward and reverse delay
both have a big jump and some significant fluctuations. These
phenomenon indicate that the delay of the optical fibers is not
stationary, from the error analysis in section III, we know that
the fluctuation of the delay may be affected by the environment
(e.g. Temperature and humidity, vibration), processing error
from optical devices, the error in clock difference measurement,
and so on. However, the very specific reason should be
investigated by following subject experiments, which will be
discussed further. In this paper, we are concerned about the
accuracy of the two-way time synchronization method when
the one-way delay fluctuates wildly.

From Table II, one finds that the reverse delay is several or
tens of picoseconds larger than the forward delay. By analyzing
the data of forward and reverse delay in principle we found, the
reverse delay is derived from the fiber transfer, and the code
error rate exists in the optical data transfer. Analyzing the
experimental results above, we have come to the following
conclusion:

- The change law of forward delay is consistent with that
  of the reverse delay. This agrees with the fact that the
two-way time synchronization method can suppress the
  influence of the link delay and can improve the
  accuracy.
- The accuracy of two-way time synchronization is better when the change of delay is more stable. The
  accuracy of scenario 1(S1) is worse more than 1 times
  than the other 3 cases. In the rest 3 scenarios, the
  change of delay is relative stable, therefore, the
  accuracy of two-way time synchronization is with little
  difference.
- For the optical fiber with length of 2km, the time time
  synchronization accuracy of all the 4 scenarios is better
  than 30ps (RMS), and the best accuracy reaches 7.6ps.

V. CONCLUSIONS

In this paper, the performances of widely used high-
precision optical fiber time frequency transfer techniques are
summarized. A novel two way time synchronization method
with optical fiber based on spread spectrum and pseudo code
ranging is proposed, the accuracy of which is improved and
extra links for data check are not needed any more. And,
the designing schemes and the implementation progress of the
engineering prototype are introduced. The experimental results
indicate that the time synchronization uncertainty is less than
30ps when the optical fiber transmission distance is within
2km. Thus an appropriate solution is provided by this paper to
the engineering applications in remote time frequency transfer.
In the future, it could be expected that time accuracy within
1ps can be achieved by optimizing the signal system and the
transfer link or by using ranging algorithms with even higher
accuracy.

REFERENCES
dissemination methods using optical fiber network.” Frequency Control
Symposium and Exposition, Proceedings of the 2005 IEEE
International.


High Precise Time-synchronization Based on Ultra-short Pulse

Fan Shi, Shengkang Zhang, Huaiying Shang, Hongbo Wang, Haifeng Wang, Hang Yi, Zhenggang Ding, Feng Nian, Keming Feng

Science and Technology on Metrology and Calibration Laboratory
Beijing Institute of Radio Metrology & Measurement, Beijing 100854, China
Email: shifan_chinese@sina.com

Abstract—We describe the idea of a high precise time-synchronization based on ultra-short pulse using auto-correlation method. The local pulse train is generated by a femtosecond laser system tightly locked to a local atomic clock. Then the optical pulse train is distributed to remote location. The backward propagating pulses and local pulses are combined and applied to the second-harmonic generation (SHG) crystal auto-correlator which is used for precise time delay measurements. The time compensation block is constructed by an optical delay line. We have measured calibration curve of the system. The time synchronization measurement is 10fs precision.

Keywords—time-synchronization, ultra-short pulse, auto-correlation method

I. INTRODUCTION

High precise timekeeping, measurement and synchronization are the most important technological tasks in many fields of science and engineering[1-3], such as distributed radars system, global navigation satellite system (GNSS), particle accelerators, free-electron lasers and phased-array antennas for radio-astronomy. With advances in modern technologies, the time-synchronization system has become even precise. Traditionally, precise time-synchronization is achieved by satellite-based techniques such as Two-Way Satellite Time and Frequency Transfer (TWSTFT) or GPS-based measurements[4-6]; but these techniques are hard to enable the synchronization accuracy at picoseconds level or even less.

Ultra-short pulse, whose pulse duration is usually several femtoseconds to several hundred femtoseconds, is generated by a femtosecond laser system. Theoretically, shorter pulse is helpful to achieve higher precise synchronization. Moreover, because femtosecond laser has ultralow noise, it has been anticipated that ultra-short pulse would be applied in scientific and engineering facilities requiring extremely high timing accuracy[7-9].

In this paper, we describe the idea of a high precise time-synchronization based on ultra-short pulse using auto-correlation method. We have measured calibration curve of the system. The time synchronization measurement is 10fs precision.

II. SYSTEM DESIGN

The experimental setups are shown in Fig. 1. The local pulse train is generated by a femtosecond laser system tightly locked to a local atomic clock. Then the optical pulse train is distributed to remote location. The backward propagating pulses and local pulses are combined and applied to the second-harmonic generation (SHG) crystal auto-correlator which is used for precise time delay measurements. The time compensation block is constructed by an optical delay line.

![Fig.1 high precise time-synchronization system based on ultra-short Pulse](image-url)

978-1-4799-8866-2/15/$31.00 ©2015 IEEE 290
According to SHG and auto-correlation theories, when local signal and backward propagating signal inject to a SHG crystal, the SHG signal intensity $S$ is dependent on the time delay $\tau$ between two signals:

$$S(\tau) = \frac{1}{2T} \int_{-\infty}^{\infty} |E_{2\omega}(t+\tau)|^2 dt$$

so time delay can be known by measuring SHG signal intensity. The theoretical relationship curves between time delay and SHG signal intensity are shown in Fig. 2. The relationship curves with different femtosecond laser pulse width are studied: (a) 200fs, (b) 400fs, (c) 600fs, (d)800fs. In Fig. 2(a)-Fig. 2(d), when the time delay is 0, the SHG signal intensity is the largest. This behavior is not depend on the femtosecond laser pulse width.

With the increasing of time delay (the area between two dot dash lines in Fig. 2), the SHG signal intensity is smaller. However, if the time delay is large enough, the SHG signal intensity never changes. There the best working area is between the two dot dash lines. The shot dash line in the Fig. 2 is the slop of the curves and larger slope indicates higher sensitivity. The theoretical results show that if the pulse width of femtosecond laser is wider, the system sensitivity is larger, while the best working area is smaller. Therefore, proper pulse width should be chosen in the system. In the laboratory, we have constructed this system using a femtosecond laser whose pulse width is 120fs. The time transfer link free space and the time synchronization measurement is 10fs precision.

![Fig.2 The theoretical relationship curves between time delay and SHG signal intensity. The relationship curves with different femtosecond laser pulse width are studied: (a) 200fs, (b) 400fs, (c) 600fs, (d)800fs.](image)

**III. CONCLUSIONS**

In summary, we describe the idea of a high precise time-synchronization based on ultra-short pulse using auto-correlation method. We have measured calibration curve of the system. The time synchronization measurement is 10fs precision. Our work is promising to be applied in the designing and operating of high-precision facilities and future communication networks.


Study on Autonomous and Distributed Time Synchronization Method for Formation UAVs

Tao Liu
National Time Service Center of Chinese Academy of sciences
Xi’an , China
Key Laboratory of Precision Navigation and Timing Technology, NTSC, CAS
Xi’an , China
University of Chinese Academy of Sciences
Beijing , China
huangej@ntsc.ac.cn

Yonghui Hu
National Time Service Center of Chinese Academy of sciences
Xi’an , China
Key Laboratory of Precision Navigation and Timing Technology, NTSC, CAS
Xi’an , China

Yu Hua
National Time Service Center of Chinese Academy of sciences
Xi’an , China
Key Laboratory of Precision Navigation and Timing Technology, NTSC, CAS
Xi’an , China

Haifeng Jiang
National Time Service Center of Chinese Academy of sciences
Xi’an , China
Key Laboratory of Precision Navigation and Timing Technology, NTSC, CAS
Xi’an , China

Abstract—In order to implement high precision time synchronization autonomously in the absence of any external time source, in this paper, the author introduced the synchronize model of fireflies into UAV formation network, and proposed a kind of distributed time synchronization method base on broadcast, and conducted some computer simulation experiments and built one test platform to prove the feasibility of the method and its performance, the results show that the method can effectively achieve time synchronization autonomously without any external time source, and the synchronization accuracy can be achieved about 50us.

Keywords—UAVs, autonomous, distributed, time synchronization

I. INTRODUCTION

Unmanned aerial vehicles (UAV) has an irreplaceable role in modern war, the high-precision time synchronization is one of the most key technologies to achieve multi-UAVs formation flight, cooperative reconnaissance and collaborative attack. However, the existing pattern that depending GNSS satellite navigation system, or ground control station for time synchronization, its independent properties are subject to certain restricted, especially in wartime, there is a serious security risk strategy. At the same time, for the distributed autonomous UAVs, in the mode of autonomous aviation, there is no fixed location or known the exact coordinates of the reference station, nor accurate external time reference, for most cooperative task, the relative position and relative clock face between the UAVs is more important than the absolute difference between the UAV and the UTC time. So, how to achieve the time synchronization with high precision for all UAVs of the whole network completely and independently in the absence of any external time source is imminent.

For the distributed architecture such as large-scale UAVs formation, the traditional master-slave time synchronization method is no longer applicable, and the firefly synchronization model which ancient origin and had been studied in biology, chemistry and mathematics fields provides a new way of thinking to solve the problem of distributed time synchronization. In this paper, the author introduced the model of synchronization for the fireflies in biology into UAVs formation network, and proposed a distributed time synchronization method based on broadcast, making use of the existing communication links between UAVs, each UAV broadcasts its current time information, after its corresponding neighbor nodes receiving the information, these received information do simple arithmetic average, put the average value as the clock tick for the next time then broadcast again, this process is repeated several times, all the nodes in the network will ultimately lead to an identical clock reaches on average, that implement the distributed time synchronous for the whole formation network.

II. UAV FORMATION STRUCTURE

A. The Structure of UAV formation

Because the working environment for UAVs formation and self-organization network is similar to the Ad-hoc network, is completely self-organization and distributed architecture, therefore, this article will build UAV formation self-organizing network based on a system of the Ad-hoc network.

The UAV formation is composed of a few UAVs as a network and can communicate with each other, each UAV acts as a node in the network, the way to form a network with no central node, the position between each node is equal. Taking one network with five UAVs as an example, the network model is shown in Figure 1.

The project is supported by “National Natural Science Foundation of China” (No. 11403033).
The five nodes can communicate with each other, and there is no central node between five nodes, when a node has been shot down or lost for other reasons, the remaining nodes can still form a network, keeping communications and do other tasks, that is, within a network, the lost of the nodes can not affect the reconstruction of the network.

At the same time, each node in the network can use their assigned a particular frequency and a spreading code to spread spectrum modulation, and sends a message to other neighbors, in addition, this process is a broadcasting-style, achieved a single point to multi-point sends multiple access. In this paper, the author build a full-connected network, that is, in the network, one node sends a message, all the remaining nodes are able to receive it, to ensure the communication is full coverage of the network [2].

B. The UWB communication for UAV formation

The UAVs formation flight is kind of multi-user and short-range wireless communication, requirements for communication systems with high transmission speed, low-power, anti-interference ability, high system capacity, multi-site resolution ability and other characteristics, and the Ultra Wide Band(UWB) wireless communication technology just has the above advantages, is very suitable for the establishment of links, as well as data transmission and other monitoring and control functions among the UAVs.

The UWB communication standard is different from the traditional, it transmits information by loading data into a very short duration (typically picoseconds to nanoseconds), low duty cycle pulse signals, it no longer has the concept of the intermediate frequency(IF) and the Radio frequency(RF), the transmitting and receiving of the signal will not undergo the like mixers, carrier modulation and demodulation process, the transmitting and receiving equipment is relatively simple, and another particular importance is that the signal propagation delay is very small [3,4].

In summary, in this article, the author make use of the UWB communication systems in UAV formation, taking the UWB signals as an information carrier, UAV formation UWB communication system as an information carrier, and gotten help from the characteristics of having a minimal transmission delay, combining the mathematical model of fireflies and achieve the distributed time synchronization for UAV formation.

III. MATHEMATICAL MODEL

Scientists have often looked to nature for inspiration. Swarms of fireflies stretching for miles can pulse in perfect unison, all without centralized control or perfect individuals. This phenomenon provides us a new thinking to obtain time synchronization in distributed formation UAVs.

In biological systems distributed synchronization is commonly modeled using the theory of coupled oscillators. For fireflies, an oscillator represents the internal clock dictating when to flash, and upon reception of a pulse from other oscillators, this clock is adjusted. Over time, synchronization emerges, i.e. pulses of different oscillators are transmitted simultaneously [5].

A theoretical framework for the convergence to synchrony in fully-connected mesh networks was proposed by Mirollo and Strogatz, before introducing the synchronization strategy adapted to formation UAVs, the mathematical model of Mirollo and Strogatz should be presented firstly as follows.

The internal clock of a firefly, which dictates when a flash is emitted, is modeled as an oscillator, and the phase of this oscillator is modified upon reception of an external flash. In the remainder, we focus on integrate-and-fire oscillators, which are also termed “pulse-coupled oscillators”. They interact through discrete events each time they complete an oscillation. The interaction takes the form of a pulse that is perceived by neighboring oscillators [6].

To demonstrate that synchrony is always achieved independently of initial conditions, each node has an internal time or phase \( t \), it is described by a phase function \( \phi_i \), which starts at zero and linearly increments from 0 to a phase threshold \( \phi_\alpha \) and periodically “fires” every T seconds. At this point the node “fires” (in the case of firefly, flashes), in a wireless communication system, the “fire” means transmitting a synchronization signal, then resets its phase to 0, and begin to linearly increments. So, we can obtain the following equation:

\[
\frac{d\phi_i(t)}{dt} = \frac{\phi_i}{T^2},
\]

In the absence of any input from neighbors, it will naturally oscillate and fire with a period T. Fig. 2(a) plots the evolution of the phase function during one period when the oscillator is isolated.

When coupled to others, the node is receptive to the pulses of its neighbors. Coupling between nodes is considered instantaneous, and when a node \( j \) (\( 1 \leq j \leq N \)) fires at \( t = \tau_j \), i.e. \( \phi_j(\tau_j) = \phi_\alpha \), all nodes adjust their phase function as follows:

\[
\phi_j(\tau_j) = \phi_\alpha \quad \text{if} \quad i \neq j \quad (\text{2})
\]

The received pulse causes the oscillator to fire early. The parameter \( \Delta \phi(\phi_j(\tau_j)) \) in Eq. (2) is the phase increment. By appropriate selection of \( \Delta \phi \), a system of N identical oscillators forming a fully-meshed network is able to synchronize their
firing instants within a few periods [7, 8, and 9]. Fig. 2(b) plots the time evolution of the phase when receiving a pulse.

![Figure 2](image)

**Figure 2.** The time evolution of the phase function

### IV. TIME SYNCHRONIZATION IN UAV FORMATION

Introducing the above model into the UAV formation network, each UAV corresponds to a node of the model, and the clock of UAV corresponds to the phase information of the node in the model, the specific time synchronization method in UAV formation is as follows:

When a node sends clock information to its neighbor nodes, all the nodes in its broadcast domain are able to receive this information, similarly, a node receives the clock information, also can receive many sets of the clock information on other neighboring nodes.

For each node in the network connectivity, \( i = 1, 2, \cdots, N \), having different initial phases, such as \( \phi_i \neq \phi_j \neq \cdots \neq \phi_N \), assuming there is no frequency deviation of the nodes in the network, that is to say, each node has the same period, such as \( T_i = T_j, \ i \neq j, \ i = 1, 2, \cdots, n, \ j = 1, 2, \cdots, n \).

Once the phase of a node increase to the threshold, broadcast its current clock information, \( \phi_i(n), \ n = 1, 2, \cdots, N \), for the receiving node \( j \), may be able to receive \( K \) sets of the clock information at a time, then the node \( j \) will calculating the arithmetic mean of the clock information, that is \( \frac{1}{K} \sum_{i=1}^{K} \phi_i(n) \), and put the average value as the clock tick for the next time, and then the node \( j \) will continue to broadcast the updated clock information to its neighbor nodes.

The above process is repeated, according to the mathematical model, every time the node updates the clock information, correspond to adjust its clock to the mean time, after several times, the clock information of every node will be infinitely close to the mean value of all the nodes, and at some time, all nodes within the network to reach a same clock tick value, that is distributed phase synchronization is achieved throughout the network.

### V. SIMULATION RESULT AND ANALYSIS

In order to validate the theoretical method presented in the preceding sections, we conducted extensive simulations in MATLAB.

In the simulation, we make use of five nodes, built up such topology as shown in Figure 1, to simplify the simulation, the periods of nodes are discretized to integers that are uniformly distributed in a certain interval, the phase of the nodes is normalized, the initial phase of each node uniformly distributed in the interval \((0, 1)\), at the same time, according to our proposed algorithm, after broadcasting the phase information, there will be a delay at the receiving nodes, we assume the random delay is evenly distributed in \((0.001, 0.01)\) interval, the vertical axis represents the phase of the nodes, and the horizontal axis represents the synchronization period, the results of the simulation is shown in figure 3.

![Phase difference of five nodes](image)

**Figure 3.** Phase difference of five nodes

As can be seen from the figure, at the start time of the synchronous process \((n=0)\), the phase difference of each node is large, through each synchronization period, the phase difference of each node is reduced, we can see, after 10 cycles \((n=10)\) synchronized, each node to reach a same phase value, and over time \((n \geq 10)\), each node in the network has a phase difference of zero, that all network nodes to achieve a phase synchronization.

### VI. TEST AND PERFORMANCE EVALUATIONS

In this section, we introduce our platform of the test, and including specific test methods. At last, we tested the synchronization accuracy and evaluated the performance for using UWB signal to achieve the distribute time synchronization based on the M&S model.

DV9110M is Wisair launched the second generation development board; the synchronization algorithm was implemented analogously to the implementation in DV9110M.
Our test environment is made up with five nodes, deployed a topology of all-to-all. In the network, each node is distributed random initial phase, and the nodes were separated by a distance of 10 meters (the program delay about 33 ns). In the test, we used a 16-bit timer to represent the synchronization interval. The timer is configured in CTC mode where the Input Capture Register represents the top value and the period of the node is stored in OCR1A register.

Assuming a nominal oscillator frequency of 8MHz, we decided to set a prescaler of 1024. As a result we get a granularity of 125000 ticks per second. This should be good enough to achieve a synchronization precision lower than one microsecond.

Whenever a node receives a pulse signal, compare the current count in TCNT1 with the counterMax value saved in OCR1A, if the threshold is reached, trigger an interrupt, transmission pulse signal, and reset its count to 0; if not, node calculates the phase increment, added to the current count, and the count continues.

When the count reaches T/2(T is the period) of the node, triggers an interrupt in INT4 pin, all other nodes send back their count value as a set of data, that is, each set of results contains 15 32-bit integer, respectively corresponding to the count value of the counter on each node, calculating the standard deviation of these 15 sets of data, used to measure the results of synchronization errors. Collect multiple sets of data, and calculating the average, then divided by the clock frequency can be synchronized precision. Take one node as an example, synchronization error was collected between the node and the remaining four nodes, and the test data is processed in MATLAB. The results are of Fig.12. The vertical axis represents the synchronization accuracy, the unit is second, the horizontal axis represents the time, and the unit is second.

The results show that, the synchronization accuracy of any two nodes is about 50 microseconds, at the same time, the accuracy is slightly larger than the theoretical results, and this may be caused by the frequency deviation between the nodes.

VII. CONCLUSION AND FUTURE WORK

For formation UAVs, traditional centralized time synchronization mechanism is applied in large-scale distributed wireless networks, exists serious error accumulation, at the same time, with the complex changes in the network topology, the robustness of the algorithm being seriously challenged. In this paper, making use of the characteristics that transmission delay is small for UWB signal, compare with conventional signal, through theoretical analysis, we consider that the M&S model is still suit for UWB signal and proposed a distributed time synchronization method based on broadcast to achieve the time synchronization in the absence of any external time source. To verify our analysis and completion, we did some simulation and built a test platform, and implemented the time synchronization algorithm on the platform. The final test results show that make use of UWB signals, the time synchronization algorithm based on the M&S model can synchronize multiple nodes indeed, and the synchronization accuracy can be achieved about 50us. Our future research will focus on the frequency deviation between the nodes, and make further improvement of the synchronization accuracy.

ACKNOWLEDGEMENT

The research work is supported by National Natural Science Foundation of China under Grant No. 11403033.

REFERENCES

Analysis of System Time Performance in BeiDou Satellite Navigation System

Jun Lu*, Ye Ren†, Xiaohui Li†‡, Ya Liu†‡, Shougang Zhang†‡
Email: {liuya}@ntsc.ac.cn
* Beijing Institute of Tracking and Telecommunication Technology, Beijing, China
† National Time Service Center, Xi’an, China
‡ Key Laboratory of Precision Navigation and Timing, National Time Service Center, Xi’an, China

Abstract—As same as other satellite navigation systems, the time difference is considered as one of the basic observations in Beiou satellite navigation system. Therefore, an accurate and stable system time is crucial to realize the system function. Based on the inverse decomposition method, this paper analyzes the requirements of Beidou system time (BDT) on behalf of the Beidou System service.

Three time transfer links, The Two-Way Satellite Time transfer, GPS Common-View Time Comparison and Beidou Common-View Time Comparison, are established between National Time Service Center and the Beidou system time center in order to obtain the time difference between BDT and UTC(NTSC). At last, the accuracy and stability of BDT are discussed based on the performance of these three links

Keywords—Beidou system time (BDT) ;Time Comparison;time difference;

I. BEIDOU SATELLITE NAVIGATION SYSTEM

Beidou navigation satellite system (BDS) is an global satellite navigation and communication system established by China, which is the third operating satellite navigation system after GPS and GLONASS. Since the official file of space signal interface control of BDS is lunched at December 27, 2012, it represented that BDS provides service of passive positioning, navigation and timing (PNT) toward Asia-Pacific region [1].

BDS is divided into three segments, which are space segment, ground segment and user segment. The space segment consists of 5 geostationary orbit satellites(GEO), 27 Medium Earth Orbit Satellites(MEO) as well as 3 Inclined Geosynchronous Orbits Satellites(IGSO). Among these satellites, 5 GEOs are currently working on the orbit and their station keeping position are 58.75°,80°, 110.5°, 140°, 160° respectively, and 3 MEOs are settled on three orbits separately.

BDS provides high quality of positioning, navigation and timing service around the world, which includes open service and authorized service. The open service offers positioning, speed measurement and timing to global user for free with 10m of positioning accuracy, along with 0.2 m/s speed measurement accuracy and 10ns timing accuracy.

Besides, the authorized service, which is designed for the users with requirement of high accuracy and reliability, provides additional communication service and system integrity information.

In order to present the time performance in BDS, the time difference between BDT and UTC is analyzed on the basis of the time difference between BDT and UTC (NTSC) provided by time comparison link between NTSC and Beidou system time center.

II. TIME COMPARISON LINK PERFORMANCE

Three types of time comparison links are established respectively between NTSC and Beidou system time center: Two-way satellite Time and Frequency Transfer (TWSTFT), GPS Common-View Time Comparison and Beidou Common-View Time Comparison.

III. TIME COMPARISON ACCURACY

During individual measurement, smoothed time difference curve in remote time comparison is treated as the real time difference. The accuracy of time comparison is defined by RMS which is regarded as type A uncertainty and is represented in the following equation [2][3],

\[ RMS = \sqrt{\frac{\sum (x - \bar{x})^2}{N}} \]

where x is observed data and \( \bar{x} \) smoothed data.

A. Accuracy of GPS Common-View

The GPS Common-View Time Comparison data is selected as the time difference between UTC(NTSC) and BDT from January 3, 2013 to august 6, 2013 (MJD56300~56510) ,211 days in total. The observed data is fitted by a smooth curve and its corresponding smoothed data is obtained. The result of GPS Common-View Time Comparison is shown in Fig. 1.
B. Accuracy of Beidou Common-View

The Beidou Common-View Time Comparison data is selected as the time difference between UTC(NTSC) and BDT between October 10, 2013 to 26 July, 2013 (MJD56453~56499), 47 days in total. The observed data is fitted by a smooth curve and its corresponding smoothed data is obtained. The result of Beidou Common-View Time Comparison experiment is shown in Fig. 2.

C. Accuracy of TWSTFT

The TWSTFT data is selected as the time difference between UTC(NTSC) and BDT from June 10, 2013 to 13, July, 2013 (MJD56453–56487), 34 days in total. The observed data is fitted by a smooth curve and its corresponding smoothed data is obtained. The result of TWSTFT is shown in Fig. 3.

D. Discussion of performance of three time Comparison links

Remote time comparison is one of the main techniques to the evaluation of system time measurement. By calculating, the accuracy of three types of time comparison links is shown in Table 1.

<table>
<thead>
<tr>
<th>Time comparison method</th>
<th>Accuracy</th>
<th>Daily stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS common-view</td>
<td>1.8ns</td>
<td>2.1E-14</td>
</tr>
<tr>
<td>Beidou common-view</td>
<td>1.3ns</td>
<td>1.5E-14</td>
</tr>
<tr>
<td>TWSTFT</td>
<td>0.5ns</td>
<td>5.8E-15</td>
</tr>
</tbody>
</table>

As Table 1 shows, among three links, TWSTFT owns the highest accuracy. The accuracy based on Beidou common-view is slightly higher than it based on the GPS common-view, which may due to the receiver performance, or the high orbit satellites utilized by Beidou common-view. Compared with GPS common-view, the error of Beidou common-view has higher correlation which leads to numerous errors canceling mutually.

IV. BDT PERFORMANCE ANALYSIS

Based on TWSTFT, the time difference between BDT and UTC(NTSC) is obtained. The data dates back from September 1, 2013 to November 13, 2013 (MJD56301–MJD56609), 308 days in total, with 18 datum for each day.

A. BDT accuracy analysis

Time difference between BDT and UTC (NTSC) is shown in Figure 4 and time difference between BDT and UTC is shown in Figure 5. Table 2 presents statistic feature of time difference data. It is illustrates that the time difference between BDT and UTC (NTSC)/UTC are less than 100ns and their corresponding RMSs are less than 30ns.
Fig. 4. BDT and UTC(NTSC) time difference

Fig. 5. BDT and UTC time difference

TABLE II. STATISTIC FEATURE OF TIME DIFFERENCE BETWEEN BDT AND UTC/UTC(NTSC)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>STD</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTC(NTSC)-BDT</td>
<td>7.9ns</td>
<td>37.3ns</td>
<td>78.8ns</td>
<td>-64.9ns</td>
</tr>
<tr>
<td>UTC-BDT</td>
<td>9.6ns</td>
<td>35.9ns</td>
<td>75.6ns</td>
<td>-64.8ns</td>
</tr>
</tbody>
</table>

B. BDT frequency accuracy

Based on the result of time comparison, it is necessary to analyze the frequency difference between BDT and UTC(NTSC). Due to five days time latency of UTC in Circular T, the frequency difference with 5 days latency is calculated based on time difference with 5 days latency and the result is presented in Fig.6 and Fig.7, along with their corresponding statistic feature in Table 3. It is shown that the frequency accuracy prior to 5e-15.

Fig. 6. Frequency difference of BDT-UTC

Fig. 7. Frequency difference of BDT-UTC(NTSC)

TABLE III. STATISTIC FEATURE OF BDT FREQUENCY DEVIATION

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>STD</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTC(NTSC)-BDT</td>
<td>6.4e-17</td>
<td>4.6e-15</td>
<td>8.9e-15</td>
<td>-1.5e-14</td>
</tr>
<tr>
<td>UTC-BDT</td>
<td>7.0e-17</td>
<td>4.2e-15</td>
<td>8.1e-15</td>
<td>-9.9e-15</td>
</tr>
</tbody>
</table>

C. BDT Frequency stability

BDT frequency stability is calculated based on both BDT–UTC(NTSC) and BDT–UTC. Table 4 presents frequency stability of BDT-UTC with the time latency of 5 days, 10 days, 20 days respectively and it illustrates that BDT daily frequency stability is around 2e-14.

TABLE IV. BDT FREQUENCY STABILITY WITH VARIOUS TIME

<table>
<thead>
<tr>
<th></th>
<th>1 day</th>
<th>5 days</th>
<th>10 days</th>
<th>20 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTC(NTSC)-BDT</td>
<td>2.1e-14</td>
<td>1.2e-14</td>
<td>1.3e-14</td>
<td>1.0e-14</td>
</tr>
<tr>
<td>UTC-BDT</td>
<td>1.5e-14</td>
<td>1.5e-14</td>
<td>1.4e-14</td>
<td></td>
</tr>
</tbody>
</table>

V. CONCLUSION

Based on statistical results of three comparison links, time difference of both BDT-UTC(NTSC) and BDT-UTC are obtained and the BDT performance is evaluated. The result shows the BDT time accuracy less than 100ns, and the accuracy and stability over 5 days can maintain at 1e-14 level.

REFERENCES

Ytterbium optical lattice clock at INRIM

Marco Pizzocaro*†, Filippo Bregolin*‡, Gianmaria Milani*†, Benjamin Rauf*†, Pierre Thoumany*†, Giovanni Antonio Costanzo*†, Filippo Levi* and Davide Calonico*

*Istituto Nazionale di Ricerca Metrologica (INRIM), Physic Metrology Division, Str. delle Cacce 91, 10135 Torino, Italy
†Politecnico di Torino, Dipartimento di Elettronica e Telecomunicazioni, C.so duca degli Abruzzi 24, 10125 Torino, Italy
‡Università di Torino, Dipartimento di Fisica, Via Giuria 1, 10125, Torino, Italy
Email: m.pizzocaro@inrim.it

Abstract— We present an optical lattice clock based on ytterbium $^{171}$Yb atoms developed in the laboratories of INRIM. In the experiment, we cool and trap ytterbium atoms in a two stage magneto-optical trap (MOT) (at 399 nm and 556 nm for the first and second stage, respectively). Atoms are then transferred in a horizontal, one-dimensional optical lattice at the magic wavelength (759 nm). Here the clock transition at 578 nm is probed by a laser stabilized on an ultra-stable cavity. We describe the generation of all the laser sources, the physic package and the operation of the clock. Lasers at 399 nm, 556 nm and 578 nm are obtained, with different techniques, using non-linear crystals starting from infrared sources. The clock laser is stabilized using a high finesse notched ULE cavity. The lattice is made with a titanium-sapphire laser. The aluminum vacuum chamber is designed for wide optical access and its temperature is measured by 8 thermistors for blackbody shift evaluation. Our system allows for fast loading of the lattice with $1 \times 10^7$ atoms trapped in the lattice in 250 ms. We obtained preliminary spectroscopy results and we locked the clock laser to the atomic line. Future perspectives are discussed.

I. INTRODUCTION

Among other the clock transition $^1S_0 \rightarrow ^3P_0$ at 578 nm of ytterbium $^{171}$Yb neutral atom is recommended as secondary representations of the SI second reflecting the measurements made at National Institute of Standards and Technology (NIST) [1], [2], at the National Metrology Institute of Japan (NMIJ) [3] and at the Korea Research Institute of Standards and Science (KRISS) [4]. Figure 1 show the Ytterbium transitions relevant for clock operations. Note that $^{171}$Yb is fermionic with nuclear spin $I = 1/2$ and further hyperfine structure is present. Ytterbium is easy to cool and trap in a double stage magneto-optical trap (MOT) exploiting the strong transition at 399 nm $^1S_0 \rightarrow ^1P_1$ (linewidth 29 MHz) and achieving microkelvin temperatures with the weaker transition at 556 nm $^1S_0 \rightarrow ^3P_1$ (linewidth 182 kHz). Ytterbium can then be loaded in an optical lattice at the magic wavelength (759 nm). A repumper at 1389 nm resonant with the $^1P_0 \rightarrow ^3D_1$ transition can be used to pump atoms from the clock state $^3P_0$ to the ground state $^1S_0$.

We are developing a ytterbium optical lattice clock based on $^{171}$Yb. The experimental setup and the laser ensemble are complete while the characterization of the clock is under way. In the following we will describe the physic package, the laser ensemble and the operation of our clock as well as first spectroscopy results and future perspectives.

II. EXPERIMENTAL SETUP

A. Physic package

Figure 2 shows a scheme of the physic package and of the vacuum chamber. The atomic source is an effusion oven at 400 °C that produces a collimated atomic beam. Atoms are trapped in a custom aluminum chamber with indium-sealed viewport, designed for wide optical access. The chamber is designed without a Zeeman slower. Instead the distance between the trapping region and the atomic source is made as short as possible to maximize the atomic flux. During operation the chamber is kept in ultra-high vacuum (pressure $< 10^{-9}$ mbar) by two ion pumps and one non-evaporable getter pump.

The water-cooled MOT coils are outside the vacuum chamber, in the vertical direction. Three pairs of Helmholtz coils are used to compensate the stray magnetic field. Eight thermistors are placed on the aluminum vacuum chamber for blackbody shift evaluation.

B. Laser ensemble

The 399 nm radiation is obtained by second harmonic generation (SHG) from a 798 nm titanium-sapphire (Ti:sapphire)
laser using a lithium triborate (LBO) crystal in an enhancement cavity [5]. The Ti:sapphire laser has an output power of 1.1 W pumped by a 8 W solid state pump laser at 532 nm. Up to 0.9 W has been obtained but a typical output of 0.5 W at 399 nm is used for the experiment. The frequency can be locked to the atomic resonance of any ytterbium isotopes by transverse spectroscopy on an auxiliary atomic beam. Polarization-maintaining optical fibers deliver the 399 nm light to the atoms for the first stage MOT, the slower beam and the detection probe beam.

The 556 nm radiation is obtained by SHG from 1112 nm amplified, ytterbium doped fiber laser using a single-pass periodically-poled potassium titanyl phosphate (PPKTP) crystal. Typically 10 mW of 556 nm light are obtained starting from 1.0 W of infrared light. The frequency of the 556 nm laser is locked to the resonance of a Corning Ultra Low Expansion (ULE) cavity, with acousto-optic modulators (AOMs) bridging the gap to the frequency of $^{171}$Yb. The green laser is sent to the second stage MOT by polarization-maintaining optical fiber.

The lattice is made by a Ti:sapphire laser pumped by a solid state pump laser at 10 W. An AOM is used as an optical isolator and for power-stabilization. Typically 1 W of light is send to the atoms using a polarization-maintaining optical fiber.

The clock laser at 578 nm is obtained by sum frequency generation (SFG) in a waveguide periodically-poled lithium niobate (PPLN) crystal using an erbium fiber laser at 1030 nm and a neodymium-doped yttrium aluminium garnet (Nd:YAG) laser at 1319 nm [6]. The output power is typically 4 mW. The frequency is stabilized using the Pound-Drever-Hall technique on a 10 cm ultra-stable cavity made by ULE, with fused-silica mirrors and ULE rings. The temperature of the cavity is stabilized by a double stage control with Peltier elements to the point of zero coefficient of thermal expansion of ULE. The temperature control is a digital implementation of the powerful Active Disturbance Rejection Control (ADRC) technique [7].

The clock light is delivered to the cavity, to the atoms and to a fiber comb by compensated optical fiber links.

A solid state laser at 578 nm is under development. The radiation is obtained using a co-doped LiLuF₄: Dy₃⁺Tb₃⁺ fluoride crystal pumped by 450 nm diode laser. [8]. This new compact system with its unique direct emission at 578 nm could replace the current laser in the future.

A pigtail distributed feedback laser at 1389 nm is used as repumper from the clock state to the ground state. It has an output on the atoms of 10 mW and a frequency variation $<100$ MHz. The power is enough to power-broadened the ytterbium line to 300 MHz so that active frequency stabilization is not needed.

### C. Clock cycle

The clock experimental setup is sketched in fig. 3. First $^{171}$Yb atoms are trapped in a 399 nm MOT from the atomic beam. The six laser beams have a total power of about 30 mW with a 1/e radius of 1 cm and a detuning of $-20$ MHz. The magnetic field gradient is 0.4 T/m along the vertical axis. A seventh beam, with a power of 50 mW and a detuning of $-360$ MHz is focused counter-propagating to the atomic beam and acts as a slower. There is no dedicated magnetic field to make a Zeeman slower but we exploit the leaking field of the MOT coils. With the slower beam, we can capture up to $4 \times 10^7$ $^{171}$Yb atoms. For clock operations, we capture typically $1 \times 10^8$ atoms in 150 ms of 399 nm MOT.

The second stage 556 nm MOT is loaded from the 399 nm MOT with an efficiency up to the 70% simply turning off the 399 nm beams. The 556 nm beams have a total power of 2 mW, a 1/e radius of 0.5 cm and are left on during the first stage. In 60 ms we apply 3 stages at different frequency, intensity and magnetic field gradient to maximize the fraction of atoms transferred in the lattice. The first stage (30 ms) is tuned to maximize the number of atoms in the 556 nm MOT and has a magnetic field gradient of 0.25 T/m. The second stage (20 ms) the frequency is brought closer to resonance and the magnetic field reduced to 0.18 T/m to minimize the atoms temperature, that is reduced to 10 µK. The lifetime of this stage is 3 s. In the third stage (10 ms) the magnetic field gradient is increased back to 0.25 T/m and the frequency is tweaked to maximize the number of atoms in the lattice.

The lattice laser is delivered to the atoms by polarization-maintaining optical fiber and is focused by an achromatic lens to a waist of 45 µm. The laser is retro-reflected by a curved mirror to form a lattice with a depth of 300 recoil energies for 1 W of power. The lattice is horizontal. We trap typically up to $5 \times 10^3$ atoms in the lattice while the maximum number of atoms we trapped was $3 \times 10^4$. Lifetime of atoms in the lattice is 3.0 s.

Atoms in the lattice are probed by the clock laser at 578 nm. The clock laser is collimated with a waist of 200 µw collinear...
to the lattice and is sent to the atoms through the lattice back-reflector, that is transparent at this wavelength. The lattice and clock polarization are aligned vertically. Other than the 578 nm light, all other radiations are stopped by mechanical shutters during spectroscopy.

The spectroscopy of the atoms is performed by detecting with a photomultiplier tube the fluorescence from the atoms by 3 pulses of resonant light at 399 nm of the duration of 1 ms. The first pulse measures the atoms left in the ground state. The second pulse is used to subtract the background from scattered light and the atomic beam. After the second pulse the repumper laser at 1389 nm is used to pump back atoms from the clock state to the ground state in 12 ms. Then the third pulse detect the number of excited atoms.

Atoms can be prepared for interrogation in around 250 ms. We used clock pulses (Rabi pulses) of typically 50 ms to 100 ms. Total cycle length with detection is between 350 ms to 450 ms.

### III. First Spectroscopy Results

Figure 4 show the spectroscopic signal from $1 \times 10^4$ atoms in the lattice, while interrogated by 100 μW pulses of 578 nm light of 100 ms. From the fit of the shape of the sidebands (red line in the figure) we can deduce a trap frequency of 70 kHz (consistent with a depth of 300 recoil energies) and an atomic temperature of 5 μK.

Applying a vertical magnetic field during spectroscopy reveals the hyperfine structure of the ytterbium clock line (fig. 5). The clock line is the average of the two π transitions, whose separation changes by 20 Hz/mT. In this figure the polarization of the clock laser was tilted respect to the vertical to show the also the σ transitions.

We used the spectroscopy signal to lock the frequency of the clock laser to the atoms using an AOM.

### IV. Conclusions

We have show the status of the ytterbium optical lattice clock at INRIM. The experimental setup is complete and we achieved first spectroscopy results. In the next step we will finish the characterization of the clock and measure its absolute frequency respect to the cryogenic fountain IT-CsF2 [9].

Moreover our ytterbium lattice clock is part of the EMRP project "International Timescales with Optical Clocks" [10], where a comparison campaign with other clocks is planned, both local and remote. The clock will be part of a proof-of-principle relativistic geodesy experiment. It will be compared to a transportable strontium clock developed at PTB moved to the Laboratoire Souterrain de Modane (LSM) in the Fréjus tunnel. The comparison will be made through a optical fiber link and with a transportable frequency comb developed at NPL at LSM. General relativity predicts a frequency shift of $10^{-16}$/m and we should measure the 1000 m elevation difference between the two clocks at the decimeter level within a few hours.

As well, the Yb clock is part of the project AQUASIM, aiming to compare INRIM’s clock to a ytterbium degenerate Fermi gas experiment at European Laboratory for Non-Linear Spectroscopy (LENS) in Florence [11]. The connection will exploit the already existing optical fiber link between the two laboratories [12] and will be useful for studies of collisions physics and quantum simulations.

**ACKNOWLEDGMENT**

The authors acknowledge funding from the EMRP Project SIB55-ITOC, MIUR Project PRIN2012 AQUASIM and ITN Marie Curie Project FACT. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.

**REFERENCES**


Two independent strontium optical lattice clocks for practical realization of the meter and secondary representation of the second

Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudziądzka 5, PL-87-100, Toruń, Poland
Email: zawada@fizyka.umk.pl

J. Zachorowski, M. Piotrowski and W. Gawlik
M. Smoluchowski Institute of Physics, Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, St. Łojasiewicza 11, PL-30-348, Kraków, Poland

F. Ozimek and C. Radzewicz
Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Pasteura 5, PL-02-093 Warsaw, Poland

Abstract—We report a system of two independent strontium optical lattice standards probed with a single shared ultra-narrow laser. This allows verification of relative stability of both optical standards. The absolute frequency of the clocks can be roughly verified by the use of an optical frequency comb with the GPS-disciplined Rb frequency standard or, more accurately, by a long distance stabilized fiber optic link with the UTC(AOS) and UTC(PL) via the OPTIME network.

I. INTRODUCTION

Ultracold neutral atoms in an optical lattice [1] are seen as an alternative to single-ions [2] for development of optical frequency standards. In particular, the best realisations of the strontium atomic clocks reached accuracy and stability at the $10^{-17}$ level or better [3]–[7].

The $^1S_0 - ^3P_0$ transition in neutral strontium was recommended by the International Committee for Weights and Measures for practical realization of the metre and secondary representation of the second. Due to limited pool of optical strontium atomic clocks working worldwide the International Bureau of Weights and Measures (BIPM) set practical relative uncertainties above $1 \times 10^{-15}$ in case of fermionic isotope $^{87}$Sr [8] and $1 \times 10^{-14}$ in case of bosonic isotope $^{88}$Sr [9]. Enlarging this pool is an essential prerequisite for a possible redefinition of the second.

II. OPTICAL LATTICE STANDARDS

Our experimental set-up has been described in detail in [10], so only its most essential elements are presented below.

A simplified scheme of the system of two optical lattice clocks is depicted in Fig. 1. Two optical frequency standards (Sr1 and Sr2) are based on the $^1S_0 - ^3P_0$ transition in neutral strontium atoms (isotope $^{87}$Sr or $^{88}$Sr). Two clouds of atoms in Sr1 and Sr2, trapped in the vertical optical lattices, are independently probed by an ultrastable laser with spectral width below 1 Hz. The laser beam is split into two optical paths. The frequencies of both beams are independently digitally locked to the narrow atomic resonances in each standard by a digital lock and acousto-optic frequency shifters. The frequency of the clock transition can be compared by the use of an optical frequency comb with the GPS-disciplined Rb frequency standard or, more accurately, by a long distance stabilized fiber optic link with the UTC(AOS) and UTC(PL) via the OPTIME network [13].

The short-time frequency reference of the optical standards, i.e. the ultrastable laser, is an ECDL laser locked to the TEM$_{00}$ mode of the high-Q cavity. The light from the ultrastable laser is transferred to the Sr1 and Sr2 standards and to the optical...
frequency comb through fibers. Each fiber has a system of active Doppler cancellation of the fiber-link noises to assure the transfer of stable optical frequencies [14].

In both Sr1 and Sr2 systems the Fabry-Perot diode lasers are injection-locked to the light from ultrastable laser. The master-slave system filters out any power fluctuations of the injection laser. The beam is passing the AOM of the digital lock and is injected to the optical lattice such that the beam is exactly superimposed with the lattice and its waist is much bigger than the size of the sample of atoms.

III. Results

The presented results were measured with both standards tuned to the \(^1S_0 - ^3P_0\) transition in bosonic \(^{88}\)Sr. Comparing two clocks using the same atomic species assures that no systematic effects have been overlooked in their individual accuracy budget, since the measured frequencies should be the same within the accuracy budgets of both standards.

The difference between the corrections in both standards gives the momentary frequency difference between the two clocks. The measured frequency stability in fractional units represented by the Allan standard deviation is presented in Fig. 2. For average times \(\tau\) greater than 60 s the Allan deviation decreased with \(\sigma_y(\tau) = 3.41(27) \times 10^{-14}/\sqrt{\tau}\). This value is close to \(2.3 \times 10^{-14}/\sqrt{\tau}\) obtained by Katori et al. [15], [16] for asynchronous operation of two clocks, a 3D lattice clock with bosonic \(^{88}\)Sr and a 1D lattice clock with fermionic \(^{87}\)Sr. The synchronous excitation, which in our system will be implemented in the near future, allowed the group of Katori to improve the stability of their bosonic clock to \(\sigma_y(\tau) = 4 \times 10^{-16}/\sqrt{\tau}\) [17].

The stability of the Sr1 was also compared with Stability of the UTC(AOS) maintained by the hydrogen maser in the Space Research Centre at Borowiec Astro-Geodynamic Observatory. The comparison was made over new branch of the OPTIME network, a dedicated 330 km long stabilized fiber optic link. The measured frequency stability in fractional units represented by the Allan standard deviation is presented in Fig. 3. The measured stability reached \(2 \times 10^{-15}\) after 500 s of averaging.

The fiber connection to the UTC(AOS) was also used to measure the absolute value of the clock frequency. We have evaluated the main contributions to the frequency shifts in both standards and compared them in Table I. Some of the shifts were measured by direct observation of the frequency of the clock line, making sets of four simultaneous (inter-laced) locks to the atomic line with four different values of the evaluated parameter. In this way we evaluated the quadratic Zeeman shift, the light shift from the clock light, and the uncertainty of the collisional shift. All other shifts were determined by measuring respective physical parameters (e.g. temperature for the blackbody radiation shift and density for the density shift) and calculating the shifts from well-known models [18]–[24]. The relative frequency difference measured between two standards, taking into account all the shifts and the accuracy budget, is equal to \(5(27)_{syst}\) Hz. The absolute frequency of the \(^1S_0 - ^3P_0\) clock transition in bosonic \(^{88}\)Sr measured in Sr2 in relation to UTC(AOS) is equal to \(429\,228\,066\,418\,015(14)_{syst} 6\) stat Hz.

IV. Conclusion

We presented the frequency stability of the two strontium optical lattice clocks and comparison with the stability of the
UTC(AOS) maintained by the hydrogen maser in the Space Research Centre at Borowiec Astrogodynamic Observatory. The comparison was made over the 330 km long stabilized fiber optic link in the OPTIME network [13]. These standards are interrogated by a shared ultra-narrow laser pre-stabilised to a high-Q optical cavity. The frequency of the clock transition in our experimental setup can be can be roughly verified by the use of an optical frequency comb with GPS-disciplined Rb frequency standard [25] or, more accurately, by a long distance stabilized fiber optic link with the UTC(AOS) and UTC(PL) via the OPTIME network.

In the current state of the experiment, the fluctuations of the relatively high magnetic field, corresponding to the clock transition line-width of few tens of Hz, limit stability of our clocks to about $2 \times 10^{-15}$. The accuracy of the Sr2 system, according to our preliminary uncertainty budget, is better than $5 \times 10^{-14}$. Given the BIPM limits are above $1 \times 10^{-14}$ for the bosonic clocks, such stability and accuracy are sufficient for local practical realization of the meter and secondary representation of the second.

ACKNOWLEDGMENT

The authors would like to thank Dr Jérôme Lődewyck and Dr Rodolphe Le Targat for valuable discussions and help in designing the optical lattice standards. This work has been performed in the National Laboratory FAMO in Toruń and supported by the subsidy of the Ministry of Science and Higher Education. Individual contributors were partially supported by the Polish National Science Centre Projects No. 2012/07/B/ST2/00235, No. DEC-2013/11/D/ST2/02663, No. 2012/07/B/ST2/00251, No. 2012/05/D/ST2/01914 and by the Foundation for Polish Science Projects Start, Homing Plus and TEAM co-financed by the EU within the European Regional Development Fund.

REFERENCES


[8] Bureau International des Poids et Mesures (BIPM) Recommended Values Of Standard Frequencies For Applications Including The Practical Realization Of The Metre And Secondary Representations Of The Definition Of The Second, Strontium 87 Atom ($f \approx 429$ THz), (BIPM, Sèvres, France, 2013)

[9] Bureau International des Poids et Mesures (BIPM) Recommended Values Of Standard Frequencies For Applications Including The Practical Realization Of The Metre And Secondary Representations Of The Definition Of The Second, Strontium 88 Atom ($f \approx 429$ THz), (BIPM, Sèvres, France, 2009)


A Magnetometer Based on Coherent Population Beating

LIU Li, WANG Yigen, ZHAO Xiaona, ZHUANG Yuxin and WANG Zhong
School of Electronics Engineering & Computer Science
Peking University
Beijing, 100871, P. R. China
E-mail: zw@pku.edu.cn

Abstract- We proposed a novel magnetic field measurement method extended from the CPT method. It is based on the coherent population beating (CPB) phenomenon. CPB occurs in a typical three-level system, when the frequency difference of the two pump laser fields have a detuning from the ground states splitting, and the CPB oscillation frequency is equal to the detuning. We are able to detect the beat frequency shift with the external magnetic field changing via digital processing, thus we can acquire the Zeeman frequency shift and then calculate external magnetic field intensity accurately.

Keywords—magnetometer; CPT; CPB; equal-precision measurement

I. INTRODUCTION

The measurement of magnetic field is being applied in many fields, such as geophysical surveying, industrial detection, medical diagnoses and so on. To obtain high sensitivity and accuracy, atom magnetometers based on quantum effects are widely used, among which the CPT magnetometer shows a great advantage. Its sensitivity is up to 10 pT, and it can be integrated to chip-scale (NIST, 2004).

Here we proposed a novel magnetometer based on a new phenomenon –coherent population beating (CPB). This CPB magnetometer can not only be digitized and integrated, but also can reach a theoretical sensitivity up to 1 pT. In experiment we have proved a good linear relationship between magnetic field intensity and oscillation frequency.

II. THEORY

A. Coherent Population Beating

The coherent population beating (CPB) is observed when two coherent optical fields transmit through an alkali atomic cell [1,2]. As is shown in Fig.1, CPB occurs in a typical three-level system, when the frequency difference (ω_{21}) of the two pump laser fields is equal to the ground states splitting (Δ_{21}), it’s a typical coherent trapping phenomenon. However, when ω_{21} has a detuning from Δ_{21}, a damping oscillation will be observed from the photodiode, and the oscillation frequency is equal to the detuning |Δ_{21}-ω_{21}| [3,4].

The CPB phenomenon enables us to directly obtain the beat frequency between the RF signal and the atomic transition frequency.

B. CPB in magnetic field

As is shown in Fig.2, for m_{F}=0 ground state hyperfine energy levels, they are insensitive to external magnetic field, which can be used to achieve a CPB atomic clock [5]. Here we move the CPB effect to the ground levels of m_{F}=1 or m_{F}=-1, which are sensitive to external magnetic field. When the atoms are placed in a magnetic field, the Zeeman sublevels of the atoms will shift with the magnetic field intensity changing. Then we are able to detect the beat frequency shift with the external magnetic field changing via digital processing, thus we can acquire the Zeeman frequency shift and then calculate external magnetic field intensity accurately..
III. SENSITIVITY ANALYSIS

In the procedure of frequency measurement, the CPB oscillation frequency Δω = |Δω₂₁ - ω₂₁|, where Δω₂₁ is the ground levels hyperfine splitting and ω₂₁ is the difference of the coherent laser fields. The CPB effect enables us to take the hyperfine splitting frequency as reference and convert the measurement to a low frequency region, which is crucial to data sampling and processing. So far we have two ways to measure the oscillation frequency – one way is to use Fast Fourier Transform (FFT) and the other one is to use equal-precision measurement based on FPGA. Here we take the FPGA measurement way as an example.

Fig. 3. Sketch map of the CPB frequency equal-precision measurement.

As is shown in Fig.3, the detected CPB signal will be shaped into a square wave with the same frequency. If we periodically excite the oscillation and add a number N of T₉₆s together, the measurement error rate is

\[ \delta \leq \frac{1}{NT_{96}t_9} \]  

In our system, NT₉₆ will be about 0.2 second and f₉ is 3 MHz, so the corresponding frequency resolution is 5x10⁻⁵ Hz. In this way the beat frequency Δω can be accurately measured through digital signal processing, which is capable of up to mHz or higher frequency resolutions (for GHz signal). Taking ⁸⁷Rb for example, the gyromagnetic constant is 7 Hz/nT, so according to Δν=γB, the magnetic field intensity resolution can reach up to 1pT.

IV. EXPERIMENT SETUP

Based on CPB phenomenon, a magnetometer system can be implemented. The experiment setup is illustrated in Fig.4.

A modulated VCSEL is used to excite the CPB phenomenon and the modulated light generated by the VCSEL is injected into an ⁸⁷Rb cell. The 795nm VCSEL laser is locked to the m₉=1 transition of D1 absorption spectrum of ⁸⁷Rb. The radio frequency, referencing to an OCXO, is used to modulate the laser frequency through a bias-tee.

The intensity through the Rb cell is detected by a photodiode and converted into electrical signal. When the magnetic field intensity (related to the coil current) is changed, Δω₂₁ is consequently changed and so is the detuning Δ. According to the CPB theory, we will get a string of oscillations. If we are able to measure the oscillation frequency accurately, we can acquire the magnetic field intensity change. If we can make sure that the OCXO is stable enough and does not contribute to the frequency shift of detected signal, we are able to make a CPB magnetometer with excellent performance.

Fig. 4. The experiment setup of CPB magnetometer system. The RF frequency from the frequency synthesizer is modulated by a square wave and the frequency periodically changed between f₀ and f₀-Δω.

V. RESULT AND CONCLUSION

In experiment we are not able to control the coil current as well as the magnetic field precisely enough yet. But we have preliminary proved a good linear relationship between coil current (which is proportional to magnetic field intensity) and oscillation frequency (Fig.5), showing that it’s a feasible way to measure magnetic intensity.

Fig. 5. The relationship between coil current (which is proportional to magnetic field intensity) and CPB oscillation frequency.

In conclusion, we have proposed a novel magnetometer based on coherent population beating. By analyzing the sensitivity of equal-precision measurement, we consequently analyzed the sensitivity of this novel magnetometer. In experiment we have preliminary proved the linear relationship between coil current and oscillation frequency, thus we can measure small changes in magnetic field in this way. The magnetometer based on CPB is easy to digitize and it has a good theoretical sensitivity performance. We believe it will be a competitive scheme for magnetic field intensity measurement.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China. We would like to thank Guo Tao for his discussion and Wan Mingyu for her related work.
REFERENCES


Monitoring the Adhesion Process of Tendon Stem Cells using Shear-Horizontal Surface Acoustic Wave Sensors

Huiyan Wu, Hongfei Zu, Qing-Ming Wang
Dept. of Mechanical Engineering & Materials Science
University of Pittsburgh
Pittsburgh, PA 15261, USA
Email: qiw4@pitt.edu

Guangyi Zhao, James H-C. Wang
MechanoBiology Lab, Dept. of Orthopaedic Surgery
University of Pittsburgh
Pittsburgh, PA, USA

Abstract—Cell adhesion to a substrate or extracellular matrix (ECM) plays an important role in a variety of cellular functions, such as cell migration, proliferation, differentiation, and tissue formation. Shear-horizontal surface acoustic wave (SH-SAW) sensors can detect cell behaviors in liquid in a non-invasive, simple and quantitative manner. As the key part of SH-SAW, acoustic-wave guiding layer plays a crucial role in improving sensor performance. Parylene-C (poly(2-chloro-p-xylene)) has been proven as ideal guiding layer due to its good uniformity, compactness and adhesion to substrate. Of comparable cell and protein compatibility to the tissue culture substrate, parylene-C films also have preferable effects as the bio-sensitive interface on SH-SAW sensors. In this study, SH-SAW sensor with parylene-C acoustic-wave guiding layer was adopted to monitor the adhesion process of tendon stem cells (TSCs), a newly discovered stem cell type in tendons. TSC suspensions of different concentrations (0.5×10^5, 1.0×10^5, 2.0×10^5, 4.0×10^5 cell/ml) were added to collagen-coated PDMS wells successively. The cells were maintained in the incubator for 10 hr, during which corresponding S_{21} spectrums were recorded every 1 min. The results indicated that there was a sharp increase in S_{21} loss in the beginning of incubation. With incubation continued, the increase rate reduced gradually, and S_{21} loss tended to be stable. S_{21} phase decreased continuously at first, and then entered a plateau with continued incubation. These changes are considered to be related to the integrin-ECM protein interactions and focal adhesion formation occurring in TSC adhesion process. In addition, as TSC suspension concentration increased, the final value of S_{21} loss change due to TSC adhesion was increased. SH-SAW sensors exhibit high sensitivity and stability in TSC adhesion monitoring, indicating their potential for investigating cell biology in general and cell adhesion in particular.

Keywords—SH-SAW; TSC; Parylene C; Cell Adhesion

I. INTRODUCTION

Tendon is fibrous connective tissue that transmits muscular force to bone thus enabling joint motion and body movement. Tendon had been considered to contain only tenocytes for a long time. However, some recent studies demonstrated that human, rat and mouse tendons also contain multi-potent adult stem cells, termed tendon stem/progenitor cells (TSCs) [1, 2]. Compared with tenocytes, TSCs exhibit distinct properties in various aspects, including cell morphology, marker expression, as well as proliferation and differentiation potential [3, 4].

Presently, most studies referring to cell adhesion are often based on the observation of cell morphologies using optical microscopy, which is time consuming, labor intensive and qualitative in measurements. In order to achieve a simple and fast monitoring, some other measurement methods have been

Previous research revealed that appropriate stimulus could direct cellular differentiation and promote extracellular matrix (ECM) development [5, 6]. Likewise, in proper circumstances differentiation of TSCs into active tenocytes could be stimulated as well, which is beneficial for maintaining tendon homeostasis. Nevertheless, excessive, improper stimulation is detrimental, as it induces differentiation of TSCs towards non-tenocyte lineages of cells, such as adipocytes, chondrocytes and osteocytes [7]. Rui et al reported that repetitive tensile loading increased the expression of BMP-2, and in addition, could enhance osteogenic differentiation of TSCs [8]. Therefore, this distinct potential of differentiation and proliferation of TSCs would be induced in certain circumstances, causing favorable or unfavorable effects on tendon functions. In order to improve the possibility of using TSCs and repair or regenerate injured tendons effectively, TSCs' behaviors, including their migration, differentiation and proliferation should be examined in detail.

Cell adhesion to a substrate or ECM plays an important role in a variety of cellular activities, such as cell migration, proliferation, differentiation, and tissue formation [9]. Cell-cell or cell-substrate adhesive interactions govern cell aggregation, polarity and migration. In addition, such interactions could alter the differentiation of cells by changes in cell morphology, proliferation, as well as programs of gene expression [10]. Generally speaking, the procedure of cell adhesion has three stages: attachment, spreading, and formation of focal adhesions and actin-containing stress fibers [11]. Focal adhesions, or focal contacts, are the localized points of attachment between cell surface and substrate [9]. For adherent cells on the substrate, focal adhesion contains clustered adhesion receptors, integrins, which span the plasma membrane and attach to other proteins both outside and inside the cell: to the outside, integrins attach to ECM proteins; to the inside, they interact with bundles of actin microfilaments via a few linker proteins. This integrin-dependent attachment enhances cell adhesion to the substrate, functioning as not only structural links between the cytoskeleton and ECM, but also triggering signal pathways that direct cell functions.

This work is supported in part by AR060920 and AR061395 (JHW).
developed for real-time detection of cell adhesion, including electric cell-substrate impedance sensing (ECIS) and Quartz thickness shear mode (TSM) acoustic wave sensing. ECIS measurement is based on electrical impedance change induced by cell attaching and spreading onto the central gold electrodes, which provides quantitative information on cell morphologies and motion of cultured cells in real time [12]. However, for electrodes are immersed in culture medium, adherent cells are required to restrict the current flow through the medium effectively to induce a detectable impedance change. As a result, there would be no reliable response from ECIS if seeded cells are of relatively low concentration and they could not cover the entire electrode surface. Without such limitation as ECIS, acoustic wave sensors show their advantages on an easy, real-time, quantitative measurement of cell activities. As the most-widely-used acoustic wave sensor, Quartz TSM resonator is also applied in some research on different types of cells [13-16]. In our previous studies, the effect of aging on viscoelastic properties of TSCs were examined through TSM, indicating an overall increase in both storage and loss shear modulus during aging process [17]. Shear-horizontal surface acoustic wave sensor (SH-SAW, Love Mode) is considered one of the most promising probing methods in fundamental biology and biomedical engineering [18, 19]. They could detect cell behaviors in liquid in a simple, non-invasive, and quantitative manner, avoiding direct contact between electrodes and liquid medium [20, 21]. Furthermore, acoustic waves excited by surface electrodes in SH-SAW performed extremely sensitivity to certain surface perturbations as well as favorable inertness to other surrounding factors, especially to the number of bonds formed within the relatively short distance of ~50 nm from the surface [22, 23]. Compared with TSM resonator, the response of SH-SAW sensor could provide more information on the interactions between cells and the substrate during adhesion process, while environmental interferences are decreased or even eliminated.

As the key part of SH-SAW, acoustic-wave guiding layer plays a crucial role in improving sensor performance. Different kinds of piezoelectric and nonpiezoelectric materials have been applied as guiding layer in the past years, such as ZnO, SiO2 and PMMA [24-27]. Nevertheless, because of a lack of proper bio-interface, most studies on SH-SAW sensors neglected their innovative applications as cell-based sensors. Parylene-C (poly(2-chloro-p-xylene)) has been proven as an ideal acoustic-wave guiding layer due to its good uniformity, compactness and adhesion to substrate [28]. More importantly, parylene-C films possess comparable cell and protein compatibility to the standard tissue culture substrate, indicating their great potentiality as a bio-sensitive interface [29]. In this study, SH-SAW sensor with parylene-C acoustic-wave guiding layer was adopted to monitor the adhesion process of TSCs. TSC suspensions with various concentrations (0.5×10^5, 1.0×10^5, 2.0×10^5, 4.0×10^5 cell/ml) were added into PDMS wells successively. The cells were maintained in wells for 10 hr, during which corresponding S21 spectrums were recorded every 1 min. Normalized S21 loss and phase changes extracted from S21 spectrums were considered to be related to the integrin-ECM protein interactions and focal adhesion formation occurring in adhesion process. SH-SAW sensors show high sensitivity and stability in TSC adhesion monitoring, indicating their potential for investigating cell biology in general and cell adhesion in particular.

II. MATERIAL PREPARATION & EXPERIMENT DETAILS

A. Isolation and culture of TSCs

In this study, TSCs were isolated from the Achilles tendons of young rats (8 week, 200–250 g). Briefly, the middle portion of tendons was retrieved by cutting the tendon samples 5 mm from the tendon-bone insertion and tendon-muscle junction. After removing the tendon sheath and surrounding paratenon, the middle tendon portion tissues were weighed and minced into small pieces (1 mm × 1 mm × 1 mm). Each 100 mg tissue sample was digested with 3 mg collagenase type 1 (Life Technologies, Carlsbad, CA, USA) and 4 mg dispase (Stemcell Technologies, Vancouver, BC, Canada) in 1 ml phosphate-buffered saline (PBS, Mediatech Inc., Manassas, VA, USA) at 37°C for 1 hr. The suspensions were centrifuged at 1500 G for 15 min, and the cell pellet was resuspended in growth medium consisting of Dulbecco’s modified Eagle’s medium (DMEM, Lonza, Walkersville, MD, USA) containing 20% fetal bovine serum (FBS, Atlanta Biologicals, Lawrenceville, GA, USA), 100 U/ml penicillin, 100 μg/ml streptomycin, and 2 μM glutamine (Thermo Fisher Scientific Inc., Waltham, MA, USA). A single-cell suspension was obtained by diluting the suspension to 1 cell/μl and then cultured in T25 flasks at 37°C with 5% CO2. After 8–10 days in culture, colonies formed as the Achilles TSCs were visible on the surface of culture flask. Individual cell colonies were collected using trypsin (Life Technologies, Carlsbad, CA, USA) and transferred to separate T25 flasks for further culture.

B. Measurement and characterization of TSCs

Fig.1 schematically shows the measurement system of SH-SAW sensor for TSCs. Cr/Au interdigital-transducer (IDT) electrodes of a dual-delay-line configuration were fabricated on 36º YX-LiTaO3 surface by photolithography. The period of IDTs is 32 μm, leading to an operating frequency around 131 MHz. Parylene-C films, as the acoustic-wave guiding layer, were deposited on these electrodes by thermal deposition system (SCS PDS 2010 Labcoter 2, SCS Equipment, Dallas, TX): the original dimer is sublimated and then undergoes pyrolytic cleavage at 690 °C; the object monomer polymerizes again on LiTaO3 surface at room temperature in vacuum, forming continuous parylene-C films. The average thickness of parylene-C is 1.6 μm, providing a good combination of both high sensitivity and low energy loss. One rectangular tube (12 mm × 10 mm × 6 mm) of poly(dimethylsiloxane) (PDMS, Dow Corning, Midland, MI, USA) was located around the upper electrode, forming the well-like structure for TSC culture. The S21 loss curves were obtained by an Agilent 34970A switch/control unit was series connected for converting among different sensors. Before real-time measurement, SH-SAW sensors with empty wells were firstly exposed to UV light overnight to sterilize. Subsequently, 400 μl 100 μg/ml collagen type I/PBS solution (Stemcell Technologies, Vancouver, BC, Canada) was added into each well and kept for 2 hr in order to cover the entire
electrode surface with collagen layer, as is commonly conducted in tissue culture. After removing the collagen/PBS solution and washing these wells with PBS three times, 400 μl TSC suspensions of different concentrations (0.5×10⁵, 1.0×10⁵, 2.0×10⁵, 4.0×10⁵ cell/ml) from rats were added. SH-SAW sensors were maintained in the incubator at 37°C with 5% CO₂ for 10 hr, and corresponding S₂¹ loss curves were detected every 1 min. Meanwhile, the S₂¹ loss curves of SH-SAW with pure culture medium were recorded as well.

In parallel experiments, TSC suspensions of identical concentrations were cultured in collagen-coated 96-well culture plate and PDMS wells on Au/Si wafer, respectively. The morphologies of TSCs in culture plate were observed through phase-contrast microscopy (Eclipse TS100, Nikon, Japan) at a series of time points (0, 1, 2, 4, 6, 10 hr). TSCs on Au/Si wafer were fixed with 4% paraformaldehyde in PBS after 10 hr, and then washed with alcohol solution of increasing concentrations (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and 100%) successively. Their morphologies were examined by scanning electron microscopy (XL-30 FEG SEM, FEI/Philips, Japan, accelerating voltage: 5 kV).

III. RESULTS & DISCUSSIONS

Fig.2 (a), (b) present the typical S₂¹ spectrum sets of SH-SAW sensor: (a) was acquired after adding TSC suspension in culture medium (1×10⁵ cell/ml) into PDMS well, and (b) represents the addition of culture medium only. The spectrum recording lasted 10 hr. As the time goes on, the color of S₂¹ spectrum is shifting from pink to black gradually. It is illustrated that with culture medium only, the intensity of resonance peak around 132.45 MHz decreased in the beginning, and then was inclined to be stable over time. Nevertheless, when TSC suspension was added, due to TSCs’ adhesion to the surface, the intensity of resonance peak exhibited a reverse increase following a slight decrease. Based on these two sets of S₂¹ spectrums, the normalized S₂¹ loss change curve for TSC adhesion process was obtained as shown in Fig.2 (c) (the pink dot curve), and normalized S₂¹ phase change curve was also plotted in pink dots in Fig.2 (d) accordingly. Fig.2 (c), (d) show the S₂¹ loss and phase variations from three independent measurements, respectively. It is indicated that both loss and phase change curves of three trials exhibited a similar tendency. There was a sharp increase in S₂¹ loss during the first 4 hr incubation (I). With continued incubation (II), the increase rate reduced gradually, and S₂¹ loss tended to be stable. S₂¹ phase decreased continuously in the initial 4 hr incubation (I), and then entered a plateau with the incubation time ascending further (II). The variations in S₂¹ loss and phase are due to surface perturbations induced by TSC adhesion. In order to get more details concerning interactions between TSCs and the substrate, morphologies of TSCs in the entire process were detected by phase-contrast microscopy.

![Fig.1 Schematic diagram of SH-SAW sensor measurement system](image1)

![Fig.2 Response of SH-SAW sensors during TSC adhesion (1×10⁵ cell/ml): (a) S²¹ spectrum set of SH-SAW sensor with addition of TSC suspension (b) S²¹ spectrum set of SH-SAW sensor with pure culture medium addition (c) Normalized loss change of SH-SAW (d) Normalized phase change of SH-SAW](image2)

![Fig.3 shows the microscopy photos of TSCs taken at a series of time points after 100 μl cell suspension of 1×10⁵](image3)
cell/ml was added into 96-well culture plate (the cell number per unit area in culture plate was designed to be same as in SH-SAW sensor). (a) - (f) were taken after 0, 1, 2, 4, 6 and 10 hr in order, respectively. As discussed before, the procedure of cell adhesion has three stages: attachment, spreading, and formation of focal adhesions and stress fibers. The stage of cell attachment, which involves integrin-ECM protein interactions, is considered as an intermediate state between weak contact and strong adhesion. The attached cells increase their surface contact area with the substrate through cell spreading and the formation of actin microfilaments, and then organize their cytoskeleton by the formation of focal adhesions and actin-containing stress fibers. In this study, the SH-SAW surface was pre-coated with collagen, which is considered a primary and structural part of ECM in tendons [9]. When suspending cells are approaching the substrate, these ECM proteins are bound by integrin receptors composed of α/β chains, initializing the cell attachment at first [10, 30]. The TSCs suspending in culture medium possessed a relatively spherical shape as shown in (a). When added into wells, TSCs began to go down and landed on the substrate. After 1 hr, as shown in (b), half of suspending TSCs had attached to the substrate (spherical ones with bright edges), and quite a few of them began to spread (flat black ones). After 2 hr, over 80% TSCs had attached to the substrate and began to spread (c). When incubation exceeded 4 hr, over 95% TSCs had settled down on the substrate (d). For attached TSCs, the first change that could be recognized was a transformation from a spherical to a flatter discoid shape. Then with these cells spreading onto the surface, significant morphological changes occurred. Through 4 hr incubation, a preliminary enlarged, polygonal cell shape could be observed on part of TSCs. With continued incubation, TSCs kept spreading and formed focal adhesions. They started to extend along specific directions but not in random directions as in the beginning. The morphology variation process was almost finished 10 hr later. It can be seen from (f) that all adherent TSCs possessed a characteristic elongated shape, exhibiting their high consistency and uniformity. With cell suspension concentration of 1×10^5 cell/ml, TSCs could form a relatively-uniform cell monolayer, and there was no obvious overlap or blank space on the substrate. In addition, SEM photos of TSCs (not shown) are in good agreement with phase-contrast photos. It can be seen that TSC monolayer was homogeneous and compact over the entire surface of the well bottom.

S21 loss and phase changes relate well to the corresponding morphology variations of TSCs in Fig.3. It can be illustrated that during the initial 4 hr incubation, due to the formation of integrin-ECM protein bonds across the entire substrate, there was a sharp increase in S21 loss change. With continued incubation, over 95% TSCs had been attached to the substrate. In this case, S21 loss change was mainly induced by the formation of focal adhesions, and as a result its increase rate was reduced gradually. As the adhesion process of TSCs was completed, loss change was inclined to be stable. In the meantime, it is found that the dominant factor of S21 phase change was the interactions in TSC attachment, for most changes occurred in the initial 4 hr. When over 95% TSCs had attached to the substrate and integrin-ECM protein bonds had formed extensively, the phase change curve entered a plateau. In comparison, the same adhesion process of TSCs was also monitored by TSM. Admittance spectrums of TSM resonator with TSC suspension were acquired every 1 min, and corresponding resonance frequencies were extracted from them (not shown). In the beginning there was a significant decrease in TSM resonance frequency as a great number of TSCs falling down and attaching to the electrodes. With more and more TSCs finishing attachment, TSC spreading and focal adhesion formation became the predominant activities, resulting in a reverse increase of TSM resonance frequency. It is revealed that compared with TSM, the trigger of SH-SAW response is mainly limited to the formation of bonds between TSCs and collagen-coated substrate, not including the weak contact in the beginning of the first stage. Due to the larger decay distance, the response of TSM might be affected by more unexpected surrounding factors. SH-SAW sensors exhibit high sensitivity to surface perturbations, especially bonds formed within the relatively short distance of ~200 nm from the surface, leading to an affirmative advantage on investigating subtle interactions between cell membrane and a substrate involved in a series of cellular activities.

![Fig.3 Phase-contrast microscopy photos of TSCs during TSC adhesion](image)

When TSC suspensions of various concentrations (0.5×10^5, 1.0×10^5, 2.0×10^5, 4.0×10^5 cell/ml) were added, corresponding electrical response of SH-SAW sensors was shown in Fig.4: (a) is normalized loss change curves extracted from S21 spectrum sets during TSC adhesion process; and (b) is normalized phase change curves. The spectrum recording endured for 10 hr. It is shown that both S21 loss and phase changes presented similar variation tendency for different concentrations. There was a fast increase in all of the S21 loss curves in the beginning of incubation. With longer incubation, the increase rate decreased to be zero gradually, and S21 loss tended to be stable. The main differences among different concentrations are merely on the duration of fast growth, as well as the magnitude of loss
variation. For TSC concentrations of $0.5 \times 10^5$, $1.0 \times 10^5$ cell/ml, the fast growth lasted approximately 4 hr. As the concentration increased to $2.0 \times 10^5$, $4.0 \times 10^5$ cell/ml, the fast-growing region was decreased to about 2 hr and 1 hr, respectively. At the same time, the growth rate in this period was significantly increased compared with that of low concentrations. This might be due to the increased bonds forming when TSCs were attaching to the substrate. As shown in Fig.3 (d), for $1.0 \times 10^5$ cell/ml, the TSCs hadn’t formed uniform cell monolayer when cell spreading just began, and a considerable proportion of substrate was available for cell attachment. Therefore, in case of high concentrations, a larger number of TSCs could form integrin-ECM protein bonds in the beginning. However, when some TSCs began to spread, there was limited space for other suspending TSCs attaching, which resulted in a higher increase rate and an earlier end of fast growing. All $S_{21}$ loss curves could be found to go into a plateau at almost the same time. The final value of loss change increased as the concentration of TSC suspension increased. As discussed above, due to the high TSC concentration, more TSCs would attach to the substrate in the first few of hours of incubation. With continued incubation, even though spreading of cells would be limited by each other to some extent and cell multilayer structure would form, the number of adherent TSCs per unit area in the nearest layer was still increased. The focal adhesions formed would be relatively denser, leading to a more compact binding structure on the substrate. In the meantime, although not as clear as $S_{21}$ loss curves, all of the $S_{21}$ phase change curves exhibited similar decreasing tendency. For TSC suspension of higher concentration, the duration of decreasing was shortened to about 2 hr. Furthermore, it could also be seen that the higher the TSC suspension concentration, the larger the $S_{21}$ phase variation from the start point to the final plateau value. These results indicates, even further, that $S_{21}$ phase change was mainly induced by the integrin-ECM protein bond formation in TSC attachment process.

IV. CONCLUSION

SH-SAW sensor with parylene-C acoustic-wave guiding layer was proposed to monitor the adhesion process of TSCs, a newly discovered stem cell type in tendons. TSC suspensions of a series of concentrations ($0.5 \times 10^5$, $1.0 \times 10^5$, $2.0 \times 10^5$, $4.0 \times 10^5$ cell/ml) were added and maintained in PDMS wells for 10 hr, during which corresponding $S_{21}$ spectrums of SH-SAW were acquired every 1 min. For TSC suspensions of different concentrations, both normalized $S_{21}$ loss and phase changes presented similarly in variation tendencies. There was a fast increase in $S_{21}$ loss in the beginning of incubation. As the incubation time increased, the increase rate reduced gradually, and $S_{21}$ loss tended to be stable. $S_{21}$ phase decreased continuously at first, and then entered a plateau with continued incubation. These changes are considered to be related to the integrin-ECM protein interactions and the formation of focal adhesions occurring in adhesion process of TSCs. Furthermore, as the concentration of TSC suspension increased, the final value of $S_{21}$ loss change due to TSC adhesion was increased. SH-SAW sensors exhibit high sensitivity and stability in TSC adhesion monitoring, indicating their potential for investigating cell biology in general, and cell adhesion in particular.

REFERENCES


A Wireless Temperature Sensor Powered by a Piezoelectric Resonant Energy Harvesting System

Peng Wang1*, Robert Gray1,2, Zenghui Wang1, Philip X.-L. Feng1*

1Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH 44106, USA
2Hawken High School, Gates Mills, OH 44040, USA
*Email: peng.wang9@case.edu, philip.feng@case.edu

Abstract—We report on experimental demonstration of a wireless temperature sensor node (WTSN) powered by a piezoelectric resonant energy harvesting system. The energy harvesting circuit stores the energy generated by a piezoelectric resonant transducer into a capacitor, and uses the stored energy to power the temperature sensor and its associated signal processing circuits for wireless signal transmission. The main functions of the harvesting system are implemented by discrete components together with a power management application specific integrated circuit (ASIC). The resulting WTSN transmits measured temperature data over a distance of 10m, in real time, and consumes ~4µW to ~13µW power (for transmission intervals from 10min to 10s, respectively), which is supplied entirely by the piezoelectric resonant energy harvester.

Keywords—wireless temperature sensor; energy harvesting; piezoelectric resonator; transducer; lead zirconate titanate (PZT)

I. INTRODUCTION

Wireless temperature sensor networks are key to energy-efficient smart buildings for more sustainable and resource-conserving development [1]. However, battery-powered large-scale temperature sensor networks have high maintenance costs due to required periodical replacements of batteries, which limit their real-world applications and deployments. Vibration energy harvesting offers a viable alternative to battery-powered wireless temperature sensor networks, as it enables self-powered sensor nodes which shall no longer require battery replacements, and thus can lead to maintenance-free systems.

Wireless sensor networks based on 802.15.4/ZigBee are promising for smart building applications, and the wireless protocol facilitates flexibility, interoperability, and large number of sensor nodes. A typical 802.15.4/ZigBee sensor network consists of front-end sensor devices (wireless sensor nodes), repeaters (wireless relays), and coordinators (wireless hubs). The repeaters and coordinators are usually powered through mains and have high operating power (thus high receiver sensitivity), which can in turn reduce the power consumption and duty cycle on the front-end devices (transmitters). The front-end sensor devices require ultra-low power budget design, which, together with vibration energy harvesting, can realize self-powered sensor nodes.

Fig. 1 illustrates the power budget analysis for a wireless sensor node. The power consumption range of a typical wireless front-end sensor device is ~5µW to ~1mW [2-5], based on transmission intervals (gaps) of 10s to 10min. The dominant power consumption of a wireless temperature sensor node (WTSN) comes from RF communication, while the averaged power for temperature reading is usually only hundreds of nW. Fig. 1 shows that with the sensor power consumption level under the current technology, battery replacement is inevitable on a yearly basis, if not monthly. Hence to extend the maintenance-free working span of a wireless sensor, self-powering solutions such as vibration energy harvesting are required.

Here we report on an experimental demonstration of using a piezoelectric resonant energy harvesting system to power a WTSN. We customize a PZT transducer, which resonates at the same frequency (120Hz) with the vibration source (a small service pump, or devices alike in a regular ventilation service room). We use discrete components and an ASIC prototype for the energy conversion and storage functions. With this energy harvesting system, we successfully demonstrate a WTSN, which is capable of transmitting measured temperature over a distance of up to 10m, and consumes ~13µW (transmitting at 10s interval) or 4µW (at 10min interval) power, all supplied by the vibration energy harvester.

II. SYSTEM DESIGN

A. System Block Diagram

Fig. 2 shows the block diagram of the WTSN system. It includes a PZT resonant energy harvester, a full-wave-bridge rectifier, an ASIC prototype, and an RF module with a microcontroller integrated with a temperature sensor. The full-
wave rectifier is built with Schottky diodes, with a total voltage drop of 0.7V. A 50nF capacitor is used as a filter, and the harvested energy is stored in a 100µF capacitor.

Fig. 2: System diagram of the wireless temperature sensor node (WTSN).

B. PZT Transducer Design

The PZT transducer takes the form of cantilever-shaped resonator (15mm×12mm×200µm), laser machined to have an optimized dimension. The PZT part extends 2/3 of the total cantilever length for minimizing the undesirable charge redistribution effect [6, 7]. Once made, the PZT transducer is characterized using a 3-axis accelerometer (ADXL325, Analog Devices, Inc.). We then fine-tune the resonance frequency of the PZT cantilever (Fig. 3) to match that of the vibration source through adjusting the proof mass, a piece of low-melting-temperature metal attached at the end of the cantilever. Upon fine tuning, the optimized PZT transducer can output up to 96µW (measured with a 51kΩ load resistor) with a peak acceleration of 1.5g at the pump surface.

![Fig. 3: PZT transducer mounted on a surface. (a) Front view. (b) Side view.](image)

C. Wireless Network

Wireless connection (over ZigBee network) is established between the WTSN (front-end device) and a receiver (access point, AP) connected to a data-taking computer.

D. Wireless Transmission Range

To estimate the wireless transmission range, we calculate the RF signal attenuation over distance, wireless link budget, and power budget on the sensor node.

The wireless signal attenuation can be generally described by the Friis equation: the received power \( P_R \) at a distance \( d \) is:

\[
P_R = P_T G_T G_R A_d \lambda^2 / (4\pi d^2).
\]

Here, \( P_T \) is the originally transmitted power at the source, \( G_R \) and \( G_T \) are the gains of the transmitting (TX) and receiving (RX) antenna respectively, \( \lambda \) is the EM wavelength, and \( n \) is an environmental parameter (\( n = 2 \) in free space; here we use \( n = 2.2 \), typical value for most indoor settings). In our experiment, both the transmitter and receiver (eZ430RF2500 and CC2500) have an average gain of 1dBi, the receiver sensitivity is -104dBm at 2.4kBaud 1% packet error rate, and \( \lambda = 0.1227 \text{m} \) at 2.445GHz. Fig. 4a shows wireless range \( d \) vs. the transmitter output power \( P_T \) as calculated using Eq. (1), with available settings on the CC2500 indicated by square dots.

The link budget is analyzed by calculating the difference between the received signal power, \( P_R \), and the sensitivity of the receiver. In a practical design of link budget, additional output power has to be added to the output power predicted by Friis equation. Sufficient link budget can reduce the RX power consumption and data packet loss rate. Fig. 4b shows the link budget for a typical indoor transmission distance \( d = 10 \text{m} \).

Fig. 4: Wireless transmission analysis results. (a) Calculated wireless transmission range as a function of transmitter output power. (b) Calculated RF link budget for wireless transmission over 10m using CC2500. The dots represent available configurations in the CC2500.

E. Temperature Sensor

The temperature sensor used in this study is a silicon (Si) bandgap temperature sensor integrated in the microcontroller (MSP430F2274) of the eZ430RF2500 module, and is read by an on-chip 10-bit ADC converter.

III. EXPERIMENTAL TECHNIQUES

A. PZT Transducer Characterization

We characterize the PZT transducer in both frequency and time domains. In the frequency domain measurement, the transmission loss is measured with a network analyzer (Agilent 4395A) as shown in Fig. 5a. In the time domain measurement, we deflect the cantilever by a fixed amount, and then suddenly release it. The PZT cantilever then undergoes a ‘ring down’ oscillation which decays over time, and its voltage output is measured with an oscilloscope (Fig. 5b).

![Fig. 5: Schemes for characterizing the PZT cantilever. (a) Frequency domain measurement. (b) Time domain measurement.](image)
B. PZT Harvester Assembly and Installation

The complete WTSN is shown in Fig. 6a, and the inset picture shows the prototype ASIC die. The fully assembled WTSN, once characterized, is mounted on the pump surface as shown in Fig. 6b. The PZT transducer energy harvester is fixed on the pump using 7036 Blanchard Wax (J.H. Young Company) for optimized mechanical coupling.

![Image of System Assembly](image)

**Fig. 6:** System assembly. (a) The WTSN. (b) WTSN mounted on the pump.

C. WTSN Power Consumption Characterization

We characterize the transient power consumption of the eZ430RF2500 front-end device. As shown in Fig. 7, we power the WTSN with a battery pack, and connect a 68Ω resistor in series. This allows us to estimate the current (and thus power consumption) in the WTSN by using a digital oscilloscope (TDS1012C-EDU, Tektronix) to monitor the voltage drop across the resistor.

During the characterization, the WSTN is in the normal operation mode. It reads the environmental temperature with the built-in temperature sensor, and sends the readings wirelessly to the wireless access point (AP). While the PZT transducer exemplified as a sharp ‘heart beat’ curve feature on top of a slowly-varying frequency dependent back ground. The data show that our fine tuning results in good frequency matching of the cantilever to the vibration source (120Hz).

The result from the time domain measurement is shown in Fig. 8b. The data are fit to the ring-down curve of a damped harmonic oscillator:

\[
a(t) = a_0 + A \exp \left(-\frac{\pi f_{res}}{Q}\right) \sin(2\pi f_{res}t + \phi),
\]

where \(a(t)\) is the time-dependent amplitude of the vibration, \(a_0\) is the offset of the response, \(A\) is the initial amplitude, \(f_{res}\) is the resonance frequency, \(Q\) is the quality factor, and \(\phi\) is the initial phase. From the fitting we extract the resonance frequency \(f_{res}\) and quality factor \(Q\) of the PZT cantilever. The results show excellent agreement with the frequency domain measurement, and again verify the precise tuning of the cantilever frequency.

![Image of Characterization](image)

**Fig. 7:** Scheme for measurement of transient power. A small resistor of 68Ω is connected in series with the power input to the WTSN (the wireless end device), which wirelessly communicates with the wireless access point (AP).

**Fig. 8:** Characterization of the resonant PZT cantilever energy harvester. (a) Frequency domain measurement. (b) Time domain data and fitting. Blue Circles: Raw data. Red Curve: Fitting of the data according to Eq. (2).

D. Wireless Temperature Reading

Once the WTSN is characterized, we demonstrate wireless temperature sensing powered entirely by vibration energy harvesting. The experimental setup is the same as shown in Fig. 7, except that the battery pack is replaced by the PZT energy harvesting system, and the current-monitoring resistor (and the oscilloscope) is removed. While the PZT transducer remains fixed on the pump throughout the experiment, the remaining part of the WTSN is tested under different temperature settings and wireless environments.

The different settings/environments include ambient open space in the lab (room temperature controlled by thermostat), on top of a hotplate (HP131725, Thermo Scientific Inc.), underneath a heat lamp, and inside two ovens with metal enclosures (model 1410 and 1410MS, VWR International, LLC.). A separate digital temperature sensor (Caliber IV Digital Hygrometer, Western Humidor, Inc.) is used as a reference. During the experiment, the WTSN wirelessly transmits the temperature reading to the wireless access point (AP) once every 10s, and the data is logged by the data-taking computer connected to the AP device (similar to the scheme during WTSN power consumption characterization, as shown in Fig. 7). The access point is located 10m away from the WTSN for the open space/heat lamp measurements, and 2m away during the oven/hotplate measurements.

IV. RESULTS AND DISCUSSIONS

A. PZT Transducer Characterization

The PZT transducer is characterized using the methods described in Section III-A (and shown in Fig. 5). The results for both measurements are shown in Fig. 8. From the frequency domain transmission/reflection measurement (Fig. 8a), we clearly observe the mechanical resonance of the PZT transducer exemplified as a sharp ‘heart beat’ curve feature on top of a slowly-varying frequency dependent back ground. The data show that our fine tuning results in good frequency matching of the cantilever to the vibration source (120Hz).

The different settings/environments include ambient open space in the lab (room temperature controlled by thermostat), on top of a hotplate (HP131725, Thermo Scientific Inc.), underneath a heat lamp, and inside two ovens with metal enclosures (model 1410 and 1410MS, VWR International, LLC.). A separate digital temperature sensor (Caliber IV Digital Hygrometer, Western Humidor, Inc.) is used as a reference. During the experiment, the WTSN wirelessly transmits the temperature reading to the wireless access point (AP) once every 10s, and the data is logged by the data-taking computer connected to the AP device (similar to the scheme during WTSN power consumption characterization, as shown in Fig. 7). The access point is located 10m away from the WTSN for the open space/heat lamp measurements, and 2m away during the oven/hotplate measurements.
B. Power Consumption of WTSN

The power consumption of the WTSN is measured using the methods described in Section III-C (as shown in Fig. 7). The measured current (and calculated power consumption) during one active transmission event (~30ms long) is shown in Fig. 9a. The peak power is around 55–65mW, and lasts about 2ms. Outside this ~30ms ‘active’ window, the idle power of the WTSN is about 4μW, which sets the ‘baseline’ of the average power consumption. Based on these measured values, we estimate the overall average power consumption of the WTSN with different transmission intervals (Fig. 9b). Example values are also shown for 10s (13μW) and 10min (4μW) transmission intervals, both of which are well within the PZT power output of 96μW (see Section II-B).

![Fig. 9: Calibration of power consumption of the WTSN. (a) Measured current in a WTSN during one transmission event (~30ms). (b) Average power consumption as a function of transmission interval.](image)

C. Wireless Temperature Reading

Experiments of vibration-powered wireless temperature sensing are performed under the settings described in Section III-D. Fig. 10 shows examples of the wirelessly transmitted temperature data. Fig. 10a shows the measurement performed when the temperature sensor is placed inside an oven, with the oven temperature going through four setpoints (26°C, 29°C, 36°C, and 42°C) with 20min duration at each setpoint. During this measurement, the wireless receiver (AP) is located at 2m away from the WTSN. Fig. 10b shows the measurement when the WTSN is placed underneath a heat lamp, showing fast response of the temperature sensor when the lamp is turned on and off. We also increase the heating duration (‘on’ time of the heat lamp) during each heating event, resulting in increasing temperature reading from the WTSN. With the open space setting (no metal enclosure), the wireless transmission distance is increased to 10m in this measurement.

![Fig. 10: Wireless temperature sensing in real time. (a) Temperature sensing from inside an oven. (b) Temperature sensing underneath a heat lamp. Blue Circles: Raw data. Red Curves: Smoothed data. Vertical dashed lines indicate changes in temperature setpoint.](image)

V. Conclusions

In summary, we have demonstrated a wireless temperature sensor node (WTSN) powered by a PZT resonant transducer that scavenges vibration energy and supplies ~10 to ~100μW to the circuits. The system is fully characterized, and the power output of the PZT is sufficient for normal operations of the WTSN. We have further demonstrated real-time wireless temperature sensing under different experimental settings, with the signal transmission over a distance up to 10m.

Fig. 10: Wireless temperature sensing in real time. (a) Temperature sensing from inside an oven. (b) Temperature sensing underneath a heat lamp. Blue Circles: Raw data. Red Curves: Smoothed data. Vertical dashed lines indicate changes in temperature setpoint.

ACKNOWLEDGMENT

We thank Ran Wei, Aman Nair, Xu-Qian Zheng, Wen H. Ko, Steve J. A. Majerus, and Christian A. Zorman for technical support and helpful discussions. We thank support from U.S. Department of Energy (DOE) EERE Grant (DE-EE0006719).

REFERENCES


The Time Validation Facility (TVF): An All-New Key Element of the Galileo Operational Phase

R. Piriz, D. Rodríguez, P. Roldán
GMV
Madrid, Spain
rpiriz@gmv.com

A. Mudrak
ESA/ESTEC
Noordwijk, The Netherlands

A. Bauch, F. Riedel, E. Staliuniene
Physikalisch-Technische Bundesanstalt
38116 Braunschweig, Germany

I. Sesia, G. Cerretto
Istituto Nazionale di Ricerca Metrologica
Turin, Italy

K. Jaldehag, C. Rieck
SP Technical Research Institute of Sweden
Borås, Sweden

P. Uhrich, G.D. Rovera
LNE-SYRTE, Observatoire de Paris, LNE, CNRS, UPMC
Paris, France

Abstract—In the Galileo FOC phase (Full Operational Capability), GMV is the prime contractor for the Time and Geodetic Validation Facility (TGVF), a contract of the European Space Agency (ESA). Within the TGVF, the Time Validation Facility (TVF) is the subsystem in charge of steering Galileo System Time (GST) to UTC, among other duties. The new TVF is operated at GMV headquarters near Madrid, Spain. TVF operations rely on the contribution of five European timing laboratories, located at INRIM, OP, PTB, ROA, and SP. This paper provides a general description of the TVF element and its related activities for the FOC phase, and presents the main results and findings of the TVF operation until now.

Keywords—Galileo, GNSS, system time, UTC steering, validation, ground infrastructure, GGTO, satellite clocks.

I. INTRODUCTION

The Time Validation Facility (TVF) is the TGVF subsystem in charge of determining and assessing timing-related performances of the Galileo system. In particular the TVF supports the validation of the FOC Galileo timing infrastructure, acts as a preliminary Galileo Time Service Provider (TSP) steering Galileo System Time (GST) to UTC, and coordinates the national timing laboratories participating in the operations. For the FOC phase, the TVF has been fully migrated to a new hardware and software infrastructure hosted and operated at GMV headquarters in Tres Cantos near Madrid, Spain. The new TVF started operations in March 2014. In addition to the data from the European UTC laboratories and the Galileo Precise Timing Facilities (PTFs), the TVF makes use of public data coming from different external sources, not part of the TGVF, such as the Bureau International des Poids et Mesures (BIPM), the International Earth Rotation Service (IERS), and the International GNSS Service (IGS). The United States Naval Observatory (USNO) contributes to the validation of the broadcast GPS-to-Galileo Time Offset (GGTO). The TVF needs data also from other TGVF elements: in particular the Orbit Validation Facility (OVF) provides products for satellite clock validation. Fig. 1 shows the TVF interfaces with the external world.

The TVF software has been developed entirely from scratch in C/C++ language, reusing part of the algorithms, experience and results from the previous IOV phase [1].

Fig. 1. TVF interfaces with the external world.
II. GST REALIZATION AND STEERING

GST is the underlying time scale to which the Galileo satellite clock parameters disseminated in the navigation message are referred to. In order to do this, the Orbit and Synchronization Processing Facility (OSPF) in charge of the Galileo message generation at the Galileo Mission Segment (GMS), processes measurements from a calibrated Galileo Sensor Station (GSS) connected to GST. GST as observed through this receiver is used as the time reference for all clock estimates in OSPF. For inter-operability with other GNSS systems, and also to facilitate timing applications, GST is closely steered to UTC modulo 1 second, i.e., GST is a continuous timescale not subjected to the introduction of leap seconds. The requirement is that GST does not deviate by more than 50 ns (modulo 1 second) from UTC.

GST is physically generated at the Galileo PTFs, also part of the GMS. Two Galileo PTFs exist: one PTF is located in Fucino, Italy (FUC or PTF1), and the second PTF is located in Oberpfaffenhofen, Germany (OBE, or PTF2). In essence, each PTF comprises a clock ensemble consisting of two Active Hydrogen Masers (AHMs) and four Caesium clocks. In nominal operations, the output of the prime AHM is tuned by means of a phase-stepping instrument according to steering corrections received from the TVF. When the PTF runs in autonomy mode (i.e., without TVF corrections), the primary AHM is steered to the ensemble of the Cs clocks. TVF computes GST-UTC offset and steering for both PTFs. The master PTF is steered to UTC by applying the correction provided by TVF, the backup PTF computes and applies the steering to the master one. This allows keeping both of them close to UTC. The physical GST realisation at the PTF output is designated GST(MC) (MC = Master Clock).

In order to evaluate the GST-UTC offset, time-transfer equipment is operated at the PTFs, allowing to compare GST with UTC(k) as realized by five European timing laboratories: INRIM in Italy, OP in France, PTB in Germany, ROA in Spain, and SP in Sweden. Common-View time transfer using signals from GPS satellites (GPSCV) and Two-Way Satellite Time and Frequency Transfer (TWSTFT) are the implemented techniques. For TWSTFT, the well-known TimeTech SATRE modem and ancillary equipment is used. Regarding GPSCV, the two PTFs have been recently upgraded with Septentrio PolaRx4 receivers. All TWSTFT and GPS equipment have been calibrated through dedicated campaigns in 2014, [2] and [3], respectively. Time-transfer measurements from the PTFs and from the timing laboratories are sent to the TVF and processed daily. Internal clock-comparison daily data files from the PTFs are also processed by the TVF.

Each laboratory maintains its own UTC(k) time scale, which is an independent physical realization of UTC. The TVF computes a weighted average of the GST-UTC(k) differences, and uses it as an approximation of GST-UTC offset, called GST-UTCapprox. Once the GST-UTCapprox has been obtained, the PTF steering correction is computed as two different terms. The first term is the frequency correction needed to compensate the frequency offset and drift of the PTF primary AHM. It is computed from historical data of the clock (typically 10 days before the date of computation). By applying this first term the frequency offset of the GST scale is removed, but the GST phase offset could still be far from UTCapprox. The second term is then used to slowly reduce the GST-UTCapprox offset. As the frequency correction should vary only slowly from day to day and no phase corrections are allowed, for continuity, the second term tries to bring the offset to zero only after a long time period (typically 30 days). Fig. 2 shows an example of the steering outcome, with GST-UTCapprox reducing gradually from -6 ns to around zero after one month, and staying around zero afterwards. Only GPSCV was enabled in the TVF during the period reported.

![Fig. 2. GST-UTCapprox evolution over one and a half months.](image)

The goal of the TVF steering is to keep GST close to UTC, but at the same time to ensure that GST is stable enough for navigation purposes (satellite clock predictions). To achieve this goal, the frequency corrections computed by the TVF are limited to a maximum of 1E-14 units per day, and the Allan deviation of these corrections is daily monitored. Fig. 3 shows that a steering correction stability better than 3E-15 (ADEV at a 1-day averaging period) is possible, depending on the stability of the steered clock (an AHM in this case).

![Fig. 3. ADEV of the steering corrections provided by the TVF (at a 1-day averaging period over a 30-day moving window).](image)
III. ALTERNATIVE GST STEERING BY SP

ATSP, the Alternative Time Service Provider, is an independent entity in support of the TVF. The design of the ATSP has been chosen to be different from the main TVF, while maintaining functional and performance compatibility and providing redundancy in the estimation of the GST steering and monitoring signals.

The ATSP is based on a clock ensemble filter that is deployed at SP for the estimation of UTC(SP) [4]. The clock combiner is a classical Kalman filter that uses clock difference measurements as input. A clock is modeled using three states: phase, frequency and possibly frequency drift representing a clock’s absolute estimate relative to UTC. The states are connected in a set by a sequence of integrated noise processes, creating a simplified clock model. The clock states are parameterized by a standard set of parameters that were derived for typical clock behavior, distinguishing H-masers and Cs clocks. The parameters are time invariant representing an equal weighting scheme, which can be considered a robust filter approach. The filter is only run in the forward direction, but can dynamically reprocess data several times in order to include delayed information. This is especially true for reference data, which is a clock’s true phase to UTC and/or UTCr. Initial state information is usually derived from reference data.

Clocks within a site, UTC(k)’s and PTFs, are related to its local time scale using time interval measurements. Time scales are in turn compared using a combined TWSTFT and GPSCV time link. At any given epoch with available clock data the ATSP creates a minimal matrix of clock differences that covers all clocks of the involved sites. The filter then consumes the measurements and in turn estimates updated clock states and uncertainties. The time scale at each site, such as GST, is estimated using its master clock estimate and the local Time Interval Counter (TIC) measurements. ATSP applies the same rules as the main TSP in order to calculate the steering suggestion that minimizes phase and frequency offset of GST. The ATSP usually combines more than 50 clocks at any epoch providing daily GST-UTC phase and frequency estimates to the TVF.

IV. DISSEMINATION OF UTC-GST AND GGTO

The UTC-GST offset and GGTO are disseminated in the Galileo navigation message through the Galileo Signal-In-Space (GALSIS). UTC-GST is typically needed by Galileo users in order to time-stamp their position obtained from the Galileo PVT (Position/Velocity/Time) solution in UTC or local time. Timing applications are using this offset to evaluate the offset of the local timescale versus UTC. GGTO is needed to align GPS and Galileo satellite clock offsets in a combined GPS+Galileo PVT solution. The GST-UTC offset is provided to the GMS by the TVF; GGTO is calculated by the GMS itself. The disseminated offset values are validated independently by the TVF. The two broadcast offsets are obtained by the TVF through navigation messages collected by the Galileo Experimental Sensor Station (GESS) network. A GTR51 calibrated GPS+Galileo receiver at PTB, named PT10 station, is also used. PTB daily sends GPS and Galileo CGGTTS files from PT10 to the TVF.

For the UTC-GST offset validation, Galileo CGGTTS files from PT10 give the UTC(PTB) offset with respect to the GST estimate derived for each of the observed Galileo satellites. The satellite average at each epoch provides the UTC(PTB) offset versus estimated GST. Since the station has been calibrated [5], we obtain UTC(PTB)-GST(GALSIS). We then apply the UTC-GST offset from the GESS navigation messages in order to obtain UTC(GALSIS)-UTC(PTB). An example is shown in Fig. 4.

![Fig. 4. Validation of the operational UTC-GST offset from the Galileo navigation message using a calibrated receiver at PTB.](image)

GGTO is defined as GST-GPSt, where GPSt is GPS Time. For validation purposes, the TVF calculates the reference GGTO using up to five different methods, described in the next paragraphs.

GGTOPTB is GGTO as seen from the calibrated GPS+Galileo receiver located at PTB (PT10). CGGTTS files from PTB give the UTC(PTB) clock offset versus GPSt for each GPS satellite, and versus GST for each Galileo satellite. Subtracting the satellite averages for each epoch, the UTC(PTB) clock offset cancels out and we obtain GST-GPSt (GGTO). GGTOPTB is the main reference for the validation of the broadcast GGTO since it corresponds closest to the GGTO use in the future.

GGTOPTFCV is the daily average of all GPS satellite records contained in the daily GPS CGGTTS file received from a PTF. The PTF CGGTTS file gives the receiver clock offset (GST(MC) by definition) versus GPSt for each GPS satellite, and versus GST for each Galileo satellite. Subtracting the satellite averages for each epoch, the UTC(PTB) clock offset cancels out and we obtain GST-GPSt (GGTO). GGTOPTB is the main reference for the validation of the broadcast GGTO since it corresponds closest to the GGTO use in the future.

GGTOUSNO(AV) is the difference GST(MC)-UTC(USNO) as seen from GPS All-in-View time transfer between the PTF and USNO. GGTOUSNO(TW) is the difference GST(MC)-UTC(USNO) as seen from indirect TWSTFT between the PTF and USNO, using PTB as pivot. UTC(USNO) is used as an approximation of GPSt in these two methods.
GGTO_{ODTS} is GGTO as seen from an Orbit Determination and Time Synchronization (ODTS) process running at GMV, using the PTF GSS station as reference so that estimated GPS satellite clocks are referred to GST. By then comparing with the GPS satellite clocks contained in the navigation message (referred to GPSt) and averaging for all satellites, we can have an estimation of GST-GPSt. Fig. 5 shows a comparison of the operational GGTO from the navigation message against the five GGTO methods provided by the TVF.

Fig. 5. Validation of the operational GGTO from the navigation message against five alternative GGTO calculations by the TVF.

V. SERVICE LEVEL AGREEMENT

An important feature of the new TVF for the FOC phase is the provision of operations under a Service Level Agreement (SLA) between the prime contractor (GMV) and ESA. The SLA is based on the fulfillment of a number of so-called Key Performance Indicators (KPIs) related to the quality and availability of the operational products delivered by the TVF to the PTFs. During the first four quarters of FOC operations with the TVF system running at GMV, the average level of fulfillment of the SLA with ESA was of 99.5%. The calculation of the TVF KPIs is done independently by PTB based on the raw product files from TVF.

In order to ensure the provision of a high-quality service by the TVF, a reliable contribution of the UTC(k) laboratories with their time-transfer data is essential. To this effect, a lower-tier SLA has also been put into place between GMV and the five UTC(k) acting as sub-contractors. The SLA guarantees the timeliness and integrity of the GPS/C and TWSTFT time-transfer files arriving daily at the TVF ftp server. For GPS/C, data from two receivers is required (prime and backup). All time-transfer daily files must arrive at the TVF before 01:00 hours UTC of the next day. All data files must comply with certain rules to ensure their completeness and quality. The SLA between GMV and the laboratories is evaluated daily and also reviewed every quarter, if needed. The daily evaluation is based on the TVF Daily Report which is automatically generated and which summarizes the time-transfer results between the PTFs and the UTC(k) realizations. An example is shown in Fig. 6.

Fig. 6. Daily time-transfer results between a PTF and the five UTC(k) providers (from the TVF Daily Report).

In general, the performance of the UTC(k) service provision by the laboratories is outstanding, and with a similar level of reliability from all of them. In a typical quarter, the average fulfillment of the UTC(k) SLA terms is around 99% for the ensemble of institutes. The UTC-UTC(k) deviation for each laboratory, although not formally subject to SLA, is also monitored using the results published in BIPM’s Circular T. An example is shown in Fig. 7. As can be seen UTC-UTCapprox is very close to zero, therefore UTCapprox as realized from averaging the five UTC(k)’s can be used as a very good real-time approximation of UTC.

Fig. 7. Performance of UTC-UTC(k) from Circular T for the five laboratories contributing to the TVF. The evolution of UTC-UTCapprox and GST-UTC is also shown.

VI. CHARACTERIZATION OF SATELLITE CLOCKS

A key task of the TVF is the performance evaluation of the Galileo satellite on-board clocks. Each satellite flies two Passive Hydrogen Masers (PHM), named PHM-A and PHM-B, and also two Rubidium Atomic Frequency Standard (RAFS), RAFS-A and RAFS-B. The OVF provides GPS and Galileo orbit and clock estimations based on the combination of global network solutions from three Processing Facilities.
The TVF analyses the satellite clocks in continuous weekly intervals based on “rapid” clock products, and in continuous quarterly intervals based on “final” products. The OVF processes a number of pre-selected stations that can be used as continuous ground clock reference in the TVF analyses. In particular, PTBB located at PTB and connected to UTC(PTB) is preferred by the TVF and normally used as reference.

The TVF produces satellite-reference clock phase, frequency, and ADEV plots. The clock frequency is calculated in an approximate way as the time-derivative of the raw phase in two consecutive values. The satellite clock characterization is limited by the “system noise” of the clock estimation and combination process, and by the stability of the ground reference clock, normally UTC(PTB) when the receiver PTBB is used. For the evaluation of the “system noise” the ADEV stability of a stable reference clock is plotted (e.g., USN4). The “system noise” does not include the apparent satellite clock instabilities originating from the orbit/clock correlation in the clock estimation.

A novel activity in the FOC phase is the evaluation of PHM clocks over long periods of time (one quarter). Fig. 8 shows the evolution of the E19 PHM-B clock offset (versus PTBB) over one quarter. The stability of the reference clock (PTBB) and the “system noise” are evaluated by comparison with USN4 clock. Fig. 8 demonstrates that PTBB is stable enough for satellite clock characterization. Fig. 9 shows the corresponding ADEV for the same period, for all active PHM clocks at the time. A typical feature seen in general in GNSS space clock analysis in these days is an increased instability (‘bump’) at tau equal to half the orbital period of the satellite and at sub-harmonics. This is related to residual uncertainty in modelling of the satellite orbit, e.g., due to solar pressure effects [7].

![Fig. 8. E19 PHM-B satellite clock offset evolution over one quarter (91 days), versus PTBB. The USN4 clock evolution is also shown.](image)

A critical activity phase for the TVF is the so-called satellite In-Orbit-Tests (IOT), a period of several weeks after the satellite launch where the satellite basic navigation functionality is evaluated. In this phase it is important to evaluate in a rapid way the performance of the on-board clocks.

The first IOT supported by the TVF in the FOC phase is the one of the fifth Galileo satellite (E18), launched on August 22, 2014. During the IOT two “hot” clock swaps were performed, from PHM-B to RAfS-B, and back to PHM-B. It is the first time that such procedure is executed by a GNSS satellite without signal interruption. As analyzed by the TVF, the clock transitions went smoothly with phase steps of the order of few cm only. An example is shown in Fig. 10.

![Fig. 10. “Hot” clock swap from RAfS-B to PHM-B during the E18 satellite In-Orbit-Tests (IOT).](image)
worldwide distributed. During the FOC phase all the GESS stations have been upgraded with new Septentrio PolaRx4 receivers for improved measurement quality. Some of the GESS are located at UTC(k) laboratories. Such “timing” stations are GIEN, GNOR, GPTB, and GUSN, located at INRiM (Italy), ESTEC (The Netherlands), PTB (Germany), and USNO (Washington DC, USA), respectively. GIEN is connected to a free-running Active Hydrogen Maser (AHM), and GNOR, GPTB, and GUSN are connected to UTC(ESTC), UTC(PTB), and UTC(USNO), respectively. These receivers are PolaRx4 timing versions accepting as input 10 MHz frequency and 1PPS timing signals. GESS timing receivers are essential for the TVF because they provide a reliable and traceable clock reference for satellite clock characterization, with direct support from the hosting laboratories if necessary. The GPTB station at PTB is new in the FOC phase. Its installation was completed at the end of 2014. Fig. 11 shows the GPTB antenna and the rack hosting the receiver and ancillary equipment.

GPTB is the only station in the GESS network that has been fully calibrated for GPS and Galileo signals. The GPS calibration has been done relative to the co-located PTBB reference station. This is possible since both GPTB and PTBB are connected to UTC(PTB). The calibration is done by processing and comparing GPS P1 and P2 pseudoranges from both stations in RINEX format. Once the GPS receiver delays have been calibrated, it is possible to calibrate the Galileo delays in E5a and E5b relative to GPS P1/P2 using the method developed by ORB [8]. The key feature of the ORB method is the cancellation of the GPS and Galileo ionospheric delays when combining pseudoranges from two satellites with a close position in the sky. The Galileo calibration is done by processing GPS and Galileo pseudoranges from GPTB in RINEX format, using a dedicated TVF software tool. Satellite positions are read from a SP3 orbit file. Details on the GPTB calibration process can be found in [5].

Fig. 11. The new GPTB station at PTB.

VIII. CONCLUSIONS

The new TVF for the Galileo FOC phase started operations in March 2014 and has been successfully running for one year now demonstrating excellent performance of GST-UTC steering down to the nanosecond level and availability higher than 99%. Its main role is the daily steering of GST to UTC, which is achieved thanks to the participation of a group of European timing laboratories providing time-transfer measurements between the Galileo PTFs and their UTC(k) realizations. The GST-UTC offset is normally maintained at a few-ns level. The TVF represents a valid precursor of the future Galileo Time Service Provider (TSP) and Galileo Reference Center (GRC).

ACKNOWLEDGMENT

The TGVF contract is being carried out under a programme of and funded by the European Union, and managed by ESA. The views expressed herein can in no way be taken to reflect the official opinion of the European Union and/or ESA. We are grateful to P. Defraigne from ORB for having provided freely information and measurement data related to the Galileo receiver calibration method relative to GPS. A. Kuna from UFE kindly supports the operation of the GTR51 receiver at PTB. W. Schaefer from TimeTech helped to formalize the support needed for the TWSTFT calibration campaign. The TVF uses International GNSS Service (IGS) products for Precise Point Positioning (PPP) calculations. National Resources Canada (NRCan) PPP online service is used often for the validation of TVF results.

REFERENCES

Front-End Receiver: Recent and Emerging Trend

Ulrich L. Rohde  
BTU Cottbus 03046 Germany

Ajay K. Poddar  
Synergy Microwave NJ USA

Enrico Rubiola  
FEMTO-ST Inst. Besancon France

Marius A. Silaghi  
University of Oradea, Romania

Abstract—This paper describes the recent trend of front-end receiver systems for the application in radio monitoring. The receiver implementation is optimized for applications such as hunting and detecting unknown signals, identifying interference, spectrum monitoring and clearance, and signal search over wide frequency ranges, producing signal content and direction finding of identified signals.

Keywords—Receivers, Real time FFT Processing

I. INTRODUCTION

The emerging security threats demand an intensive data gathering for fiber or radio communication [1]. Besides the fiber technology, there are varieties of wireless activities, which are typically analyzed by off the air monitoring [2]-[9]. The spectral density of signals these days is very high and therefore such monitoring receivers require high performance. The technique in designing such receivers is a composition of microwave engineering of the building blocks preamplifier, mixer, synthesizer and necessary filters, this paper describes critical aspects of important building blocks of radio monitoring receivers [10]-[20].

What is needed: The following is a proposed data sheet for a surveillance receiver. The dynamic range requirements are typically the highest in the frequency range between 80 and 160 MHz, as the broadcast band is full of strong signals (80 MHz to 109 MHz) and the frequency range above covers the aircraft radio band, the amateur radio band: 144-148 MHz (80 MHz to 109 MHz) and the frequency range above covers and 160 MHz, as the broadcast band is full of strong signals which are typically the highest in the frequency range between 80 and 117 MHz, as the broadcast band is full of strong signals. Monitoring receivers require high performance. The technique in designing such receivers is a composition of microwave engineering of the building blocks preamplifier, mixer, synthesizer and necessary filters, this paper describes critical aspects of important building blocks of radio monitoring receivers [10]-[20].

II. DATA FOR RADIO MONITORING RECEIVERS [21]

<table>
<thead>
<tr>
<th>Frequency range</th>
<th>650 MHz to 1300 MHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 MHz to 650 MHz</td>
<td>V&lt;sub&gt;in&lt;/sub&gt; = 117 dBm (0.3 μV) ≥10 dB</td>
</tr>
<tr>
<td></td>
<td>V&lt;sub&gt;in&lt;/sub&gt; = 47 dBm (1 mV) ≥50 dB</td>
</tr>
<tr>
<td></td>
<td>LSB/USB, IF bandwidth 500 Hz</td>
</tr>
<tr>
<td></td>
<td>AF=500 Hz</td>
</tr>
<tr>
<td></td>
<td>0.5 MHz to 20 MHz, V&lt;sub&gt;in&lt;/sub&gt; = 0.4 μV ≥10 dB</td>
</tr>
<tr>
<td></td>
<td>20 to 30 MHz, V&lt;sub&gt;in&lt;/sub&gt; = 0.5 μV ≥10 dB</td>
</tr>
<tr>
<td></td>
<td>LSB/USB, IF bandwidth 2.5 kHz</td>
</tr>
<tr>
<td></td>
<td>AF=1 kHz</td>
</tr>
<tr>
<td></td>
<td>0.5 MHz to 20 MHz, V&lt;sub&gt;in&lt;/sub&gt; = 0.6 μV ≥10 dB</td>
</tr>
<tr>
<td></td>
<td>20 to 30 MHz, V&lt;sub&gt;in&lt;/sub&gt; = 0.7 μV ≥10 dB</td>
</tr>
<tr>
<td></td>
<td>V&lt;sub&gt;in&lt;/sub&gt; = 100 μV ≥46 dB</td>
</tr>
<tr>
<td></td>
<td>AM, IF Bandwidth 2.5 kHz, f&lt;sub&gt;mod&lt;/sub&gt;=1 kHz, m=0.5</td>
</tr>
<tr>
<td></td>
<td>0.5 MHz to 20 MHz, V&lt;sub&gt;in&lt;/sub&gt; = 1 μV ≥10 dB</td>
</tr>
<tr>
<td></td>
<td>20 to 30 MHz, V&lt;sub&gt;in&lt;/sub&gt; = 1.2 μV ≥10 dB</td>
</tr>
<tr>
<td></td>
<td>Crossmodulation interfering signal 2.5 V (+21 dBm), SINAD ≥20 dB</td>
</tr>
<tr>
<td>Demodulation</td>
<td>AM, FM, LOG, PULSE; SSB and CW optional</td>
</tr>
<tr>
<td>Squelch</td>
<td>signal controlled, adjustable −10 dB V or to 80 dB V (max. 110 dB V, 120 dB V)</td>
</tr>
<tr>
<td>AGC range</td>
<td>90 dB; 1 μV to 10 mV makes ≤4 dB difference in AF level</td>
</tr>
<tr>
<td>RF attenuator 30 dB selectable or AGC speed for 90 dB range</td>
<td>Attack Decay</td>
</tr>
<tr>
<td>AFC speed for 90 dB range</td>
<td>Attack Decay</td>
</tr>
<tr>
<td></td>
<td>AM/B=15 kHz &lt;15 ms 15 ms Pulse/B=100 kHz &lt;0.1 ms 3 s, corr. to SSB/B=2.5</td>
</tr>
<tr>
<td></td>
<td>kHz &lt;1 ms, 3dB/100ms</td>
</tr>
<tr>
<td>Range of MGC (manual gain control), EGC (external gain control) by analog voltage</td>
<td>90 dB 90 dB</td>
</tr>
<tr>
<td>COR Decay Attack</td>
<td>adjustable 1 s to 10 s ≤25 ms</td>
</tr>
<tr>
<td></td>
<td>digital tuning for signals of unstable frequency graphic using tuning markers, numeric in 50 Hz steps (B ≤100 kHz)</td>
</tr>
<tr>
<td>AFC Offset indication Signal-level indication</td>
<td>graphically as level line or numeric from −10 dB V to 80 dB V (110 dB V), with tuner 0 120 dB V graphic 1 dB, numeric 0.1 dB ≤2 dB for level ≥ 0 dB V 1000 definable memory locations, each location may be allocated a complete set of receive data, up to 250 ch/s five definable start/stop frequency spans with separate receive data sets (5 jobs), up to 250 ch/s full receive range (max. 650 MHz) or any expanded</td>
</tr>
</tbody>
</table>

III. TECHNICAL SPECIFICATIONS

<table>
<thead>
<tr>
<th>Frequency range</th>
<th>9 kHz to 26.5 GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 kHz, 100 Hz, 10 Hz, 1Hz</td>
<td>≤1.5 x 10⁻⁵ (6°C to +55°C) ≤1 x 10⁻⁵</td>
</tr>
<tr>
<td>5 x 10⁻⁷ per year</td>
<td>≤138 dB (10 kHz)</td>
</tr>
<tr>
<td>≤1 ms</td>
<td>≤≤ 113 dBm</td>
</tr>
<tr>
<td>Immunity to interference, nonlinearities</td>
<td>Image frequency rejection typ. 110 dB, ≥90 dB</td>
</tr>
<tr>
<td>IP2</td>
<td>typ. 50 dBm, ≥40 dBm</td>
</tr>
<tr>
<td>IP3</td>
<td>typ. 35 dBm, ≥28 dBm</td>
</tr>
<tr>
<td>Spurious</td>
<td>≤113 dBm</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>measurement using telephone filter to CCITT</td>
</tr>
<tr>
<td>Total Noise Figure (incl. AF Section)</td>
<td>U&lt;sub&gt;c&lt;/sub&gt; = −103.5 dBm (1.5 μV) ≥10 dB</td>
</tr>
<tr>
<td>(S+N)/N ratio</td>
<td>U&lt;sub&gt;c&lt;/sub&gt; = −47 dBm (1 mV) ≥47 dB</td>
</tr>
<tr>
<td>FM, B=15 kHz, f&lt;sub&gt;mod&lt;/sub&gt;=1 kHz, deviation 5 kHz</td>
<td>9 kHz to 26.5 GHz</td>
</tr>
<tr>
<td>Vin=−107 dBm (1 μV) ≥25 dB</td>
<td></td>
</tr>
</tbody>
</table>
Radio monitoring receivers must be able to process antenna signals with high cumulative loads and wide dynamic range [1]-[2]. Unknown signals are normally detected using high-speed scans over wide frequency ranges and then analyzed in detail in fixed frequency mode. A radio monitoring receiver’s scan speed and probability of intercept (POI) are determined by its real-time bandwidth, sensitivity and the type and speed of signal processing employed. In particular, seamless real time processing is a requirement that other receivers do not meet. To support real-time processing, large-bandwidth and dynamic range is needed without compromising sensitivity. Some radio monitoring receivers feature multiple, switchable broadband receive paths [3].

The digital signal processing (DSP) module is the key to monitoring the required task and performance. High performance radio monitoring receivers feature DSP computing power so high that up to four times the number of FFT points actually needed is available, depending on the selected real time bandwidth [4]. By selecting an appropriate FFT length, even closely spaced channels can be reliably detected as discrete channels [5].

By utilizing the higher number of FFT points available, the FFT can be expanded by up to four times. The high computing power can also be used to perform FFT calculation using overlapping windows. This makes even short pulses clearly discernible in the spectrum’s waterfall display [6].

As shown in Figures (1)-(4), signal processing is generally referred to as seamless (gapless), although pulses may go undetected if they are very short and located at an unfavorable position with respect to the FFT frame (see upper processing step in the figure "Overlapping FFT").
Therefore, some receivers offer overlapping FFT. Two FFTs whose frames are shifted with respect to one another are calculated in parallel from the data stream [7].

A sample located in the minimum of the Blackman filter curve of one FFT will then be found in the maximum of the other. For a real time bandwidth of 10 MHz as used in this example, minimum signal duration of 240 µs is required to ensure 100% reliable signal acquisition and correct level measurement [8].

For shorter pulses, the level may not be displayed correctly, but only very weak signals may go undetected. It is evident that the use of digital signal processing in a radio monitoring receiver offers great advantages [9]. Extremely high sensitivity (due to very fine resolution) combines with a broad spectral overview and high scan speed to significantly increase the probability of intercept over analog receivers or spectrum analyzers [18].

In the panorama scan mode, the spectrum is displayed across a frequency range far wider than the radio monitoring receiver’s real time bandwidth. This mode provides users with a quick overview of the spectrum occupancy. The principle of the fast spectral scan (panorama scan) is described in the following using a receiver with up to 20 MHz real time bandwidth (such as the R&S®EB500) [21]. During the scan, frequency windows of a maximum of 20 MHz width is linked in succession, so that the complete, predefined scan range is traversed shown in Figure (5), illustrates the “signal processing in panorama scan mode”.

As is done for the IF spectrum, an FFT is used to process the broad window with a finer resolution. The width of the frequency window and the FFT length (number of FFT points) are variable and are selected by the receiver automatically. The user can select among 24 resolution bandwidths from 100 Hz to 2 MHz. The resolution bandwidth corresponds to the width of the frequency slices (bin width) mentioned under “IF spectrum”.

Based on the selected bin width and start and stop frequency, the monitoring receiver automatically determines the required FFT length and the width of the frequency window for each scan step. The receiver selects these internal parameters so that the optimum scan-speed is achieved for each resolution bandwidth shown in Figure (6), illustrates “Resolution in panorama scan mode”.

The highest resolution bandwidth of 2 MHz yields the maximum scan speed, while the smallest resolution bandwidth of 100 Hz yields maximum sensitivity. The resolution bandwidth (bin width) for the panorama scan (selectable between 100 Hz and 2 MHz) therefore corresponds to the resolution bandwidth (BW_{bin}) used in the displayed noise level (DNL) calculation for the IF, and can be used for calculating the DNL for the panorama scan as:

\[
DNL = \text{–}174 \text{ dBm} + NF + 10 \times \log(BW_{bin}/\text{Hz}) \tag{1}
\]

The quantity NF represents the overall noise figure of the receiver. The user selects the resolution bandwidth to obtain the desired frequency resolution shown in Figure (7), illustrates “Bin width and channel spacing”.

A receiver’s available IF bandwidth has a direct influence on the achievable panorama scan speed. Doubling the IF bandwidth (i.e. using 20 MHz instead of 10 MHz in this example) will also double the achievable scan speed. If the IF bandwidth is increased from 20 MHz to 80 MHz, the scan speed can be boosted by a factor of four.

The above explanations show that the use of digital signal processing in a radio-monitoring receiver offers decisive advantages. Extremely high sensitivity (due to very fine resolution) combined with a broad spectral overview and maximum scan speed significantly increases the probability of intercept as compared with an analog receiver.

As the number crunching of analog to digital converters improves, 1 GHz clock frequency has already been achieved, better performance will be possible. The systems performance is determined by hardware, firmware and software. Even if everyone has access to the same chipset the implementation may differ depending upon the market price of the equipment. Not everyone needs the high-end systems.
III. RECEIVER CONCEPT

The rapidly advancing in development of components for digital signal processing and the rapid rise of processing power enables more and more new concepts in the implementation of so called Software Defined Radios (SDR). The SDR approach includes transmitter and receiver concepts, whereas the signal processing is largely done in programmable devices, such as field programmable logic arrays (FPGA) and digital signal processors (DSP). Thereby, components that have been typically implemented in hardware (e.g. amplifiers, mixers, filters, modulators, etc.) are instead implemented by means of software on a digital signal processing platform. Thus, hardware is replaced by software. A software defined radio platform has the advantage of being able to make functional changes on a receiver or transmitter to meet new requirements quickly. Thus it is possible to keep these products up to date for a long period. A typical example is the mobile telephone according the various GSM, UMTS and LTE standards. Even inexpensive amateur radio use more and more of this modern technology, because of the increasing integration of available DSP and FPGA modules and the performance enhancement of microprocessors, accelerating the application in this price segment.

The SDR approach is making use of time discrete signal processing by sampled signals. Such sampled systems have been published first in 1985 [9]-[10]. The typical receiving frequency range is 9 KHz to 55 MHz given a typical sampling frequency of 125MHz, with a resolution equal to 16 bit. The latest 16 bit Analog-to-Digital Converters (ADC) offer a range up to 110MHz, at a sampling frequency of 250MHz. The key element in a SDR receiver is the analog-to-digital converter. This device is placed ideally as close as possible to the antenna. For VHF (>50MHz) to SHF frequencies however, there is a need to place an analog down converter in front of the AD converter. All signal processing components following the AD converter are widely free from tolerances, noise, unwanted couplings and profit from a high reproducibility and a zero drift. The software configurable hardware components and DSP algorithms allow a maximum product flexibility.

The analog front-end in Figure (8) may consist of pre-selector filters, a pre-amplifier and, if needed, a frequency converter. The preconditioned signal is then passing the AD-Converter, which is producing a sampled discrete time representation of the analog, time invariant input signal. The output data rate of the AD Converter (i.e. the product of sample rate times resolution) is in the range of 1…4Gbit/s. This data stream may consist of the whole spectrum from DC up to the half sample rate. For most communications purposes, only a small fraction of this bandwidth is of interest, unless the implementation of several receivers in the system is wanted in special cases. Therefore it needs a single, or a number of frequency conversions with a subsequent reduction of the sample rate. This function block is designated to as the Digital Down Converter (DDC). The preconditioned digital signals are then processed at a lower data rate in the Digital Signal Processing unit. The task of the DDC is to perform the digital down conversion, decimation of the channel rate, baseband IQ generation, channel filtering, and offset cancellation, using commercially available ASICs (application specific integrated circuits), or a programmable hardware in the form of FPGAs (field programmable gate arrays). Subsequently, the further processing as demodulation, clock and carrier synchronization, decryption, audio processing, spectrum analysis, etc. can be performed by the DSP.

A typical front-end consists of an input stage, first mixer stage including the necessary synthesizer, a possible second mixer stage, and then the output is fed to signal processing down at the IF level of choice. As seen in section II, there are varieties of important receiver parameters. The noise figure determines the minimum discernible signal sometimes also called minimum detectable signal, typically expressed in dBm and overall dynamic range, the key intermodulation distortion products [1]-[2]. System noise and noise floor defines the spurious free dynamic range. The 2\textsuperscript{nd} order intermodulation distortion product can degrade the system performance; the appropriate input filters reduce this as depicted in Figure (9).

Figures (10) and (11) show the typical scheme of up/down converted for high performance receiver applications. This signal is then up-converted to an IF of about 20 GHz using a highly cleaned-up LO chain. The up-conversion of signal is achieved with the help of frequency doubler and large bank of filters, as shown in Figure (10). Spectral purity of the oscillator chain is of the essence. Finally, a second converter is used to down convert the signal to the IF level. This arrangement allows a very high performance signal analysis, as illustrated in Figure (11).
Figure (12) shows the typical block diagram of wideband radio monitoring receiver. Figure (13) shows a frequency panorama display and a waterfall time event. On the left hand of the frequency display, there are a large number of FM broadcast stations and on the right side are an air-traffic control frequency range and police and other mobile radioactivity.

As depicted in Figure (13), the waterfall display shows the various transmissions that are useful for frequency occupancy analysis over time, and the spectral display shows a range where demodulation of transmission is possible while observing all the activity.

It can be seen that Figure (13) demonstrates the capability of high dynamic ranges; Aircraft Radio Communication Receiver can be monitored and demodulated in the presence of strong FM Radio signal. Figure (14) shows the multi-channel operation capability, 5-channel arrangement for signal analysis. The wideband monitoring receiver can handle five individual channels simultaneously, including transmission monitoring. In this case, two ATC frequencies and two FM broadcast stations.
It can be seen that the spectral density of signals these days is very high and therefore such monitoring receivers require high performance. Figures (12)-(14) demonstrate the capability of very high performance radio monitoring receiver systems.

IV. CONCLUSION

The systems performance is determined by hardware, firmware and software. Even if everyone has access to the same chipset the implementation may differ depending upon the market price of the equipment. Not everyone needs the high-end systems.

ACKNOWLEDGEMENT

We would like to thank Rohde & Schwarz for making the latest hardware information and specifications available. We actually have one of those surveillance systems, which allow us to monitor the radio amateur frequencies and other bands of interest. The surveillance system is connected to a variety of active antennas as passive antennas are too frequency selective. The dynamic range of active antennas is another topic of interest for further investigation.

REFERENCES

[21] ESMD Wideband Monitoring System, Rohde & Schwarz

Fig. 13 Wide bandwidth waterfall (courtesy: R & S)

Fig. 14: Five channel arrangement for signal analysis

Fig. 15: Five channel arrangement for signal analysis
Oscillator Phase Noise: A 50-year Retrospective

D. B. Leeson
Department of Electrical Engineering
Stanford University
Stanford, CA 94305-9515, USA
leeson@stanford.edu

Abstract—Fifty years ago emerging developments in oscillator applications led to the formation of an IEEE committee to unify time- and frequency-domain definitions of frequency stability. As a member, the author had the good fortune to participate and contribute. This paper is a personal recollection of events and impressions of the committee's 1964 IEEE-NASA Symposium, our 1966 Proc IEEE special issue on frequency stability (with comments on this author's oscillator-model paper), and our 1971 "Characterization of Frequency Stability" paper that was written to provide a basis for IEEE Std 1139.

Keywords—phase noise, short-term stability, frequency stability, 1964 IEEE-NASA Symposium, history

I. INTRODUCTION

Oscillators, the sources of signals in electronic systems for time keeping, radio communications and radar, are characterized by frequency stability. Today the understanding of both noise-like and environmentally induced frequency instabilities in oscillators is both rigorous and readily applied. Information on the subject is widely accessible. Citation records show some 17,000 publications found in searches for "phase noise" (11,000) and "frequency stability" (6,000).

Today, standards such as IEEE 1139 offer concise definitions of concepts for the optimization of performance and interpretation of measurement [1]. Fifty years ago, differing applications had arisen in relative isolation. There was insufficient interaction to consolidate basic concepts and terminology. The current beneficial outcome was formally initiated then by an IEEE Standards subcommittee formed to unify time- and frequency-domain definitions of frequency stability. I was fortunate to participate and contribute, and this paper offers personal recollections and comments on a process that led to today's standards.

II. FREQUENCY STABILITY STANDARDS TODAY

A. Frequency Domain

Today almost the entire picture can be reduced to a few simple expressions and graphics [2]. Two sets of parameters are well-known components of standards. For an oscillator whose output is described as \( V(t) = A \cos(\omega_0 t + \phi(t)) \), the power spectral density (PSD) \( S_\phi(f) \) of the phase \( \phi(t) \) due to random noise is modeled in frequency domain as a power-law sum \( S_\phi(f) = \Sigma_b \nu_f^{-n} \), where \( f \) is the Fourier frequency and \( n \) ranges from 4 > \( n > 0 \). The exponents correspond to white phase (\( n = 0 \)), flicker of phase (\( n = 1 \)), white frequency (\( n = 2 \)), flicker frequency (\( n = 3 \)) and random walk frequency (\( n = 4 \)). Equivalent forms are PSD of frequency \( S_f(\nu) \) or normalized frequency \( S_f(\nu) = (1/\nu_0^2) S_\phi(f) = (f^2/\nu_0^2) S_\phi(f) \).

The frequency-domain standard defines the measure of phase noise as \( \mathcal{A}(f) = (1/2) S_\phi(f) \). \( \mathcal{A}(f) \) is seen, for the common case \( 1 \ll |a| \) and AM \ll FM, to be a useful approximation to the RF spectrum. Dynamic range is typically limited in cell phones (the "near-far" issue) and Doppler radar (subclutter visibility) in this portion of the RF spectrum conforming to the small-angle and minimal-AM assumptions. \( \mathcal{A}(f) \) is commonly presented as in Fig. 1, the log-log plot of \( S_\phi(f) \) vs. \( f \), which concisely reveals the power-law terms.

Phase noise from power supply, frequency modulation inputs and vibration or acoustic exposure has long been recognized as a mechanism that can potentially dominate quiescent noise performance. While this area represents a significant part of the author's own experience, and the literature is extensive, it is outside the scope of this paper.

B. Time Domain

The time-domain definition of stability over a time interval \( \tau \) is a specific two-sample variance \( \sigma_\tau^2(\tau) \) of normalized random frequency \( y = (1/\nu_0)df/dt \), known as the Allan Variance. The Allan Variance, also a power-law sum, can be derived from, and exists for, all spectral density power laws encountered in physical oscillators. This permits conversion from frequency domain measurements to time domain. A modified form is better suited for showing this relationship, also shown in Fig. 1, a log-log plot of \( \sigma_\tau(\tau) \) vs. \( \tau \), whose segments correspond to the exponents of the PSD \( S_\phi(f) \). It can be seen that over longer times, the low-frequency terms of phase PSD dominate. [after NIST Pub 1065]

Fig. 1. Log-log plots of \( S_\phi(f) \) vs. \( f \) and \( \sigma_\tau(\tau) \) vs. \( \tau \), showing segments
III. HISTORY

A. Frequency Stability before 1960

This enlightened state of affairs did not exist fifty years ago. Before 1940, physics and radio made use of frequency standards such as the WWV stations that employed the best quartz crystal techniques of the day. Instrumentation followed to provide portable metrology.

The need for large quantities of quartz crystals arose at the time of WWII with the introduction of channelized radio communications for mobile and airborne warfare. Problems of volume manufacturing and aging were identified and resolved to the necessary degree [3].

B. Frequency Stability at the Beginning of the 1960's

After the war's end, promising new developments such as semiconductor devices, quantum-physics frequency and timekeeping devices, television, mobile radio communication, microwave Doppler radar and even space rocketry were ripe for development. By the 1960's, applications of stable oscillators fell into two broad classes.

One, precision time and frequency standards and metrology, found the expression and measurement of frequency instabilities most natural in time-domain terms. The other, multi-signal systems such as Doppler radar and radio communications, with their dynamic range limitations due to spectrum, turned to frequency-domain definition and measurement as more applicable. Many systems were newly exposed to more stringent environments, as well.

The annual Symposium on Frequency Control, sponsored for many years by the U.S. Army, and then by the IEEE, served well as a common forum. In electronic circuits and quantum devices, the frequency-domain papers of this period focused on the RF spectrum of the oscillator itself, or on linewidth, rather than the power spectrum of phase [4]-[7]. The result was that the predicted theoretical spectra were not in complete agreement with complex observed spectra [8].

The tools appropriate for each application (long-term and time-domain in time-keeping and frequency standards, short-term and frequency-domain in radar, and a combination of both in communications) evolved along divergent paths as if in something of a guild system. Then the revolutionary developments of the transistor, the integrated circuit, digital computing and communications techniques and even the large space rocket gave rise to new requirements.

IV. IEEE SUBCOMMITTEE 14.7

A. The Formation of the Subcommittee

The emergence of communications and ranging techniques in the space program created a need for understanding and advances in both time and frequency domain concepts, and a means to convert from one to the other. These requirements stemmed from the rise of digital modulation techniques, as well as ranging. This created a new constituency whose unique issues that were not addressed by the existing communities.

At this point, early in the Apollo, satellite and planetary exploration programs, it became apparent to the several communities that they were experiencing the parable of the blind men and the elephant, and that some effort was required to pool the independent reservoirs of knowledge. The urgent NASA interest in finding common terminology for oscillator and system specifications found fertile ground in the IEEE, and a subcommittee of Technical Committee Standards 14 — Piezoelectric and Ferroelectric Crystals was proposed to explore a cross-discipline standard.

In response to this impetus, the Technical Subcommittee, Standards 14.7 — Frequency Stability was established to serve as a focal point for information in the field. The ultimate aim of the Subcommittee was an IEEE standard on the definition and measurement of both short-term and long-term frequency stability. It was at this time that I was fortunate to receive an invitation to join the Subcommittee through a mentor and sponsor, W. K. Saunders, who was familiar with my prior publications and my work on Doppler radar at Hughes Aircraft Co. My early contact with oscillator noise came as solid-state signal sources began to be applied to the radars that had been under development since the days of the MIT Radiation Laboratory. I was initiated into the phase-noise requirements of airborne Doppler radar as I applied the nonlinear frequency multipliers of my graduate theses.

Subcommittee 14.7 attracted members from the full range of user and instrument communities, and in discussion it was realized that a cross-specialty symposium would be very useful for exchange of viewpoints and techniques that would promote convergence. The committee sensed that the separate use of frequency-domain and time-domain definitions stood in the way of development of a common standard. We hoped to find a common language to discuss frequency stability, one that could be understood by everyone in the discipline. The Subcommittee focused first on the short-term frequency stability regime, in which there were greatest differences of viewpoint among the multiple user communities.

B. IEEE-NASA Symposium on Short-term Frequency Stability

The first step of the program to craft a standard that would define frequency stability was to understand and meld the frequency- and time-domain descriptions of phase instability to a degree that was mutually accepted, and that permitted analysis and optimization. To promote focused interchange as an extension of its own discussions, the Subcommittee acting as a program committee sponsored the November 1964 IEEE/NASA Conference on Short Term Frequency Stability. That conference, with some 350 attendees, was an opportunity for the cross-fertilization of ideas, and featured papers on all aspects of generation, application and measurement of short-term frequency stability. The Symposium proceedings give an insight into ripening questions and authoritative answers [9].

Of particular interest to me is the record of four panel discussions, led by prominent scientists and engineers of the time. The tension between rigorous theory and practical experiment came out often, as did the concern that adoption of a single standard would leave a portion of the community without the tools for its specific applications. Also evident was
the full range of individual specialization and experience, and even of personality types.

Specific questions were raised about the uncertainties of the correspondence between near-carrier linewidth and RF spectrum and an underlying spectral density of phase or frequency. The conundrum of the origin of flicker noise with its lack of convergence of integrals at zero frequency received substantial questioning and discussion. Additionally, the subject arose whether higher order effects or amplitude noise were adequately recognized, and it was concluded that experimental evidence supported the idea that these were not significant in then-current applications.

One of my own curiosities on reviewing the Symposium proceedings from this remove was to identify how and where certain key concepts were conclusively identified in the papers. For example, there was substantial discussion of power-law descriptions of PSD of phase and its relationship to RF spectrum, but my review found no graphic that explicitly showed the segments. The issue of converting from frequency to time domain was explored, but not resolved at that early time. A number of authors noted flicker noise in amplifiers and other physical devices. There were several efforts to relate the output spectrum of an oscillator to the characteristics of the resonator feedback network and the active device, but the full connection between the amplifier PSD, resonator and amplifier parameters and the output PSD remained to be clarified.

### C. 1966 Proceedings Special Issue on Frequency Stability

With the success of the 1964 Symposium, in order to consolidate the gains and promote further exchange of information, the Subcommittee was invited to serve as editorial committee of the February 1966 Special Issue on Frequency Stability of the Proceedings of the IEEE [10]. This issue attracted many submittals, including updated papers from the 1964 Symposium, including several by committee members who were also among the most active in the field.

We were most pleased to receive a paper by D. W. Allan that settled issues of time domain definitions and techniques, as well as showing how to convert from frequency domain definitions [11]. By this time it had become accepted that spectral density of phase or frequency, rather than RF spectrum or linewidth, was the more fundamental frequency-domain measure of short-term stability. PSD could be directly related both to the time domain definition that became identified as the Allan Variance, and within limits it predicted the RF spectrum. The IEEE 1139 standard now applies the small-angle limitation in reverse, such that the RF spectrum is *defined* as half the PSD of phase except where the small-angle condition is not met.

In preparation for this paper, I revisited the special issue, again curious to find when and where key points became clear. Although there were many instances of power-law spectra, I still found no example of the now-accepted multi-segment PSD of phase. Papers that dealt with determining oscillator output PSD from input PSD were restricted to a subset of the overall question. Questions remained regarding flicker noise, nonlinearities, the interrelation of PSD of phase and RF spectrum, and AM noise. A number of papers dealt with flicker noise, including one that specifically mentioned flicker noise in resonators. Papers on oscillator-multipliers suggested a choice of higher oscillator frequency because of the multiplication of modulation index. The radar community, responding to vibration problems, was adopting the ribbon-mounted quartz crystal developed at the Bell Telephone Laboratories.

### D. Model of Feedback Oscillator Phase Noise Spectrum

In our final deliberations to settle the contents of the special issue, it seemed to me that we had not received a paper on frequency-domain techniques that was as clear as the Allan paper was on time-domain issues. I thought I could see a way to create such a paper.

A paper selected for our special issue showed that, subject only to conditions that were typically met in oscillators, for a nonlinear circuit driven by a periodic input, the AM and PM noise could be treated as strictly linear and stochastic, and thus could be described in terms of spectral densities [12]. Its author and I had been graduate students together at MIT, sharing a thesis advisor. Adding to my own background in nonlinear circuits, this encouraged my interest in synthesizing a simple quasi-linear model of oscillator noise behavior.

At this point it seemed enough was known to assemble a model that used the power-law forms of PSD, with graphics to provide additional clarity. The input and output PSD could be related by a transfer function reflecting the key parameters of the active element and resonator. This transfer function would be an extension of those of Edson and Baghdady to include the model for white noise outside the resonator bandwidth from my own Doppler-radar papers. I felt that a quasi-linear model of phase noise as a small perturbation of the oscillator steady-state signal, even in a nonlinear oscillator, would have broad applicability. Since this was very late in the editing process, I was encouraged by the committee to submit a concise paper that would be published as correspondence, since that section was held open for late submittals.

That was the origin of my 1966 paper on the oscillator noise model [13]. Looking back, I am satisfied with what I was able to shoehorn into two pages, submitted after discussion with colleagues by Dec. 29, barely a month before our publication date. Proceedings correspondence was not archived, so for a number of years the paper remained obscure except to insiders. During that same time I was fully occupied with founding and managing a new company, so the paper led something of a life of its own. I am pleased to find its continuing utility has raised it to the most cited paper in the "phase noise" category.

In the intervening fifty years there have been advances in the clarity with which the concepts could be expressed, and the model has been extended to new frequency-determining and active elements. Much later, novel requirements and solutions would arise from the emergence of the integrated circuits that would completely reshape what was possible in electronics, and would require new approaches. Many of the questions about nonlinearity and RF 1/f spectrum have been resolved, through physical argument or mathematical rigor.
V. COMMENTS ON THE PHASE NOISE MODEL PAPER

A. Spectral Models of Phase Variations

As was already the practice, the oscillator output was taken as \( v(t) = A \cos \{ \omega_0 t + \phi(t) \} \), where \( \phi(t) \) is treated as a zero-mean stationary random process. In Symposium papers and discussions among attendees and committee members, it had been concluded that power spectral density of phase \( S_\phi(\omega_0) \), or its equivalent, the PSD of frequency \( S_f(\omega_0) = \omega_0^2 S_\phi(\omega_0) \), represented the most suitable definition of phase noise instabilities (as opposed to RF spectrum or linewidth, from which one could not necessarily determine a unique spectrum of phase).\(^1\) Here \( \omega_0 \) was taken as the Fourier frequency associated with the noise-like variations in \( \phi(t) \). Subject to the limitation AM \( \ll \) FM and to the small angle approximation \( \phi \approx \sin \phi \ll 1 \), the normalized RF power spectrum was related by a constant to \( S_\phi(\omega_0) \).

B. Power Spectral Density of Oscillator Internal Phase Noise

For the VHF overtone crystal feedback oscillator used as the basis for the model, the spectrum \( S_\phi(\omega_0) \) of the input phase uncertainty \( \Delta \phi(t) \) was taken to have two components, flicker 1/f модуляtion and additive white noise around the oscillator frequency, including "noise at other frequencies mixed into the pass band of interest by nonlinearities."

The spectral density of input phase due to additive white noise was known from modulation theory to be the ratio of noise power to signal power. For a feedback oscillator with an effective noise figure \( F \) (giving effect to nonlinear mixing), the two-sided input spectrum \( S_\Delta \phi(\omega_0) = 2FKT/P_s \), where \( P_s \) was taken to be the signal level at the input of the oscillator active element. The factor of 2 was deleted in subsequent papers in a change to one-sided spectra.

The second component of input phase spectrum was seen to be parameter variations that modulate the internal phase at video or baseband rates. This modulation, whose PSD typically varies as 1/f, impresses its effect on the oscillator signal without any appeal to nonlinearity. This modulation component is independent of signal amplitude. The flicker variation of the resonator itself, which was beginning to be appreciated by those working with lower frequency oscillators, was not observed then in VHF crystals and was disregarded by me at that time.

A suitable expression for total spectral density of input phase errors was (and is) of the form \( S_\Delta \phi(\omega_0) = \alpha/\omega_0 + \beta \), where \( \alpha \) is a constant determined by the magnitude of 1/f flicker variations and \( \beta = FKT/P_s \) for one-sided spectra. The already modest noise figure \( F \) of those days was raised by "corrections necessary to account for nonlinear effects, which must be present in a physical oscillator." This line of thought was expanded then in a following section on nonlinearity.

C. Relation to Oscillator Internal Noise

A key intent of the model was "to show clearly the relationship of the output spectral density of phase \( S_\phi(\omega_0) \) to the known or expected noise and signal levels and resonator characteristics of the oscillator," \( F, P_s, Q \) and \( \omega_0 \). The simplest model was that of a linear feedback oscillator.

To deduce from physical reasoning the transfer function from input phase spectrum to output phase spectrum, the paper considered a single-resonator feedback network of fractional bandwidth \( 2B/\omega_0 = 1/Q \), with \( Q \) the loaded quality factor. For small phase variations that fall within the feedback half-bandwidth \( \omega_0/2Q \), a phase error at the oscillator input due to noise or parameter variations would result in a frequency error determined by the phase-frequency slope of the feedback network, \( \Delta \omega = 2Q\Delta \phi/\omega_0 \).

For modulation rates large compared to the feedback bandwidth, the feedback loop has no effect, so for this regime the output power spectral density \( S_\phi(\omega_0) \) was seen to equal the input spectral density \( S_\Delta \phi(\omega_0) \). Thus a suitable expression for the transfer function was given as \( |H(\omega)|^2 = [1 + (\alpha/\omega_0 + FKT/P_s)]^2 \).

D. Power Spectral Density of Output Phase

The PSD of output phase was then shown to be just the product of the input spectrum and the transfer function, so \( S_\phi(\omega_0) = S_\Delta \phi(\omega_0) |H(\omega)|^2 \). This could be shown simply in the graphic construction of Fig. 2 that identified the feedback bandwidth and the breakpoint of the flicker segment. Note that flicker noise is typically modulative, and hence does not vary with \( F \) or \( P_s \).

The example given was the case in which 1/f effects predominate only for frequencies small compared with the feedback loop bandwidth. It was noted that if flicker predominate, the breakpoints would be interchanged. The caption noted that the RF spectrum could be derived subject to the limitations (small-angle, AM \( \ll \) FM) in the text. One-sided rather than two-sided spectra became the norm at a later time, thus resolving a factor of two in the white noise level. A significant point is that this figure and the measurement made graphically explicit the power-law segments of oscillator PSD.

---

1 As a note, the original paper used a dot over the symbol \( \phi \) to denote the time derivative ("Newton's notation"). Recalling the frustration of chasing dots pasted to the galley proof, the author heartily approves of the modern use of \( v \) or \( \Omega \) for frequency. This problem also plagued other papers in the issue.
E. Output PSD Experimental Verification

A measurement was presented to validate the theoretical model. The model and data are compared in Fig. 3, reproduced from the original paper. The agreement was reassuring.

F. Video Frequency Range of Interest

Space systems and Doppler radar were of particular interest to the author. Space data links used narrow bandwidth, and so the low Fourier frequencies of the flicker segment were seen as critical. Radar requirements ranged up to 100 kHz. The name "Hertz" as the unit of frequency had just been adopted in 1965, and was not yet in common use.

G. Choice of Oscillator Frequency

Frequency multipliers were known to increase modulation index by a factor equal to the multiplication ratio , so PSD is increased by . Our work was at 10 GHz. For a given output frequency, the choice of oscillator frequency is significant.

The graphical construction in Fig. 4, alluded to but not included in the paper, shows that a higher oscillator frequency yields lower noise for Fourier frequencies above the resonator bandwidth. From comparisons such as this, it was also seen that the most favorable PSD segments of oscillators could be combined by use of phase lock loops in synthesizers.

H. 1/f noise in the active element

Since 1/f variations and nonlinearity compromised the achievable PSD, it was suggested that AGC oscillators with large-area high-power transistors could provide simultaneous improvements in flicker and nonlinear effects. Later it was found that bipolar devices were better in this aspect.

It was pointed out that output PSD could be modified by subsequent bandlimiting filtering and by the noise of following amplifiers. Last, the potential was noted for coupling the signal directly from the resonator to filter the white noise component.

I. Nonlinear Effects

To raise the estimate of noise figure above the status of a fitting factor, I had estimated an added 4 dB above the published small-signal value "to account for nonlinear mixing of noise at third harmonic and higher frequencies." This is shown schematically in Fig. 6. The paper was directed at the VHF oscillator type then typical of Doppler radar applications I worked with, with the expectation that the result would have more general applicability.

VI. "Characterization of Frequency Stability"

By the end of the 1960's, it was felt that sufficient progress had been made that we could prepare a paper to summarize understanding that would underlie a future standard. Produced in the days before the Internet and email, this involved numerous discussions and correspondence among the ten authors. Despite the complexity of responding to all viewpoints, we deemed the result to be a useful step forward, and it was published in 1971 [14].

VII. Limits and Extensions of the Simple Model

Over time, the following questions have properly been raised about the limits of applicability of the simple model:

- Does reflect nonlinearity and circuit impedances?
- Extension to other frequency-determining elements?
- Is flicker of resonator recognized?
- Near-carrier limits of conversion from PSD to RF?
- Is AM to PM conversion recognized?
A. Nonlinearity

In time, strongly nonlinear oscillators arose from the proliferation of semiconductor integration. These were treated by quite different fundamentally nonlinear analyses in more recent papers, in one case by close colleagues at Stanford [15]-[16]. However, measurements of high-Q oscillators continue to confirm the persistent utility of the simple model [17].

B. Extended frequency determining elements

A range of new frequency determining elements has arisen over time, including delay lines, bandpass filters and multiple resonators. In many cases, the simple model is extended by determining the phase slope from group delay \( \tau \) rather than \( Q \).

C. Flicker in Active Elements and Resonators

As proposed in the phase noise model paper, the reduction of flicker noise by feedback and choice of active element resulted in substantial improvement in oscillator stability in a relatively short time [18]. As can be seen from a graphical construction, this effect would be much greater in high-Q HF oscillators, as opposed to lower-Q VHF overtone oscillators [19]. Flicker noise in the resonator itself had been suggested in a 1966 paper [20]. Investigations confirmed the significance of flicker in resonators [21]. This noise source was less observed in VHF crystals and was disregarded by me in the simple model, which has been extended to recognize this [22].

The nature and effect of flicker noise has been the subject of substantial attention in subsequent years. A confounding problem was the infinity at zero frequency for PSD rising as \( 1/f^6 \). It has been suggested that finite bandwidth and measurement time create the equivalent of a bandpass filter that acts to truncate the PSD [23].

D. Near-carrier Large Modulation Index

A related issue is the regime of large modulation index where the small-angle assumption is not valid, in which the spectral density of phase grows without limit, typically very near the carrier. This issue has been termed the "infrared catastrophe" by allusion to the ultraviolet catastrophe of pre-quantum radiation physics. Papers before the 1960's modeled the RF spectrum of an oscillator frequency modulated by noise with components down to zero frequency. More recent papers confirm mathematically that the output power of the modulated oscillator remained constant as expected, and that the close-in RF spectrum shape is Lorentzian or Gaussian [24].

E. AM-PM Conversion

The effect of AM-PM conversion remains a concern that must be considered. Oscillators generally meet the criterion AM « PM, and experiment has shown it not to be a primary issue in many systems of interest. By modulation theory, equal RF sidebands confirm that one or the other form dominates. From experience, this is typically phase noise in an oscillator with limiting or frequency multiplication.

ACKNOWLEDGMENT

The author gratefully acknowledges respectful interactions with the members of the IEEE Subcommittee, especially the late J. A. Barnes, L. S. Cutler, D. J. Healey and W. K. Saunders. The author also thanks E. Rubiola, T. H. Lee, J. Everard and M. M. Driscoll for very helpful discussions.

REFERENCES

Least-Square Fit, $\Omega$ Counters, and Quadratic Variance

F. Vernotte$^\triangledown$, M. Lenczner$^\triangledown$, P.-Y. Bourgeois$^\triangledown$, E. Rubiola$^\triangledown$

$^\triangledown$ UTINAM/Observatory THETA, University of Franche-Comté and CNRS, 41 bis, avenue de l’observatoire, BP 1615, 25010 Besançon Cedex, France
$^\triangledown$ FEMTO-ST Institute, UMR 6174 : CNRS/ENSMM/UFC/UTBM, Time and Frequency Dpt., 26 ch. de l’Epitaphe, 25030 Besançon Cedex, France

Abstract—This work is motivated by the wish to have the most precise measurement of a frequency $\nu$ and of the variance $\sigma_x^2$ of its fractional fluctuations in a given time $\tau$, out of high-end general-purpose instruments.

Thanks to the progress of digital electronics, new time-interval analyzers have been made available in the last few years. Such instruments measure the time stamp of the input events at high sampling speed (MS/s), and with high resolution (10–100 ps).

We propose the linear regression as a means to estimate the frequency from time stamps of the input signal. The frequency counter based on the linear regression is called $\Omega$ counter. The linear regression is interpreted as a finite impulse response filter which takes the frequency samples as the input, and delivers the estimated frequency at the output. We derive the transfer function of such filter, which turns out to be parabolic shaped.

As compared to the $\Pi$ and $A$ counters, the $\Omega$ counter features better rejection of the background noise.

We define the quadratic variance (QVAR), a wavelet variance similar to the Allan variance, and we derive its statistical properties. The QVAR is superior to the AVAR and MVAR in the rejection of the background noise.

I. THE $\Omega$ COUNTER

For our purposes, the time analyzer takes the input signal as the ‘start,’ and the reference as the ‘stop,’ and also as the time stamp. Let us assume that reference signal (stop) has frequency $1/\tau_0$, that input signal (start) has the nominal frequency $\nu_0$, and that $\nu_0$ is greater or equal than the sampling frequency $1/\tau_0$. In this conditions, at each ‘stop’ event the instrument measures the time interval $t_{stop} - t_{start}$. The time tag associated to the measure is $t_{stop}$. Broadly speaking, this is equivalent to the ‘Picket Fence’ scheme introduced by Greenhall [1], [2].

We define the phase time, denoted with $x$, as either $t_{stop} - t_{start}$ or its fluctuation. Similarly, we define the fractional frequency $y$ either as $\nu/\nu_0$ or its fluctuation $(\nu - \nu_0)/\nu_0$. Of course, the use of ‘value’ and ‘fluctuation’ must be consistent.

We introduce the $\Omega$ counter as a new type of frequency counter that uses the linear regression on the time series \{$x_k$\} as the means to estimate the input frequency. Presently, the $\Omega$ counter is a proposed implementation using time-tag time interval counters. The game of the name $\Omega$ will be explained later.

General purpose time-tag instruments have wide input bandwidth (usually 0.2...2 GHz, or more), thus the instrument background is chiefly white noise, with negligible contribution of flicker and other coloured noise types. In this conditions, the linear regression is an obvious choice as it gives the best rejection of the background.

Let us denote with $\hat{A} = \nu/\nu_0$ either the fractional frequency or its fluctuation, and with $\hat{A}$ its estimate. Using the linear regression in the time interval $\tau = N\tau_0$ centered at $t = 0$, $\hat{A}$ reads

$$\hat{A} = \frac{12}{N^3\tau_0^3} \sum_{-(N-1)/2}^{(N-1)/2} k x_k. \quad (1)$$

Equation (1) gives the fractional frequency when we take $x = t_{stop} - t_{start}$, and its fluctuation when we take $x$ as the fluctuation or the background noise.

A. Phase-time filter interpretation

We rewrite Equation (1) as

$$\hat{A} = \sum_k h_{\Omega x}(t_k)x_k, \quad (2)$$

where

$$h_{\Omega x}(t) = \frac{12t}{\tau^3} = \frac{12t}{N^3\tau_0^3} \quad \text{(Fig. 1)} \quad (3)$$
Equation (2) states that the linear regression can be interpreted as a filter applied to the phase-time data. Alternatively, the linear regression can be seen as a weighted measure. This alternate interpretation is made easy by the symmetry of \( h_{\Omega y} \).

B. Frequency filter interpretation

We rewrite Equation (1) as

\[
\hat{A} = \sum_k h_{\Omega y}(t_k) y_k
\]

where \( h_{\Omega y}(t) \) is given by

\[
h_{\Omega y}(t) = \frac{d}{dt} h_{\Omega y}(t),
\]

thus

\[
h_{\Omega y}(t) = \frac{6}{\tau^3} \left( \frac{\tau^2}{4} - t^2 \right).
\]

Equations (4)–(6) rely on the property of the convolution integral, that \( f' * g = f * g' \).

As before, the linear regression can be interpreted as a filter. However, the filter now processes the fractional frequency \( \{y_k\} \) associated to \( \{x_k\} \), and delivers the estimated frequency. The impulse response of such filter is \( h_{\Omega y} \). Alternatively, the linear regression is seen as a weighted measure applied to \( \{y_k\} \). The relation between these two interpretations is made easy by the symmetry of \( h_{\Omega y} \).

C. The game of the name

In a previous paper [3] we introduced terms ‘\( \Pi \) counter’ and ‘\( \Lambda \) counter’ for the instruments with rectangular and triangular weight function, respectively. In the same way we use the term ‘\( \Omega \) counter’ for the linear-regression counter, because the impulse response has parabolic shape, which is broadly similar to the Greek letter \( \Omega \). (Fig. 2 C).

II. COMPARISON OF THE \( \Pi \), \( \Lambda \) AND \( \Omega \) COUNTERS

We introduce two classical operators, the mathematical expectation \( \mathbb{E}\{x\} \), and the mathematical expectation of the variance \( \mathbb{V}\{x\} = \mathbb{E}\{(x - \mathbb{E}\{x\})^2\} \) of a random variable \( x \).

Figure 2 shows the weight functions of the \( \Pi \) counter, the \( \Lambda \) counter, and the new \( \Omega \) counter. We compare the performance of these counters in the presence of white PM noise with zero mean and variance \( \sigma_y^2 \). Basic statistics (details are omitted) gives the results summarized in Table I.

The \( \Pi \) counter is clearly inferior to the other two because the white PM noise is rejected as \( 1/\tau^2 \) instead of \( 1/\tau^3 \). The \( \Lambda \) looks superior to the \( \Omega \) counter only because the measurement is allowed to take twice the time. If we constrain the measurement time to \( \tau \) for both, we get

\[
\mathbb{V}\{\hat{A}_\Lambda\} = \frac{2\tau_0^2\sigma_y^2}{\tau^3} = \frac{4}{3} \mathbb{V}\{\hat{A}_\Omega\}.
\]

The supports are chosen for the two-sample variance (defined later) to be the same for the three counters.

III. THE TWO SAMPLE VARIANCE

The two-sample variance is defined as

\[
\sigma_y^2 = \frac{1}{2} \mathbb{E}\{(y_2 - y_1)^2\},
\]

where \( y \) results from the appropriate estimator. We get the Allan variance AVAR with \( y = A_{\Pi} \) (\( \Pi \) estimator), the modified Allan variance MVAR with \( y = A_{\Lambda} \) (\( \Lambda \) estimator), and the new quadratic variance QVAR with \( y = A_{\Omega} \) (\( \Omega \) estimator).

IV. THE QUADRATIC VARIANCE QVAR

A. Time domain

Using phase measurements as the input data, the QVAR is given by

\[
\sigma_y^2(\tau) = \left\langle |x(t) * h_{\Omega x}(t)|^2 \right\rangle
\]
A: Phase-time data

\[ h_Q(t) = \frac{6\sqrt{2}}{\tau^3} \left(t + \frac{\tau}{2}\right) \quad t \in [-\tau, 0] \]

\[ h_Q(t) = \frac{6\sqrt{2}}{\tau^3} \left(-t + \frac{\tau}{2}\right) \quad t \in [0, \tau]. \]

B: Frequency data

\[ h_Q(y(t)) = 3\sqrt{2} \tau^3 \left(t - \frac{\tau}{2}\right) \quad t \in [-\tau, 0] \]

\[ h_Q(y(t)) = 3\sqrt{2} \tau^3 \left(t - \frac{\tau}{2}\right) \quad t \in [0, \tau]. \]

B. Frequency domain

Using frequency measurements as the input data, the QVAR is given by

\[ \sigma_Q^2(\tau) = \left\langle \left| y(t) \ast h_Q(y(t)) \right|^2 \right\rangle \] (10)

with (Fig. 3-B)

\[ h_Q(y(t)) = \frac{3\sqrt{2}t}{\tau^3} \left(-t - \tau\right) \quad t \in [-\tau, 0] \]

\[ h_Q(y(t)) = \frac{3\sqrt{2}t}{\tau^3} \left(t - \tau\right) \quad t \in [0, \tau]. \]

In the frequency domain, the QVAR is given by

\[ \sigma_Q^2(\tau) = \int_{0}^{\infty} S_y(f) |H_Q(y(f))|^2 \, df \] (11)

Figure 3. Time domain computation weight of QVAR for phase data (above) or frequency deviations (below).

Table II

<table>
<thead>
<tr>
<th>( S_y(f) )</th>
<th>( \text{MVAR}(\tau) )</th>
<th>( \text{QVAR}(\tau) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_{-2}f^{-2} )</td>
<td>( \frac{11\pi^2 h_{-2}^2}{20} )</td>
<td>( \frac{26\pi^2 h_{-2}^2}{35} )</td>
</tr>
<tr>
<td>( h_{-1}f^{-1} )</td>
<td>( \frac{[27\ln(3) - 32\ln(2)] h_{-1}}{8} )</td>
<td>( \frac{[2\ln(16) + 1] h_{-1}}{5} )</td>
</tr>
<tr>
<td>( h_0f^0 )</td>
<td>( \frac{h_0}{4\pi} )</td>
<td>( \frac{3h_0}{5\tau} )</td>
</tr>
<tr>
<td>( h_{+1}f^{+1} )</td>
<td>( \frac{[24\ln(2) - 9\ln(3)] h_{+1}}{8\pi^2\tau} )</td>
<td>( \frac{3[\ln(16) - 1] h_{+1}}{2\pi^2\tau} )</td>
</tr>
<tr>
<td>( h_{+2}f^{+2} )</td>
<td>( \frac{3h_{+2}}{8\pi^2\tau} )</td>
<td>( \frac{3h_{+2}}{2\pi^2\tau} )</td>
</tr>
</tbody>
</table>

For a linear frequency drift:

\[ y(t) = D_1 \cdot t \]

\[ \frac{1}{2} D_1^2 \tau^2 \]

\[ \frac{1}{2} D_1^2 \tau^2 \]

\[ |H_Q(f)|^2 = \frac{9}{2} \frac{\sin^2(2\pi f\tau) - \pi f \sin(2\pi f\tau)}{(\pi f)^3}. \]

The transfer function \(|H_Q(f)|^2\), shown in Figure 4, is an octave band pass filter similar to that of the AVAR, but with significantly smaller side lobes. Table II shows the responses to the different types of noise. The same information is shown in Fig. 5.

C. Convergence

For small \( f \), it holds that

\[ |H_Q(f)|^2 \approx \frac{3}{(\pi f)^2}. \]

Thus, \( \sigma_Q^2(\tau) \) defined in Eq. (11) converges for \( f^{-2} \) FM.

At large \( f \), \( |H_Q(f)|^2 \) decreases as \( f^{-4} \), and \( \sigma_Q^2(\tau) \) converges for \( f^{-4} \) FM.
V. EQUIVALENT DEGREES OF FREEDOM

For a general introduction to the problem of the degrees of freedom in the two-sample variance, the reader can refer to [4].

As done with the other variances, we assume that the QVAR estimates are approximately $\chi^2$ distributed

$$\tilde{\sigma}_Q^2(\tau) = \alpha \chi_n^2$$

(14)

where $\alpha$ is a scale coefficient, and $n$ is the Equivalent Degrees of Freedom (EDF). The mathematical expectation and the variance of a $\chi^2$ distribution are proportional,

$$\mathbb{E}\{\chi_n^2\} = n$$

$$\mathbb{V}\{\chi_n^2\} = 2n.$$  

(15)  

(16)

The fractional dispersion of the QVAR estimates is given by

$$\Delta \tilde{\sigma}_Q^2(\tau) = \frac{\sqrt{\mathbb{V}\{\tilde{\sigma}_Q^2(\tau)\}}}{\mathbb{E}\{\tilde{\sigma}_Q^2(\tau)\}} = \frac{\sqrt{2\alpha^2\nu}}{\alpha\nu} = \sqrt{\frac{2}{n}}.$$  

(17)

Figure 6 shows the histogram distribution in the case of $10^4$ realizations.
A: White FM noise

B: Frequency drift

Figure 8. Detection of white FM noise and drift in presence of white PM noise

Figure 7 shows the EDF in a typical representative case, where \( N = 2048 \). For \( f^2 \) and \( f \) FM noise (white and flicker PM noise), the AVAR is of little interest because it rejects the background proportionally to \( 1/\tau^2 \) and \( 1/\tau^2 \ln \tau \), instead of \( 1/\tau^3 \). For all the noise types shown, the QVAR has higher EDF than the MVAR, which is desirable. For the \( f \) and \( f^2 \) noise types, the QVAR is substantially equivalent to the AVAR.

VI. DETECTION OF WHITE FM NOISE AND DRIFT

Comparing the variances, the capability to detect white FM noise and drift in the shortest time \( \tau \), for a given white PM noise (instrument background) is a relevant criterion. The latter can also be expressed as the lowest white FM noise and drift that can be detected in a given \( \tau \).

The white FM noise is present in all atomic standards, and the drift is ubiquitous.

Figure 8 A shows a variance plot with PM noise (instrument background), the error bars at 95% confidence level, and a white FM noise (black line). Figure 8 B shows the same variance plot, with a drift (black line) The FM noise (or the drift) is detected with a probability of 95% when the upper point of the error bar hits the FM noise (drift) line. In both cases the MVAR wins.

VII. EXAMPLE OF APPLICATION

Let us consider a time interval analyzer that receives a perfect reference at 1 MHz at the ‘stop’ (time stamp) input, and the 10 MHz output from a H maser at the ‘start’ input. In this condition, the maser phase time is sampled at a rate equal to \( 1/\tau_0 = 1 \) MHz. We assume that each time-interval measurement is affected by white noise (instrument background) with \( \sigma_\epsilon = 10 \) ps. From Table I, the background translates into \( \sigma_y = 3.5 \times 10^{-14} \) at \( \tau = 1s \), decreasing as \( \tau \sqrt{\tau} \). For comparison, the typical noise \( f \) a H maser is \( \sigma_y = 10^{-13} \) at \( \tau = 1s \), decreasing as \( 1/\sqrt{\tau} \). Thus, the \( \Omega \) counter implemented with such time interval analyzer is in principle sufficient to monitor the maser, out of the box, with no dedicated down-conversion hardware.

ACKNOWLEDGEMENTS

This work is supported by the ANR Programme d’Investissement d’Avenir in progress at the TF Departments of Femto-ST Institute and UTINAM (Oscillator IMP, First-TF and Refimeve+), and by the Région de Franche Comté.

REFERENCES

SDR and Self-Focusing Radar Techniques for milliHerz Measurement of Multi-Component Phase Noise Spectra?

Michael J Underhill
Underhill Research Ltd
Lingfield, UK
e-mail: mike@underhill.co.uk

Abstract — Synthetic Aperture Radar (SAR) Self-Focusing post-processing techniques can be used for faster higher precision Phase Noise and Allan Variance measurement on more than one signal component at a time. Software Radio (SDR) techniques [1, 2] are used to sample, acquire decimate and process the signals, on a continuous basis, for record lengths of up to a few days and with resolution bandwidths down to a few milliHerz.

Keywords — software radio measurements; phase noise; Allan variance; Synthetic Aperture Radar self-focusing; milli-Herz bandwidths; multiple signals

I. INTRODUCTION

Software Defined Radio (SDR) techniques provide spectrum analyzers with state of the art performance at low cost. Being software based, such spectrum analyzers are very flexible and can incorporate the latest signal processing techniques available from new research, and from other fields of signal processing such as Radar.

One Radar technique is ‘Super-Resolution’. Here this is used to propose a faster measurement of the precise frequency of an oscillator. The technique is shown to give up to two orders of magnitude reduction in the measurement time for a given frequency measurement precision.

A second Radar technique is Synthetic Aperture Radar (SAR) Self-Focusing. This post-processing technique is shown to give faster higher precision Phase Noise and Allan Variance measurement on one or more signal component at the same time.

II. SOFTWARE DEFINED RADIO (SDR) RECEIVERS AS DIGITAL SPECTRUM ANALYSERS

The use of Software Radio (SDR) techniques for low cost spectrum analysis and phase noise measurement has been previously examined [1, 2]. Here SDR techniques are further refined for measurements down to a few mHz resolution bandwidth. The data collected in such measurements can easily be processed into any type of Allan Variance with improved accuracy and/or reduced measurement time. Here we use off-the-shelf SDR receivers to sample, acquire and process the signals, on a continuous basis, for record lengths of up to a few days.

The basis of an SDR is a multi-bit ADC (Analog-to-Digital Converter) sampling at a high frequency. Long record lengths are required for milliHerz (mHz) resolution measurements. As a rule of thumb a 1000 second record length is required to measure with a precision of 1mHz. This implies that very large files of data have to be stored. To avoid this in practice, data processing is operated in real time to reduce the data storage requirements to a manageable level and designed to achieve the desired measurement outcome. The RFSpace SDR-IP has 16 bit ADC sampling at 80MHz and this creates data at a rate of 160 megabytes/second, or a gigabyte in 6.25 seconds, or 14 terabytes a day. In practice this data is processed (or decimated) in real time while it is being collected, so that only the required information is recorded and stored with the maximum accuracy, resolution and precision depending and the available equipment and measurement time. Easily available PCs of modest specifications can then host several simultaneous Allan variance or Phase noise measurement processes all in real time.

Fig. 1 shows the three SDR receivers used for the results in this paper. All receivers are computer (PC) driven and all signal processing is performed by software hosted in the PC. In this case the two PCs used are a Dell Studio and a Samsung NC10 notebook. The SDRs are always run as background tasks in general leaving more than enough capacity for usual PC tasks to be run simultaneously. SpectraVue free-ware has been used on both the RFSpace receivers to provide spectrum analysis with resolution bandwidths down to a few milliHerz, 3.9mHz for the SDR-IQ and 15mHz for the SDR-IP.

III. SUPPRESSION OF SPECTRUM ANALYSER MEASUREMENT SPURS BY USE OF SDR TECHNIQUES

A useful new discovery is that the spurious signals that typically appear around 100Hz to 1kHz when sweeping or
phase-locked spectrum analysers are used, do not appear if an SDR receiver is used as a spectrum analyser.

Fig. 2. Spectrum Analyser Measurement ‘Spurs’ around 100Hz. On left: 10MHz signal from HP 8642 fed via ‘AJC Concept Demonstrator’ to HP8560A Digitally transferred from HP8560A via HPB to PC with Toric Custom Software and to an Excel spread-sheet. On right: 9600 Symmetricron Ultra-miniature Space and Military OCXO Series OCXO Series at 10MHz

Fig. 2 shows unwanted spurious components or ‘spurs’ on the of two different spectra of 10 MHz sources measured by conventional sweeping type spectrum analysers or ones containing switching types mixers or phase detectors.

Fig. 3 shows a set of three spectrum plots simultaneously taken using an SDR receiver as a spectrum analyser with a resolution bandwidth of 3.9mHz. Note that there are no spurious measurement spurs or fluctuations in the spectra.

Fig. 3. Example of three simultaneous SDR phase noise spectra plots of 10MHz signal from an HP8457D source. The RFSpace SDR-IQ was used with SpectraVue software to display spans of 20Hz, 200Hz and 2Hz, with 3.9mHz resolution (RBW), and using Blackman-Harris windowing function. 1db vertical scale steps on bottom right plot and 10dB steps on top and bottom left plots. No spectrum or time averaging was used. The data record length was about 260 seconds. Decimated sample rate was 8137 Hz and a two million point FFT was used.

One possible source of fluctuations is the inevitable non-linearities of mixers, counter sampling and (one bit) quantisation, and (switching) phase detectors in conventional frequency sweeping types of spectrum analysers and Allan variance counter/timers. We find that SDR techniques overcome most if not all of such problems.

However we also find that SDR methods can overcome a further source of spectrum measurement fluctuations. SDR spectrum analysers use FFTs (Fast Fourier Transforms) with ‘windowing’ functions. ‘Windowing’ implements filters for the spectrum samples which can much improve the selectivity of each spectrum sample against unwanted contributions from adjacent spectrum samples but at the cost a broadened filter response and less precision of frequency measurement.

A useful discovery is that any windowing function with poor far out selectivity is likely to suffer severe fluctuations of the total spectrum, depending on the phase relationship between the signal and the sampling frequency comb. The average rate of the fluctuations is observed to be about 6 to 10 times the FFT sample rate which the inverse of the resolution bandwidth or FFT frequency sample separation.

Fig. 4 shows these ‘windowing’ effects. All spectrum plots are of a 20MHz crystal oscillator with a display bandwidth of 220Hz and a resolution bandwidth of 1Hz. SpectraVue software was used for all the plots. For the top two plots the simplest ‘rectangle’ window was selected in the SpectraVue software. The top left plot captures the maximum of the plot variation and the top right one captures the minimum. The fluctuations amount to about 30 to 40dB between these extremes. The fluctuations are biased towards the upper values and can be reduced to less than about 10dB by averaging four plots. But then the time required to make the measurements goes up by a factor four. The bottom left plot uses the ‘Hanning’ window. This retains a good accuracy of frequency measurement as shown by the reasonably sharp peak of the plot, but the medium far out selectivity is compromised and this part of the spectrum does suffer reasonably small fluctuations of less than 10dB. The bottom right plot in Fig. 4 shows the use of the ‘Blackman-Harris’ windowing function. This has very good selectivity with negligible spectrum fluctuations. It is highly recommended.

The use of an SDR FFT with the Blackman-Harris windowing function is therefore highly recommended for low cost high precision spectrum analysis with essentially no spurious fluctuations in the measured spectra.

Fig. 4. Unwanted spectrum fluctuations for three FFT windowing functions (WFs). Top left and top right show the maximum and minimum of fluctuations of the ‘rectangle’ WF. Bottom left is using the ‘Hanning’ WF. Bottom right is using the recommended Blackman-Harris WF.
IV. Fast Frequency Measurement by a SDR ‘Super-Resolution’ Technique

Most signal sources have a well defined ‘bell’ shape with a defined bandwidth for at least the central part of the spectrum. The Leeson model can be used to find the 3dB bandwidth of an oscillator spectrum from measurements of sideband phase noise. For example an oscillator with a phase noise of 120dBc/Hz at 1kHz has a 3dB bandwidth of about 2mHz. The existence of this well defined shape allows a phased array radar ‘super-resolution’ technique to be used to give one or two orders improvement in the speed of making a frequency measurement to a given precision.

Fig. 5 shows two examples of four markers on adjacent 0.5Hz spaced FFT spectrum samples of a crystal oscillator. The SpectraVue software allocates finds the exact amplitudes of each marker and displays them in the top left corner of the spectrum plot. With knowledge of the mathematical shape of the Blackman-Harris windowing function, used in this case, interpolation can be used to estimate the frequency of the oscillator from the marker information. The left hand plot shows a case where the oscillator frequency is almost exactly half way between the two highest markers. The right hand plot shows the case where the oscillator frequency is slightly to the right of the highest marker.

The expectation is that a single measurement using this super-resolution technique should give at least a ten times more precise estimate of the actual oscillator frequency. Two or three measurement sets taking two or three times longer should arguably give a further order of magnitude improvement in accuracy.

This interpolation method could also be used to get a direct estimate of the mHz or less bandwidth of a crystal oscillator from the measurements shown (without markers) in the bottom right plot of Fig. 3. In this case all four markers are needed in the computation to provide the spectrum bandwidth as well as the spectrum peak frequency.

V. Self-Focussing Drift Correction

Fig. 6 shows 70 second FFT sample of the spectrum of a 20MHz crystal oscillator on top of a five hour waterfall showing its drift history. A 12 Hz drift in 5hrs (18000sec) is 70 × 12/18000 = 46mHz drift during the 70 seconds of a single FFT spectrum sample. Thus the measured oscillator spectrum appears to have a 46mHz wide spectrum rather than the expected 1 to 2mHz. However all the information needed to correct for the drift lies in the FFT data of the present and a few previous complete FFT samples. It can be extracted and used for drift correction by suitable signal processing.

Fig. 6. 20MHz XO spectrum sample at top of display, with 10dB/div vertical scale, 20 Hz span, and 15mHz resolution bandwidth. Waterfall display below shows frequency drift over previous five hours. Sample update time of 70sec.

Fig. 7. Synthetic Aperture Radar (SAR) uses the motion of the radar to create a long synthetic antenna, which therefore has a narrow effective beam width along the track. Successive radar pulses form the ‘elements’ of the antenna. SAR image formation and focussing requires the exact radar track to be known typically to a few mm. Picture from [3].

Fig. 8. From SAR raw data (left) to a focused image (right). From UN FAO publication [4].
The required signal processing is essentially the same as is used for the Self-Focussing of Synthetic Aperture Radar (SAR) images. Fig. 7 shows the principle of SARs. And Fig. 8 compares the raw data with the radar image after application of the radar focussing algorithm. The algorithm is in two parts. The first part takes the assumed trajectory of the radar in the aircraft. The second part is ‘self-focussing’ where small phase corrections are made to each pulse return to compensate for errors in the radar antenna position relative to the flight path predicted by the aircraft inertial navigator.

This latter part of the algorithm can be applied to the drifting oscillator FFT data. It can be considered to be a three step process. First the drift profile is extracted from the data, and second the profile is thereafter applied to the same (stored) dataset to correct the FFT process according to the drift measure drift profile. A further fine-tuning can if necessary applied as a third step. In this way the full precision of the FFT spectrum filtering may be regained for the long FFT data records required for milliHerz spectrum accuracy and resolution.

VI. MULTIPLE SIGNAL SEPARATION AND MEASUREMENT

The SDR implementation of milliHerz resolution spectrum measurement allows the phase and frequency characteristics of each component of a complex multi component signal to be measured essentially simultaneously from the same data record. An example is the reception of multiple components around the nominal carrier frequency of shortwave frequency standards as shown in Fig. 9 and taken from a paper in this same conference [4].

![Fig. 9. 10MHz Signals received on a WinRadio Excalibur SDR at Lingfield, Surrey, UK, at 2147 UTC on 2 June 2014 from WWV Fort Collins USA (4700 mile NW path), BPM China (5200 miles NE path), PPE Rio de Janeiro, Brazil (5700 miles SW), WWVH Kekaha, Hawaii (7200 miles NNW) and Italian ‘pirate’ time signal located in Tuscany. Split into four main components, at +6, -6, -12, and -48Hz. Top left: 19 second reverse (upwards) waterfall with 1Hz resolution bandwidth. Snapshot at top right. 0 to 30MHz spectrum at bottom. 50 and 60Hz ‘hum’ sidebands are visible [5].](image)

Fig. 10 shows a three day horizontal right-to-left waterfall of 10Hz span of the received spectrum of components around 1680kHz in the new USA broadcast band allocation as received in the UK in Lingfield, Surrey, on an SDR-IQ receiver. 1680kHz is one of the AM carrier frequencies allocated in the USA.

![Fig. 10. Three day USA to UK propagation on 1680 kHz up to 0019 UTC on 17 Mar 2015. 4mHz resolution bandwidth used. The SDR-IQ reference oscillator temperature drift can be seen on the main signal. But variable excitation of upper and lower ‘sideband modes can be seen, with spacings of about 0.4Hz. Solar flare on 11 March caused auroral effects? From [5].](image)

VII. CONCLUSIONS

Software Defined Radios (SDRs) can be used for the measurement and display of phase noise spectra of single or multi-component spectra with milliHerz resolution. SDRs therefore have a number of significant advantages:

1. SDRs are low cost.
2. SDRs do not generate measurement ‘spurs’.
3. Direct milliHerz spectra can be plotted.
4. Long time records of complex spectra are feasible.
5. Complex multi-component spectra may be displayed and analysed.
6. ‘Self-focussing’ may be used to improve measured spectrum resolution and accuracy.

REFERENCES

Characterization of a set of Cryocooled Sapphire Oscillators at the $10^{-16}$ level with the three-cornered hat method

Christophe Fluhr, Serge Grop, Timothée Accadia, Ahmed Bakir, Yann Kersalé, Enrico Rubiola and Vincent Giordano
FEMTO-ST Institute
Besançon, France
Email: giordano@femto-st.fr

Benoît Dubois
FEMTO Engineering
Besançon, France

Abstract—In this paper, we present the characterization results of three Cryogenic Sapphire Oscillators (CSO) by using the three-cornered hat method. The three-cornered hat method permits us to extract the individual frequency instabilities. Thus this powerful tool helps us to choose the best mechanical and thermal CSO configurations. We tested two frequency counters requiring two different data processing and get almost the same results. The three CSOs reach a frequency instability better than $7 \times 10^{-16}$ between 1 s and 3,000 s integration times. The Allan deviation of the best CSO reaches a noise floor around $1.5 \times 10^{-16}$ at 200 s integration time. Although, the new CSO incorporates a Kyropoulos instead of a HEMEX sapphire resonator, it presents the almost the same frequency stability than the two other ones.

I. INTRODUCTION
The latest improvements of the Cryogenic Sapphire Oscillators (CSO) were realized in the frame of the ULISS project [1] and presented during the last EFTF in Neuchâtel [2]. Since, a third instrument, codenamed Absolut, was built. This set of three nearly identical CSOs is a part of the OSCILLATOR-IMP platform [3]. It constitutes a unique tool permitting to reach an unprecedented noise floor in frequency stability measurement. Their deployment is not totally finished: only the first CSO is completely optimized. At this step the three-cornered hat method has been applied to extract the CSO individual instabilities, which will help to determine the best CSO configuration. Indeed although the 3 CSOs are based on the same principle they slightly differs from design details and are in different state of progress.

II. CSO DESIGNS
A. Cryogenic Sapphire Oscillator
The CSO incorporates a frequency reference constituted by a whispering gallery mode sapphire resonator placed in the center of a cylindrical copper cavity that can be cooled down to 4 K. High order whispering gallery modes that can be excited in this resonator are characterized by a high energy confinement in the dielectric due the total reflection at the vacuum-dielectric interface. The different resonators we designed operate on quasi-transverse magnetic whispering gallery modes as WGH$_{m,0,0}$ where $m$ is the number of wavelength in the resonator along the azimuth. The sapphire resonator is placed into a cryostat and in thermal contact with the second stage of a pulse-tube cryocooler delivering typically 0.5 W of cooling power at 4 K. The resonator temperature can be stabilized above 4 K at ±200 μK. Providing the azimuthal index is sufficiently high ($m > 13$ typically) the Q factor can achieve $1 \times 10^9$ at the liquid-He temperature. The CSO is basically a transmission oscillator: the two-port cryogenic sapphire resonator being inserted in the positive feedback loop of a microwave amplifier. A Pound-Galani and a power servos complete the system to control the phase and the power of the oscillating signal.

In our most advanced CSO, a 54 mm diameter and 30 mm thick resonator is operated on the WGH$_{15,0,0}$ mode. The frequency of the WGH$_{15,0,0}$ mode is intentionally set to 10 ± 4 MHz off the nominal 10 GHz [4]. The CSO synthesizer output can be adjusted precisely with a DDS [5].

B. New cryostat
The development of a third CSO was the opportunity to tests other designs and materials. The new cryostat has been developed by Absolut System [6] in collaboration with FEMTO-ST. Absolut is shown in figure 1. It is based on a PT407 Cryocooler from Cryomech, which is more powerful than the two others CSOs equipped with PT405. Conversely to the two first ones, the mechanical decoupling of the cryocooler head (bellows) is implemented. The internal mechanical decoupling is realized with copper foil thermal straps instead of copper braids. This cryostat also incorporates different copper thermal shields (instead of aluminium covered with MLI). A thermal ballast is placed between the resonator and the 2nd stage of the Pulse-Tube to filter temperature fluctuations.

Absolut incorporates a Kyropoulos sapphire crystal provided by the company Precision Sapphire Technology...
(PST) [7] instead of the common HEMEX sapphire resonator [8]. This resonator is a preliminary version attended to test another source of high quality sapphire crystals. To limit the cost of this very preliminary sample, its mechanical tolerances were relaxed leading to a higher offset frequency. Moreover the resonator is just a sapphire cylinder with a 5 mm hole along its axis. A screw passing through is used to maintain the sapphire in the cavity. In the two other CSOs, the resonator is equipped with a spindle, which limits the stress induced in the effective resonator volume. A higher long-term drift is expected with the simple Absolut resonator mounting. Unloaded Q-factors as high as 700 million have been observed in preliminary experiments. The current resonator adjustment is totally optimized and the loaded Q-factor is 400 million. Eventually and in order to test different components, the oscillating loop has been implemented on a table and linked to the cryostat with commercial microwave 1m length cables. It thus experiences all the temperature variations and vibrations of the experimental room.

The status of each CSO is given in the table I.

---

**TABLE I. MAIN CHARACTERISTICS OF EACH CSO**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Marmotte</th>
<th>Uliss</th>
<th>Absolut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loaded Q factor</td>
<td>1 Billion</td>
<td>350 Million</td>
<td>400 Million</td>
</tr>
<tr>
<td>Crystal type</td>
<td>HEMEX</td>
<td>HEMEX</td>
<td>Kyropoulos</td>
</tr>
<tr>
<td>Turnover point temperature</td>
<td>6.2379 K</td>
<td>5.7656 K</td>
<td>5.2217 K</td>
</tr>
<tr>
<td>Cryocooler</td>
<td>PT405</td>
<td>PT405</td>
<td>PT407</td>
</tr>
<tr>
<td>Cryostat manufacturer</td>
<td>Oxford Instruments</td>
<td>Oxford Instruments</td>
<td>Absolut System</td>
</tr>
<tr>
<td>Thermal Ballast</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Status</td>
<td>Optimized</td>
<td>Resonator needs cleaning and coupling adjustment</td>
<td>Resonator needs cleaning and coupling adjustment, Oscillating loop on table</td>
</tr>
</tbody>
</table>

---

**Fig. 1. Mechanical and thermal Cryostat specifications**

---

**III. THREE-CORNERED HAT**

**A. Setup and data acquisition**

With this third CSOs, Absolut, we implement three-cornered hat method [9]. This method is useful for determining the individual stabilities of units having similar performance. Despite the three oscillators have a different design of cryostat, we expect that their noises level are equivalent. Each has a different frequency, this one being specific to sapphire dimensions. The three beatnotes frequency are obtained by mixing the output of each CSO with the two others CSOs. In order to not be limited by the counter's resolution we chose to increase the heterodyne factor of the two beatnotes Absolut-Marmotte and Absolut-Uliss. We make a down conversion by mixing beatnotes with a signal from a direct digital synthesizer (DDS). DDSs are clocked on a 100 MHz MASER source manufactured by T4Science [10]. This three beatnotes are connected to a frequency distribution amplifier product by Timetech [11]. The frequencies are counted by two different instruments, the K+K FXE SCR and the GuideTech GT668. The entire setup is described on Fig 2.

a) **K+K FXE SCR**: is a multi-channel phase recorder [12]. It uses the picket fence measurement method. Frequency is determined by counting period’s occurrences during a certain length of time. Reference signal defines the time interval. All channels are operated in the same time. The sampling rate is 1,000 samples per second. An averaging is applied to get one data per second for each channel. This averaging permit to increase the resolution. All data are post-processed on Stable32 to extract frequency stability of each CSO.

b) **GuideTech GT668**: is a high-speed continuous time interval analyzer (CTIA) which logs the time of occurrence of
events at its inputs [13]. Event are defined as a signal voltage
crossing a specified threshold. These 'time-tags' are used to
calculate the signal frequency over any time windows \( \tau \) (1 s in
our case) and with no dead time. In order to improve the
GT668 resolution, the processing involves a triangular
averaging of frequency measurements. The sampling rate is
40,000 samples per second and could be improve to
4 MSa/s [13]. The instrument have only two input channels
which seems to be a problem for three-cornered hat method. A
third 'virtual' channel is generated by software. Equation (1)
shows the relative frequency stability, \( \sigma \), of one channel with
due devices A,B,C described how to the third channel is
reconstructed.

\[
\sigma_{AB}^2 = \sigma_A^2 + \sigma_B^2
\]

Then with two channels, the third can be reconstructed with :

\[
\sigma_{AC}^2 = \sigma_{AB}^2 - \sigma_{BC}^2
\]

according that all CSO are totally uncorrelated. The
implementation of this instrument is only on preliminary
startup. We use Stable32 to calculate the Allan deviation.

The shapes given by the GuideTech's results are
 equivalent. Nevertheless the values are slightly different. The
first possible origin of error is that the data were not acquired
at the same time. The second source of error is the data post
processing of the GT668's. The preliminary results are traced
on Fig. 4.

<table>
<thead>
<tr>
<th>Integration time ( \tau ) in second</th>
<th>( \sigma ) (Marmotte)</th>
<th>( \sigma ) (Uliss)</th>
<th>( \sigma ) (Absolut)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.32x10^{-16}</td>
<td>5.27x10^{-16}</td>
<td>7.29x10^{-16}</td>
</tr>
<tr>
<td>8</td>
<td>3.98x10^{-16}</td>
<td>6.45x10^{-16}</td>
<td>4.55x10^{-16}</td>
</tr>
<tr>
<td>64</td>
<td>2.38x10^{-16}</td>
<td>3.82x10^{-16}</td>
<td>4.32x10^{-16}</td>
</tr>
<tr>
<td>256</td>
<td>1.55x10^{-16}</td>
<td>4.58x10^{-16}</td>
<td>6.55x10^{-16}</td>
</tr>
<tr>
<td>1024</td>
<td>3.81x10^{-16}</td>
<td>1.08x10^{-15}</td>
<td>1.00x10^{-15}</td>
</tr>
</tbody>
</table>

Fig. 3. Allan deviation for each CSO from \( \text{K+K} \).

Fig. 4. Allan deviation for each CSO from GuideTech.

II. RESULTS

The results of individual Allan deviation for the three
CSOs are given in Fig. 3 concerning the K+K, the data are
drift removed.

Three-cornered hat show us that Marmotte reaches a noise
floor about \( 1.4\times10^{-16} \) at 100 s integration time. The difference
between the oscillators from 1 s to 500 s, could be explain by
the Q factors variation. All CSOs are equivalent at 1 s
integration time about \( \sigma(1s)=5 \) or \( 7\times10^{-16} \). For Uliss, the hump at
8 s is due to the sapphire temperature regulation. A better
adjustment of the control loop parameters should reduce the
effect. We notice that the three oscillators present a long term
drift after 1,000 s, it is the effect of room's air conditioning.
The values of Allan deviations are given in table II.
III. CONCLUSION

The results show an Allan deviation of $1.5 \times 10^{-16}$ at 100 s integration time measured. Two instruments were used. The three-cornered hat method allows to discriminate among the CSO characteristic offering news possibilities to improve CSOs performances. Thereby different solutions can be tested and analyzed. This solutions may concern the thermal configuration, the mechanical design or even the oscillator control. It will be possible to reach better frequency stability by combining the best choices. The startup of our new CSO, Absolut, allows us to propose an alternative cryostat design. This oscillator achieve the same results than the two others.

Acknowledgment

The authors would like to thank the Programme Investissements d'Avenir (PIA).

References

[8] Precision Sapphire Technologies, Ltd.: http://www.sapphire.it
Generating Entanglement between Atomic Spins with Low-Noise Probing of an Optical Cavity

Kevin C. Cox, Joshua M. Weiner, Graham P. Greve, and James K. Thompson
JILA and Department of Physics
University of Colorado and National Institute of Standards and Technology
Boulder, Colorado USA 80309
Email: jkt@jila.colorado.edu

Abstract—Atomic projection noise limits the ultimate precision of all atomic sensors, including clocks, inertial sensors, magnetometers, etc. The independent quantum collapse of $N$ atoms into a definite state (for example spin up or down) leads to an uncertainty $\Delta \theta_{SQL} = 1/\sqrt{N}$ in the estimate of the quantum phase accumulated during a Ramsey sequence or its many generalizations. This phase uncertainty is referred to as the standard quantum limit. Creating quantum entanglement between the $N$ atoms can allow the atoms to partially cancel each other’s quantum noise, leading to reduced noise in the phase estimate below the standard quantum limit. Recent experiments have demonstrated up to 10 dB of phase noise reduction relative to the SQL by making collective spin measurements. This is achieved by trapping laser-cooled Rb atoms in an optical cavity and precisely measuring the shift of the cavity resonance frequency by an amount that depends on the number of atoms in spin up. Detecting the probe light with high total efficiency reduces excess classical and quantum back-action of the probe. Here we discuss recent progress and a technique for reducing the relative frequency noise between the probe light and the optical cavity, a key requirement for further advances.

I. INTRODUCTION

Atoms and molecules make excellent sensors of time [1], [2], accelerations [3], and fields [4], and enable precise tests of fundamental physics [5]–[7]. This is because quantum mechanics provides certainty that the atoms can be made nearly identical using modern optical pumping and laser-cooling and trapping techniques. The value of the physical quantity to be sensed is most often encoded in the impact of the physical quantity on the rate at which a quantum phase develops between quantum states. Here, we will consider spin states, but quantum phases can also be measured between other types of quantum states, such as momentum states in matter wave interferometers. The high accuracy provided by quantum mechanics must be balanced against the fundamental quantum uncertainty that quantum mechanics imposes on the measurement of the evolved phase. The uncertainty can be mitigated by using many independent atoms $N$ in parallel, as is done in optical lattice clocks or matter wave interferometers. The independent quantum collapse or projection noise [8] of each atom is averaged down to an uncertainty in the estimate of the evolved phase scaling as $\Delta \theta_{SQL} = 1/\sqrt{N}$ radians, a limit known as the standard quantum limit.

Quantum entanglement between the $N$ atoms can allow the randomness in the measurement-induced collapse of one atom to be partially cancelled by biasing the collapse of other atoms [23], [24]. Theoretically, it is possible to produce entangled states such that quantum projection noise is cancelled by this compensation. However, the last atom to be measured will not have another atom present to cancel its projection noise. The “noise of the last atom” leads to a more fundamental phase estimation uncertainty $\Delta \theta_{HL} \approx 1/N$, known as the Heisenberg limit. For large numbers of atoms, this enhancement is potentially dramatic.

Small numbers of ions of order 10 or fewer [19]–[22], [25], [26] have achieved phase imprecision close to the Heisenberg limit. However, it is only recently that larger ensembles of atoms have been entangled and shown to improve phase estimation beyond the standard quantum limit. The approaches have included twisting operations using atomic collisions [14], [15], [17], [27], [28] or an optical cavity [18], parametric two-mode squeezing driven by atomic collisions [13], [16], [29], and coherence preserving collective measurements of atoms in optical cavities [9], [10], [30], [31] and in free space [11], [12],
Figure 1 attempts to partially summarize the observed enhancements in phase estimation sensitivity relative to the standard quantum limit. The results are plotted versus atom or ion number since the rms phase estimation sensitivity of the standard quantum limit scales as $1/\sqrt{N}$.

One heuristic way to think about the importance of scaling entanglement to larger atom number is to consider the following scenario: Imagine one can entangle 10 atoms to reach the Heisenberg limit of 100 mrad, i.e., a factor of 10 reduction in noise variance with respect to the original 10 atom SQL of 320 mrad. In comparison, recent entanglement generation results in large ensembles approaching $10^6$ atoms are very far from the $10^6$ atom Heisenberg limit of 0.001 mrad, but they do realize a factor of 10 reduction in noise variance with respect to the original $10^6$ atom SQL of 1 mrad. To achieve the same phase estimation imprecision with collections of 10 atoms, one would have to perfectly entangle each collection of 10 atoms to their Heisenberg-limited sensitivity and then successfully repeat this in parallel $10^3$ times. It is clear that this would be daunting compared to the simple operations presented here.

II. COHERENCE PRESERVING MEASUREMENTS

Here, we will focus on making highly precise collective measurements using atoms laser-cooled and trapped inside of an optical cavity. This approach was used to obtain an enhancement in the phase variance by $(\Delta \theta_{SQL}/\Delta \theta)^2 = 10$ (or $10 \log_{10}[(\Delta \theta_{SQL}/\Delta \theta)^2] = 10$ dB) relative to the standard quantum limit [9]. More recently, we have made a preliminary observation of $(\Delta \theta_{SQL}/\Delta \theta)^2 = 23(5)$. These are the directly observed enhancements with no background subtractions or corrections for imperfections or readout noise. Thus, this is the enhancement one would expect in a Ramsey measurement relative to using a non-entangled collection of atoms, although one must be careful to consider the effects of single particle and collective dephasing and decoherence [33]–[35]. We report the reduction in the phase variance as this reflects the reduction in the amount of resources required to achieve the same imprecision: 23 times less Ramsey evolution time or 23 times fewer atoms. These reduced resource requirements might be applied to increase measurement bandwidth or reduce systematic errors due to atomic collisions.

To reduce the quantum noise, we essentially measure it and subtract it out, as shown in Fig. 2. This approach was first proposed for free space probing in [36], and our experimental and theoretical work in optical cavities is described in Refs. [9], [10] and [37]. We make a pre-measurement of the quantum spin projection noise of the total atomic ensemble, with measurement outcome labeled $J_{zp}$. The measurement outcome is then subtracted from a final measurement of the total spin projection, with measurement outcome labeled $J_{zf}$. The atomic projection noise cancels in the differential quantity $(J_{zf} - J_{zp})$. If the measurement leaks no information to the environment about which atoms are in spin up versus spin down, then each atom remains in a superposition of spin up and spin down. In a Bloch vector picture, the first measurement localizes the initially blurry quantum state into a region with less blurriness along $\hat z$ and thus $\theta$, at the expense of enhanced blurriness in an orthogonal spin projection. To utilize this state to make a precise measurement, one would insert a Ramsey pulse sequence between the two measurements.

Fig. 2. (a) Measurement sequence for probing the atoms. A coherent spin state is prepared by optically pumping all of the atoms into spin down. Each atom is rotated into a superposition of spin up and down, corresponding to the total spin or Bloch vector oriented along the equator. Two consecutive measurements of the spin projection $J_z$ are then performed (100 $\mu$s each), with the pre and final measurement outcomes for a single trial labeled $J_{zp}$ and $J_{zf}$. The sequence is then repeated many times. (b) Measurement outcomes versus trial number. Classical rotation noise in the $\pi/2$ pulse causes classical excess noise in the observed spin noise fluctuations, so here we display simulated Gaussian noise with rms distribution equal to the predicted projection noise level fluctuations $\Delta J_{zCSS}$ (left axis, red open circles). The measured differential quantity $J_{zf} - J_{zp}$ (blue, right axis) shows partial cancellation of both the quantum projection noise and the excess classical noise. The differential quantity’s noise variance is 50 times below the projection noise level (or 171 dB). (c) (left) The red data can be visualized as arising from a fundamental blurriness of the orientation of the collective Bloch vector. (right) The measurement process projects the Bloch vector into a squeezed state with reduced uncertainty in the polar angle, at the expense of increased uncertainty in the azimuthal angle. Accounting for a reduction in the size of the Bloch vector due to free space scattering and dephasing, allows a net preliminary improvement in angular variance of $23(5)$ or $13.7(10)$ dB below the standard quantum limit for unentangled atoms.

Measuring the spin projection $J_z = (N_u - N_d)/2$ is achieved by attempting to count the number of atoms in spin up $N_u$ versus spin down $N_d$. In our latest work, the effective spin-half system is composed of the two ground hyperfine states of $^{87}$Rb, $|\uparrow\rangle = |5^2S_1/2, F = 2, m_F = 2\rangle$ and $|\downarrow\rangle = |5^2S_1/2, F = 1, m_F = 1\rangle$ (see Fig. 3). We convert the atom counting problem into a frequency measurement by placing the atoms inside of an optical cavity of finesse $F = 2700$ and arranging things such that the cavity mode’s resonant frequency is dispersively shifted by an amount that depends on the atomic population $N$. Classically, the atoms act as a medium with index of refraction $n(N)$. As a result, the optical path length between the mirrors is modified, causing the cavity resonance frequency to shift. Standard microwave $\pi$ pulses can be used to swap the populations between spin states, allowing the population $N_u$ to also be determined if needed.

To achieve a state-dependent shift of the cavity resonance frequency, we tune the length of the cavity such that a TEM$_{00}$ is 500 MHz blue detuned from the transition frequency between $|\uparrow\rangle$ and the optically excited state $|e\rangle = |5^2P_{3/2}, F' = 3\rangle$ with transition wavelength $\lambda = 780$ nm. The $|\downarrow\rangle$ to $|e\rangle$ transition is much further off-resonance with
the cavity mode due to the ground state hyperfine splitting of 6.8 GHz, and its interaction with the cavity can thus be largely ignored.

The cavity has a power decay rate $\kappa = 2\pi \times 3.05(5)$ MHz, a mirror separation $L = 1.85(1)$ cm, a free spectral range 8.10(2) GHz, mirrors with radius of curvature $R_c = 5$ cm, a mode waist $w_0 = 69 \mu m$, and mirror power transmission coefficients of $T = 2010 \times 10^{-6}$ and $T = 130 \times 10^{-6}$. The peak single-atom Jaynes-Cummings coupling on the cycling $|\uparrow\rangle$ to $|\downarrow\rangle$ transition is $g = 2\pi \times 526$ kHz. We typically load an effective atom number $N = 4 \times 10^5$ into the lattice [30], [38].

To sense small applied rotations, it is sufficient to precisely measure changes in the population $N_\uparrow$ only (see [9] for details.) Therefore, we concentrate on the case of making two consecutive measurements of the dressed cavity resonance frequency with measurement outcomes $f_{cp}$ and $f_{c,f}$ from which we can extract a pre-measurement and final measurement of the population $N_\uparrow$.

The goal then is to make the rms noise in the difference frequency $\Delta f_d = \Delta(f_{c,f} - f_{cp})$ less than the fluctuations in the individual frequencies due to quantum projection noise $\Delta f_{P|JN}$. For the above experimental parameters we find an rms fluctuation $\Delta f_{P|JN} = 97$ kHz. The experimental challenge is to achieve $\Delta f_d / \Delta f_{P|JN} \ll 1$, while also disturbing the atomic system as little as possible. In the following subsections, we provide details on our current locking scheme and characterize its performance both in terms of the technical noise contribution to $\Delta f_d$ and the associated small loss of atomic coherence.

III. CANCELING LASER-CAVITY FREQUENCY NOISE

A. Experimental Scheme

The cavity frequency shift can be easily blurred by the typical 200 kHz FWHM Lorentzian linewidth of standard external cavity diode lasers (ECDLs). In addition, vibrations can cause the empty cavity’s resonance frequency $\omega_c$ to jitter. We use the optical lattice beam at 823 nm that is used to trap the atoms in the cavity, to also stabilize the cavity resonance frequency with a few kHz bandwidth. In previous work, we then used DBR lasers narrowed to a few kHz to probe the laser resonance frequency [39]. Low frequency relative noise limited the technical noise to 17(2) dB below the projection noise level. Here we will describe a robust approach in which we start with 200 kHz linewidth ECDLs and demonstrate a technical noise floor on $\Delta f_d$ that is as much as 27 dB below the projection noise level.

Two ECDLs are used to probe two different longitudinal modes of the optical cavity. This approach is related to the approach described in Ref. [30]. Figure 4a depicts the physical layout used to probe the cavity. Figure 4b outlines the various frequency components on the two probes and how the frequency chain is locked.

The cavity-probe is tuned to a mode far from resonance with the atomic transition, while the atomic-probe is tuned to the original mode that is close to resonance with the atomic transition. The cavity-probe laser’s frequency is Pound-Drever-Hall locked using phase modulation sidebands at $f_{PDH} = 9.7$ MHz. The cavity-probe light is $\sigma^-$ polarized when it hits the cavity.
the cavity, allowing it to be polarization-separated from the atomic-probe light that is $\sigma^+$ polarized. The cavity-probe is detected using an avalanche photodiode with a gain of 100. The laser is served to the cavity with $\sim 1$ MHz unity gain frequency.

The atomic-probe laser is then phase-locked to the cavity-probe laser with unity gain frequency also close to 1 MHz. The direct beat frequency of the two lasers is approximately 122 GHz, too high of a frequency to easily detect. Instead, we partially bridge the frequency gap by using a fiber phase modulator driven at 13.6 GHz to place high order sidebands on light derived from the cavity-probe laser. The microwave modulation source is derived from the low phase noise 6.8 GHz local oscillator used to drive atomic rotations \cite{40}. The 9th order sideband is within 1 GHz of the atomic-probe laser and can be directly detected. The phase of the detected signal is then phase locked to a frequency reference supplied by an AD9959 direct digital synthesizer (DDS).

Having established a chain to stabilize the atomic-probe laser frequency relative to the optical cavity, we now consider the atomic-probe detection scheme. The atomic-probe light is reflected from the cavity and detected using a heterodyne local oscillator (LO) also derived from the atomic-probe laser. To ensure good common mode cancellation of the atomic-probe’s phase noise in the heterodyne detection, the total path lengths between the point of separation from the LO path to the point of re-overlap with the LO path are made equal to within 8 cm. Before striking the cavity, the atomic-probe light is shifted by an acousto optic modulator (AOM) by +82.5 MHz and then phase modulated at 137.5 MHz. This produces a small frequency sideband that is tuned to resonance with the optical cavity mode by adjusting the frequency of the DDS used to phase lock the atomic-probe laser to the cavity-probe laser.

The frequency component that interacts resonantly with the cavity mode appears in the detected rf spectrum at 55 MHz. Having established a chain to stabilize the atomic-probe laser frequency relative to the optical cavity, we now consider the atomic-probe detection scheme. The atomic-probe light is reflected from the cavity and detected using a heterodyne local oscillator (LO) also derived from the atomic-probe laser. To ensure good common mode cancellation of the atomic-probe’s phase noise in the heterodyne detection, the total path lengths between the point of separation from the LO path to the point of re-overlap with the LO path are made equal to within 8 cm. Before striking the cavity, the atomic-probe light is shifted by an acousto optic modulator (AOM) by +82.5 MHz and then phase modulated at 137.5 MHz. This produces a small frequency sideband that is tuned to resonance with the optical cavity mode by adjusting the frequency of the DDS used to phase lock the atomic-probe laser to the cavity-probe laser.

The frequency component that interacts resonantly with the cavity mode appears in the detected rf spectrum at 55 MHz. The signal is IQ demodulated using phase coherent channels from the same DDS board that are used to provide the various frequency shifts and modulations. We determine the phase of the reflected light from the two quadrature signals. The change in phase as a function of frequency detuning from the cavity resonance sets the conversion factor between detected phase and detuning. From the measured phase noise of the reflected field, we can then estimate the relative frequency noise between the atomic-probe laser and the cavity mode that it is probing.

A representative power spectral density of instantaneous frequency noise $S_v(f)$ is shown in Fig. 5a. In the central flat region $S_v \approx 1.5 \times 10^3$ Hz$^2$/Hz. This corresponds to the instantaneous frequency noise of a laser with Lorentzian FWHM $\Delta \nu = \pi \times S_v = 5$ kHz, significantly smaller than the initial laser linewidth of 200 kHz.

The roll off at high frequency results from 300 kHz anti-aliasing low pass filters after the IQ demodulation. The rise at low frequencies is largely due to uncontrolled relative path length changes between the atomic-probe LO and atomic-probe paths. We have found it unnecessary to stabilize this path length phase for the preliminary results presented here because it only required 200 $\mu$s to measure the differential quantity $f_{cf} - f_{cp}$.

We are interested in characterizing the noise in $f_{cf} - f_{cp}$. Each measurement is the average of the measured frequency in a window of length $T_m$ and the two measurement windows have a time gap $t$ between them. The variance is obtained by integrating $(\Delta f_{d})^2 = \int_0^{\infty} S_v(f)T(f) df$. The transfer function is $T(f) = 4 \sin^2(\pi f(T_m + t)) \sin^2(\pi fT_m)/(\pi fT_m)^2$. Figure 5b shows the measured noise variance $\Delta f_{d}$ as a function of the measurement window length $T_m$ with $t = 0$ $\mu$s. For this data, the minimum is $\Delta f_{d} = 6$ kHz at $T_m = 100$ $\mu$s. The rise at longer times is dominated by path length fluctuations.

The path length noise has recently been suppressed by actively stabilizing the relative path length. This is achieved using the much stronger probe component that is detuned from the cavity mode by 137.5 MHz, and produces a signal in the atomic-probe heterodyne detector at 82.5 MHz. Appropriate demodulation allows us to derive an error signal to phase-lock the relative phase between probe path and the heterodyne reference path. The phase is stabilized by adjusting the phase of the 82.5 MHz driving the AOM that shifts the atomic-probe.
B. Degree of coherence preservation

The previous measurements were made at very high probe powers with the no atoms in the cavity. As a result, the photon shot noise and technical noise sources of the detectors were negligible compared to contributions from other technical noise sources. However, as the power in the cavity-probe is increased, the amount of squeezing may become limited by additional scattering of light from the atoms, potentially leading to single-particle wavefunction collapse (loss of signal) and Raman transitions to other ground hyperfine states (a source of additional noise.) Inhomogeneous differential light shifts of the spin transition frequency can also lead to dephasing that can be spin-echoed away, but perhaps imperfectly.

Figure 6a shows the noise variance $\langle \Delta f_d \rangle^2$ between two measurements of the cavity resonance frequency versus the cavity-probe power incident on the cavity $P_c$. Here the atomic-probe power is increased such that its photon shot noise contribution is negligible. The right hand axis translates the noise variance into an equivalent uncertainty relative to a contribution is negligible. The right hand axis translates the noise variance scales as $\langle \Delta f_d \rangle^2 \propto 1/P_c^2$, indicating that lower powers $P_c$ might be utilized with improved photodetection.

The cavity-probe induces a differential light shift of the transition frequency $\omega_{\uparrow \downarrow}$ between down and up. We measure the differential shift of the clock transition frequency $|F = 1, m_F = 0\rangle$ to $|F = 2, m_F = 0\rangle$ versus $P_c$ in Fig. 6b. Appropriately rescaling for transition strengths and detunings gives an average shift of $\omega_{\uparrow \downarrow}$ by 8.5 kHz per $\mu$W. The shift is highly inhomogeneous and leads to dephasing. However, preliminary spin-echo measurements have shown that the atomic contrast (i.e., length of the Bloch vector) is reduced by less than 5% at $P_c = 0.5 \mu$W with 40 $\mu$s measurement windows. Thus, the atomic state is not significantly decohered by the cavity stabilization presented here.

IV. Conclusion

Coherence preserving measurements are a powerful technique for producing large amounts of entanglement in large atomic ensembles. Here the figure of merit is the enhancement in the estimation of a quantum phase relative to the standard quantum limit. This figure of merit is particularly compelling because it directly connects to the application of entanglement to precision measurements with atoms and ions. The geometry used here for proof-of-principle experiments may be amenable to enhancing optical lattice clocks [1], [2] beyond the standard quantum limit and may also allow for reduced dead time due to the non-destructive nature of the readout [41]–[43]. Additionally, the scheme presented here may allow enhancements to atom interferometers [44] used for rotation sensing [3], measurements of gravitational acceleration [45], or even searches for gravitational waves [46]. Because the probe light can be switched on and off, the entanglement can be generated without perturbing the atoms during the critical Ramsey phase evolution time. Here we have presented a scheme to realize large amounts of entanglement with standard external cavity diode lasers. Work is currently underway to improve the net quantum efficiency for detecting the atomic-probe for even greater amounts of entanglement enhancement of phase estimation.

ACKNOWLEDGMENT

The authors would like to thank Matthew A. Norcia for helpful discussions. All authors acknowledge financial support from DARPA QuASAR, ARO, NSF PFC, and NIST. K.C.C. acknowledges support from NDSEG. This work is supported by the National Science Foundation under Grant Number 1125844. Part numbers are given as technical information only, and do not represent endorsement by NIST.

REFERENCES


Laser Stabilization on Velocity Dependent Nonlinear Dispersion of Sr Atoms in an Optical Cavity

Bjarke T. R. Christensen*, Stefan A. Schäffer*, Martin R. Henriksen*, Philip G. Westergaard†, Jun Ye‡ and Jan W. Thomsen*
*Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark
†Danish Fundamental Metrology, Matematiktorvet 307, DK-2800 Kgs. Lyngby, Denmark
‡JILA, National Institute of Standards and Technology and University of Colorado, Boulder, CO
Email: bjarkesan@nbi.ku.dk

Abstract—The development of simple and reliable high stability clock lasers is of great importance for future state-of-the-art optical clocks [1]–[5] and for future transportable optical clocks [6], [7]. Further development of clock lasers with better stability has so far been hindered by thermal noise in the reference cavity used for laser stabilization and conventional approaches for improvements may be technically challenging. It has been proposed [8]–[11] to improve the stability and reduce the complexity of state-of-the-art laser frequency stabilization by exploiting cavity QED systems consisting of atoms with a narrow optical transition coupled to a single mode of an optical cavity.

The laser stabilization performance of a cavity QED system is affected by a number of system parameters such as the finite temperature of the atoms, the number of involved atoms and the laser power [12]–[14]. However, the dynamics of those elements have not yet been fully explored. Here we present a simple cavity QED system consisting of laser cooled strontium-88 atoms coupled to an optical cavity. We relate measurable quantities to the complex transmission coefficient which relates the input field to the output field. The optimal input power for stabilizing a laser to this system is experimentally determined and the optimal shot-noise-limited linewidth of the system is evaluated to 500 mHz. Furthermore, theoretical shot-noise-limited linewidths of similar cavity QED systems are evaluated for a number of different two electron systems.

I. INTRODUCTION

State-of-the-art atomic clocks rely on highly coherent light sources to probe narrow optical transitions [1]–[5]. Stability and accuracy at the $10^{-18}$ fractional level has been reported by several groups in the recent years [1]–[4]. However, surpassing the stability levels of the current state-of-the-art atomic clocks has so far been hindered by thermal noise in the reference cavity used for laser stabilization [15]–[17], and further improvements may be technically challenging. Novel proposals suggest an alternative strategy by employing so-called active optical frequency standards, where atoms are probed on narrow clock transitions inside an optical cavity, [8]–[10], [18]. This brings non-linear effects into the system dynamics that could potentially lead to a laser stability comparable to or better than the current state-of-the-art [14]. However, this non-linear dynamics makes the determination of the optimal parameters non-trivial [14], [19] and detailed characterization of the system behavior must be carried out. Originally the proposed active optical frequency system requires atoms trapped in optical lattices [10] or the use of beam-line systems with very low atomic beam velocities [20], increasing the complexity of the system and making it less attractive for transportable optical atomic clock systems.

In this work, the laser stabilization performance of laser cooled strontium-88 atoms located inside an optical cavity is experimentally investigated. The non-linear enhancement of the systems performance as a frequency discriminator is exploited, while the complexity of the system is kept to a minimum by operating at relatively high sample temperatures of the order of a few mK. Recently a theoretical model has been developed for our cavity QED system that takes into account the finite temperature of the atoms [14]. This model shows significant deviation from the zero temperature model [10]. In the limit of $T = 0$ K the laser stabilization performance of such a cavity QED system is limited by the so-called bi-stability region, where the input field intensity may yield several solutions for the steady state intra-cavity field. This is not the case at the finite temperatures considered here and this enables access to an interesting experimental region. In the following we explore the dynamics and identify optimal parameters relevant for laser stabilization in our system. In addition, based on recent theoretical methods we evaluate the potential laser linewidths achievable by a shot noise limited laser stabilization scheme for a number of different two electron systems.

II. EXPERIMENTAL SETUP

Our cavity QED system consists of a sample of laser cooled strontium-88 atoms with temperatures of few mK coupled to a low-finesse optical cavity ($F = 85$). The strontium atoms are trapped and cooled in a Magneto optical Trap (MOT) on the $^{1}S_{0} \rightarrow ^{3}P_{1}$ transition at 461 nm. The cooled atoms are trapped inside a vacuum chamber while the mirrors of the optical cavity are placed outside the vacuum chamber. The system is probed on the $^{1}S_{0} \rightarrow ^{3}P_{1}$ intercombination transition at 689 nm while the cavity is forced to be in resonance with the probe laser at all times. The probing time is 100 ms and the 461 nm light is turned off during the probe period. The cavity transmitted field amplitude and non-linear phase shifts induced by the atoms in the cavity is measured by cavity-enhanced FM spectroscopy [21], see Figure 1.

The collective cooperativity $C$ of this system, describing how strongly the atomic sample is coupled to the cavity mode, has been measured to $C = \frac{g^2}{\kappa \gamma} = 630$, with a collective atom-light coupling parameter, $g = 2.8$ MHz, a cavity decay...
Amp

parameter, $\kappa = 5.8$ MHz, and a natural linewidth of the intercombination transition $\gamma = 7.6$ kHz. In this parameter regime, the system experiences no bi-stable region limiting its potential performance. The saturated dispersion feature of this system has been proposed as a promising error signal for laser stabilization [10]. The slope of the dispersion feature around resonance is important for the determination of the achievable frequency stability of the system. Hence, the non-linear input power dependence of this slope value is measured and characterized.

Experimentally, we employ cavity enhanced FM spectroscopy for high contrast phase dispersion measurements by using the so-called noise-immune cavity-enhanced optical heterodyne molecular spectroscopy (NICE-OHMS) technique [21]. In the NICE-OHMS technique, sidebands are generated on the probe laser carrier by phase modulation with a modulation frequency $\Omega$ corresponding to the Free Spectral Range (FSR) of the cavity, $\Omega = \text{FSR}$, here 500 MHz. Notice we operate in a regime where the natural linewidth $\gamma$ is significantly smaller compared to the FSR of our cavity. The beat components of the transmitted carrier and sideband intensities are then measured by a high bandwidth avalanche photo detector (APD) and demodulated by frequency mixing with the drive frequency $\Omega$ or $2\Omega$. We have developed the phase technique previously to measure the weak phase dispersion of narrow line systems with high signal-to-noise ratio [14] and it is here extended to also include measurement of the attenuation of cavity transmitted signals. Due to the back action of the atoms on the cavity, the signal demodulated at $2\Omega$ is not equivalent to direct cavity transmitted signal, i.e., the direct attenuation of the carrier plus sidebands, and care must be adopted when comparing measured signals to theoretical models.

A. Theory of measurements

In order to establish a connection between our measured quantities and the complex transmission coefficient of the input field we perform an analysis of the cavity transmitted field phase and attenuation. Our input probe field can be written in terms of a carrier and two sidebands:

$$E_n = E_0 \left( J_0(x)e^{i\omega_c t} + J_1(x)\epsilon^{i(\omega_c + \Omega)t} - J_1(x)\epsilon^{-i(\omega_c - \Omega)t} \right),$$

where $J_0(x)$ and $J_1(x)$ are the zeroth and first regular Bessel functions signifying the amount of power transferred from the carrier to the sidebands, $x$ the EOM modulation index, $\omega_c$ is the carrier frequency and $\Omega$ is the free spectral range of the cavity. Each transmitted component of the cavity field is modified by the cavity according to the well know expression [22]:

$$E_{\text{out}} = \frac{T e^{i\beta}}{1 - R e^{i2\Omega}} E_n = \chi E_n,$$

where $\chi$ is the complex transmission coefficient describing the phase change and the attenuation of the field, $T (R)$ is the power transmission (reflection) coefficient of the cavity mirrors, and $\beta$ is the complex phase of the cavity transmitted field. The carrier is close to the atomic resonance while the sidebands are far off resonance ($\Omega = 500$ MHz). Since the cavity is kept on resonance with the carrier at all times we find:

$$E_{\text{out}} = E_0 \left( J_0(x)\chi_0 e^{i\omega_c t} + J_1(x)\chi_1 e^{i(\omega_c + \Omega)t} - J_1(x)\chi_1 e^{-i(\omega_c - \Omega)t} \right),$$

where $\chi_0$ is a real quantity taking into account attenuation of the carrier and $\chi_1 = \chi_{-1}$ is the complex phase picked up by the two sidebands due to the atoms. The NICE-OHMS signal, as seen from Figure 1, is the demodulated signal at $\Omega$ (phase of cavity transmitted field) or $2\Omega$ (attenuation of cavity transmitted field). In our case this becomes:

$$S_{1\Omega} \propto J_0(x) J_1(x) \chi_0 \cdot \text{Im}(\chi_1),$$

The signal described in Eq. (4) is proportional to the phase dispersion induced by the atoms for small laser detunings.

By demodulating the transmitted field at $2\Omega = 2 \cdot \text{FSR}$, one obtains:

$$S_{2\Omega} \propto J_0(x) J_2(x) \chi_0 \cdot \text{Re}(\chi_1) - J_1(x)^2 |\chi_1|^2. \quad (5)$$

The complete information on the complex transmission coefficient, $\chi$, can be found by combining the quantities obtained from demodulating at $\Omega$ and $2\Omega$ via Equation (4) and (5).

The theoretical model [14], [19] for this system provides the mean value of the annihilation operator of the cavity mode $\alpha = -i \langle \hat{a} \rangle$, which in the classical limit corresponds to the measured quantity $\chi$ as follows: $\frac{\kappa}{\eta} \alpha = \chi$. Here $\kappa$ is the cavity decay rate and $\eta$ is the classical drive amplitude of the cavity field.

III. Results

The NICE-OHMS signals for demodulation at $\Omega$ and $2\Omega$ are presented in Figure 2. The signal demodulated with the frequency $\Omega$, shown in Figure 2(a), corresponds to the phase dispersion induced by the atoms for small phase shifts. This signal is the signal expressed in Eq. (4) and deviates slightly from the pure phase shift for larger detunings since small-angle approximations are no longer valid. The slope and the signal-to-noise ratio of this measured phase dispersion at resonance determine the ultimate frequency stability of the system.

The Doppler energy scale in the investigated system is several orders of magnitude larger than the narrow linewidth
The inset in Figure 2(a) shows experimental data of a frequency scan of the cavity transmitted phase shift close to resonance. Two theoretical plots are also shown. The blue solid line takes the full theory (all doppleron orders) into account while the black dashed line is based on a theoretical prediction where the velocity dependent doppleron resonances are not taken into account. We observe good agreement with the theoretical model including all doppleron orders and notice the doppleron resonances tend to decrease the phase dispersion slope around resonance slightly.

The signal demodulated with the frequency 2Ω, shown in Figure 2(b), shows the attenuation of the cavity transmitted field. This signal corresponds to Equation (5) from which positive values are expected. However a background signal corresponding to the empty cavity signal is subtracted from the experimental data, and this results in negative values. The absolute value of this signal after background subtraction represents the degree of field attenuation. The theoretical empty cavity value is also subtracted from the theoretical prediction in Figure 2(b). The central region of Figure 2(b) shows reduced attenuation due to saturation, modified by velocity dependent processes. The inset in Figure 2(b) shows the detailed structures of this modified saturation peak. This modification is not predicted by the $T=0$K theory model [10].

The shot-noise-limited linewidth of the system, $\Delta\nu$, can be estimated by assuming a perfect locking scheme and detectors with unity quantum efficiencies as follows [10]:

$$\Delta\nu = \frac{\pi \hbar \nu}{2 P_{\text{sig}} \left( \frac{d\phi}{d\nu} \right)^2 \left( 1 + \frac{P_{\text{sig}}}{2P_{\text{sideband}}} \right)},$$

(7)

where $P_{\text{sig}}$ is the input carrier power, $P_{\text{sideband}}$ is the sideband power and $\frac{d\phi}{d\nu}$ is the dispersion slope at resonance. Whereas multi-photon scatterings are found to degrade the ultimate linewidth achievable, the shot-noise-limited linewidth of the current non-optimized system is estimated from Equation (7) to be 500 mHz.

The same system with optimized parameters is estimated to have a shot-noise-limited linewidth comparable with state-of-the-art laser stabilization [15], [16], [23]. Furthermore, the theoretical model shows that the gain, in terms of frequency stability, by further cooling of the atomic sample is minimal. This opens the prospects for implementations of a simple and compact transportable optical atomic clock with high phase stability.

The phase dispersion slope depends non-linearly on the optical intra-cavity power, which makes the optimal choice of parameters non-trivial. The non-linear input power dependence of the phase dispersion slope is measured and shown in Figure 3(a). The data and theory are in great agreement and of-the-art laser stabilization [15], [16], [23]. Furthermore, the theoretical model shows that the gain, in terms of frequency stability, by further cooling of the atomic sample is minimal. This opens the prospects for implementations of a simple and compact transportable optical atomic clock with high phase stability.

The inset in Figure 2(a) shows experimental data of a frequency scan of the cavity transmitted phase shift close to resonance. Two theoretical plots are also shown. The blue solid line takes the full theory (all doppleron orders) into account while the black dashed line is based on a theoretical prediction where the velocity dependent doppleron resonances are not taken into account. We observe good agreement with the theoretical model including all doppleron orders and notice the doppleron resonances tend to decrease the phase dispersion slope around resonance slightly.

The signal demodulated with the frequency 2Ω, shown in Figure 2(b), shows the attenuation of the cavity transmitted field. This signal corresponds to Equation (5) from which positive values are expected. However a background signal corresponding to the empty cavity signal is subtracted from the experimental data, and this results in negative values. The absolute value of this signal after background subtraction represents the degree of field attenuation. The theoretical empty cavity value is also subtracted from the theoretical prediction in Figure 2(b). The central region of Figure 2(b) shows reduced attenuation due to saturation, modified by velocity dependent processes. The inset in Figure 2(b) shows the detailed structures of this modified saturation peak. This modification is not predicted by the $T=0$K theory model [10].

The shot-noise-limited linewidth of the system, $\Delta\nu$, can be estimated by assuming a perfect locking scheme and detectors with unity quantum efficiencies as follows [10]:

$$\Delta\nu = \frac{\pi \hbar \nu}{2 P_{\text{sig}} \left( \frac{d\phi}{d\nu} \right)^2 \left( 1 + \frac{P_{\text{sig}}}{2P_{\text{sideband}}} \right)},$$

(7)

where $P_{\text{sig}}$ is the input carrier power, $P_{\text{sideband}}$ is the sideband power and $\frac{d\phi}{d\nu}$ is the dispersion slope at resonance. Whereas multi-photon scatterings are found to degrade the ultimate linewidth achievable, the shot-noise-limited linewidth of the current non-optimized system is estimated from Equation (7) to be 500 mHz.

The same system with optimized parameters is estimated to have a shot-noise-limited linewidth comparable with state-of-the-art laser stabilization [15], [16], [23]. Furthermore, the theoretical model shows that the gain, in terms of frequency stability, by further cooling of the atomic sample is minimal. This opens the prospects for implementations of a simple and compact transportable optical atomic clock with high phase stability.

The phase dispersion slope depends non-linearly on the optical intra-cavity power, which makes the optimal choice of parameters non-trivial. The non-linear input power dependence of the phase dispersion slope is measured and shown in Figure 3(a). The data and theory are in great agreement and the non-linear tendency is evident. The theory curve predicts a maximum absolute value for the slope for lower input powers than the measured values. However, the shot-noise-limited linewidth does not depend solely on the phase dispersion slope.
It is also governed by the input power in accordance with Eq. (7). The shot-noise-limited linewidth for each measured parameters are calculated and shown in Figure 3(b). The data sets in Figure 3(a) and (b) are identical dataset. Figure 3(b) shows, that a system with the optimal input power with the minimum shot-noise-limited linewidth has been realized and measured. Note, that the horizontal axis is logarithmic, allowing a change in input power of one order of magnitude without degrading the shot-noise-limited linewidth significantly. This indicates, that it is not necessary to probe with technically challenging low input powers down to $<100 \text{nW}$, and that the system is robust to changes in the input power.

![Graph](image_url)

**Fig. 3.** The input power dependence of the slope of the phase dispersion around resonance and the corresponding shot-noise-limited linewidth for a total number of atoms of $N = 4.0 \times 10^9$ and atom temperature of $T = 5.0 \text{ mK}$. (a): The slope of the phase dispersion around resonance measured for different input powers. The solid line is the theoretical prediction while the dots are experimental data. Every experimental value of the slope is evaluated by fitting a theory curve to a scan of the phase dispersion around resonance for a given input power. The data and theory are in great accordance and the non-linear tendency is evident. Notice that the power-axis is logarithmic. (b): The shot-noise-limited linewidth for different input powers corresponding to the measured parameters in (a). The solid lines are theoretical predictions and the dots are experimental values. The different colors corresponds to different ratios of carrier and sideband input powers. This has no influence on the plotted data in (a).

**IV. PROSPECTS**

The understanding of relevant motional effects investigated here has direct implications for atomic clocks and superradiant laser sources involving ultranarrow transitions.

In order to evaluate the performance of the laser stabilization method presented here, we have estimated the shot noise limited linewidth using equation (7) for stabilization on the $^1S_0 \rightarrow ^3P_1$ intercombination transition for a number of different two electron systems. It is assumed for each element that the atomic sample is cooled close to the Doppler limit with reference to temperatures obtained experimentally [24]–[26], see Table I.

A realistic atom number overlapping the cavity mode has been estimated to $N_{\text{cav}} = 2.5 \times 10^7$ and has been used for all elements. For several elements larger atom numbers have been reported in literature which makes our estimates somewhat conservative. Cavity dimensions are based on the experimental values obtained in our setup [14]. The cavity waist diameter is chosen to 1.0 mm and the empty cavity finesse to $F = 250$, corresponding to an empty cavity decay rate of $\kappa = 2\pi \cdot 2.0 \text{ MHz}$. In each case the laser decoherence has been assumed to be $\gamma_{\text{laser}} = 2\pi \cdot 2.0 \text{ kHz}$. The input power corresponding to the minimum obtainable shot noise limited linewidth, and the corresponding shot noise limited linewidth has been found for each element, see Table I. Whereas stabilizing a probe laser on the $^1S_0 \rightarrow ^3P_1$ transition of $^{25}$Mg may be slightly more experimentally challenging compared to other two electron systems, the stabilization of a probe laser on the $^1S_0 \rightarrow ^3P_1$ transition of $^{88}$Sr with $F = 250$ and $N_{\text{cav}} = 2.5 \times 10^7$ promises shot noise limited linewidth of 152 mHz. This linewidth is already comparable with the linewidths of current state-of-the-art frequency stabilized lasers [14], [16].

For all elements the shot noise limited linewidth may be reduced further by increasing the value of the collective cooperativity $C = NC_0$. This can be achieved in two ways: by increasing the atom number or the cavity finesse. For strontium-88 the laser linewidth may be reduced to $\Delta \nu = 9.5 \text{ mHz}$ by increasing the cavity finesse to $F = 1000$ and increasing the atom number by a factor of two to $N_{\text{cav}} = 5.0 \times 10^7$. See Table I.

Improvement of the cavity finesse and implementation of the techniques developed in this work into beam-line experiments, without meeting the severe technical requirements as described in [20], seems realizable. This also opens for a continuous interrogation of narrow optical transition.

**V. CONCLUSION**

We have experimentally investigated the velocity dependent processes and dynamics of a cavity QED system consisting of laser-cooled strontium-88 atoms coupled to a low finesse optical cavity.

The NICE-OHMS technique has been applied to perform FM spectroscopy and a complete connection is established between the measured quantities and the complex transmission coefficient of the input field. Both the cavity transmitted phase shift and the attenuation of the cavity transmitted field are measured and both quantities show significant modifications
TABLE I. SHOT NOISE LIMITED LINESWIDTHS $\Delta \nu$ ESTIMATED THEORETICALLY FOR A NUMBER OF $^{171}$Yb $\rightarrow ^{3}P_{1}$ INTERCOMBINATION LINES AT EXPERIMENTALLY REALIZABLE PARAMETERS: TRANSITION WAVELENGTH $\lambda$, NATURAL LINESWIDTH $\gamma$, PROBE LASER DECOHERENCE $\gamma_{\text{Laser}}$, ATOMIC TEMPERATURE $T$ AS REPORTED IN LITERATURE [24]–[26] AND OPTIMAL INPUT POWER $P_{\text{In, optimal}}$. THE RATIO OF THE CARRIER AND SIDE BAND POWER IS CHOSEN TO $\frac{P_{\text{Side Band}}}{P_{\text{Car}}}$ = 1 IN ALL CASES. THROUGHOUT THE CALCULATION WE HAVE ASSUMED A PRESTABILIZED PROBE LASER DECOHERENCE OF $\gamma_{\text{Laser}}/2\pi = 2.0$ kHz TO SPECIFY AN ORDER OF MAGNITUDE. THE SHOT NOISE LIMITED LINESWIDTHS ARE CALCULATED FOR SYSTEMS WITH EMPTY CAVITY FINESSE OF $F = 250$ AND $N_{\text{CW}} = 2.5 \times 10^{7}$ ATOMS OVERLAPPING THE CAVITY MODE, AND A IMPROVED CAVITY WITH $F = 1000$ AND $N_{\text{CW}} = 5.0 \times 10^{7}$.

<table>
<thead>
<tr>
<th>Atom</th>
<th>$\lambda$ (nm)</th>
<th>$\gamma/2\pi$ (kHz)</th>
<th>$F$</th>
<th>$T$ (mK)</th>
<th>$P_{\text{In, optimal}}$ (mW)</th>
<th>$\Delta \nu$ (mHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{171}$Yb</td>
<td>556</td>
<td>182 kHz</td>
<td>250</td>
<td>6.5</td>
<td>128</td>
<td>128 $\mu$W</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td></td>
<td>1000</td>
<td>318 $\mu$W</td>
<td>210 mHz</td>
<td></td>
</tr>
<tr>
<td>$^{40}$Ca</td>
<td>657</td>
<td>400 Hz</td>
<td>250</td>
<td>1.7</td>
<td>3.7</td>
<td>448 mHz</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td>1000</td>
<td>17.4</td>
<td>0.6</td>
<td>0.6 W</td>
<td>28 mHz</td>
</tr>
<tr>
<td>$^{24}$Mg</td>
<td>457</td>
<td>34 Hz</td>
<td>250</td>
<td>3.0</td>
<td>0.4</td>
<td>4.0 Hz</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>1000</td>
<td>3.0</td>
<td>0.2</td>
<td>0.2 mW</td>
<td>585 mHz</td>
</tr>
<tr>
<td>$^{88}$Sr</td>
<td>689</td>
<td>7.6 kHz</td>
<td>250</td>
<td>3.0</td>
<td>57</td>
<td>350 W</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td></td>
<td>1000</td>
<td>69</td>
<td>9.5 mHz</td>
<td></td>
</tr>
</tbody>
</table>

due to velocity dependent processes in accordance with the theoretical model.

The ideal shot-noise-limited linewidth of a laser stabilized to our system depends on the phase dispersion slope around resonance. The non-linear input power dependence of the phase dispersion slope is measured and the input power for minimum shot-noise-limited linewidth is determined. Parameters corresponding to a shot-noise-limited linewidth down to 500 mHz are measured.

The ideal shot-noise-limited linewidth of similar cavity QED systems with different two electron transitions in other elements are calculated. It is evaluated based on the state-of-the-art theoretical model of the non-linear dynamics of such systems, that a shot-noise-limited linewidth below 10 mHz is achievable for a strontium-88 based system with realistic experimental parameters.

The understanding of relevant velocity dependent effects presented here has direct relevance for atomic clocks and superradiant laser sources involving ultra-narrow transitions. Specifically, the understanding of the dynamics of cavity QED systems with single-stage MOT temperatures obtained for different two electron systems in this work will prove valuable for future transportable stable lasers and atomic clocks employing warm atoms for out-of-lab operation under more noisy environments.

ACKNOWLEDGMENT

We would like to acknowledge support from the Danish Research Council and ESA Contract No. 4000108303/13/NL/PA-NP1272-2012. J. Y. also wish to thank the DARPA QuaSAR program, NIST, and the NSF Physics Frontier Center at JILA for financial support.

REFERENCES


Ten Years of Active Optical Frequency Standards

Duo Pan, Wei Zhuang*, Xiaobo Xue, Xiaogang Zhang, Mo Chen, Zhichao Xu, and Jingbiao Chen
State Key Laboratory of Advanced Optical Communication System and Network, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China.
*National Institute of Metrology, Beijing10013, China.
E-mail: jbchen@pku.edu.cn

Abstract—The concept of active optical clock was proposed ten years ago. In this paper, after a simple review, we will mainly present the most recent experimental progresses of active optical frequency standards in Peking University, including 4-level Cesium active optical frequency standards and active Faraday optical frequency standards.

Keywords—Active optical clocks; bad-cavity laser; 4-level active optical frequency standards; active Faraday optical frequency standards.

I. INTRODUCTION

To reduce the clock linewidth to millihertz level [1] experimentally is still a challenge due to the cavity inevitable thermal Brownian-motion noise. Active optical clock[2-3], utilizing optical stimulated emission on ultra-narrow atomic transition line in bad-cavity regime with coherent weak optical feedback to maintain collective phase information from different atoms, could greatly reduce the influence of the mechanical or thermal vibrations of the cavity mirrors on the emitted optical frequency. Since the concept of active optical clock has been conceived [2-3], several experimental schemes in different configurations have been proposed and demonstrated.

Neutral atoms at 2-level [3-9], 3-level [2,10-12], 4-level [4-6,13,14] energy structures with thermal and laser cooled and trapped configurations [5-9], Raman laser [15-16], and sequential coupling configuration [17], and moving optical lattice [18] have been investigated recently. Different approaches to the creation of a high-stability active optical frequency standard with cold atoms have been discussed carefully in the most recent paper [18]. For laser cooled and trapped atoms, to make continuous-wave active optical clock not at 4-level configuration, one has to apply sequential coupling technique [17], or synchronization between two ensembles of atoms [19-20].

The theoretical quantum limited linewidth of the active optical clock [3,7,10] is narrower than mHz, and to reach this unprecedented linewidth is possible since the effect of thermal noise on cavity mode can be suppressed dramatically with the cavity-pulling suppression mechanism of active optical clock [2,3]. The Raman laser in bad-cavity regime has been intensively investigated with cooled Rb atoms recently [15,16,21-23] with very beautiful results [15]. However, the light shift due to pumping laser may be a limitation of high performance for active optical clocks with laser cooled and trapped atoms in 3-level configuration [2,18]. The effect of interaction between atoms on frequency stability and accuracy of active optical clock has been calculated theoretically [24].

Four-level active optical frequency standards have been proposed to avoid the sensitivity of light shift due to pumping laser [13,14,25-27]. Recently, the lasing action, cavity pulling suppression, and the output linewidth of Cs four-level active optical frequency standards have been investigated experimentally [28-31]. For alkali atoms like Cs, Rb and K, the available diode lasers and mature techniques of laser cooling and trapping will be very helpful in the research of active optical frequency standards based on alkali atoms.

However, in all above mentioned approaches [2-31] to the creation of active optical clock, the quantum reference of frequency standard and the stimulated emission of gain medium are from the same atomic transition of the same atoms. The requirement strongly limits available quantum system. The active Faraday optical frequency standard, as demonstrated by a recent experiment [32], spatially separates the quantum reference of frequency standard and the stimulated emission of gain medium. In the active Faraday optical frequency standard [32], a narrow linewidth Faraday atomic filter is used as quantum reference of frequency standard [32-35], while the stimulated emission of gain medium can be provided by Ti: sapphire and dye, besides semiconductor diode materials[32]. In this way, the Faraday effect [35] found 170 years ago, starts to play a quantum reference role in modern optical clocks.

In this Review, we mainly focus on the four-level Cesium active optical frequency standards in the first part. It was shown that when Cesium atoms within a FP cavity with a finesse of 4.3, which is at deep bad-cavity regime, pumped by 455.5 nm and 459.3 nm laser, the stimulated emission of Cs four-level active optical frequency standards is realized. The measured threshold, linewidth, cavity-pulling reduction, and saturation
effect of pumping laser are presented. The second part emphasizes on the active Faraday optical frequency standard, which has advantages of separated the quantum reference of frequency standard and the stimulated emission of gain medium, besides narrowed linewidth and reduced cavity pulling effect. Measured by the optical heterodyne beat between two similar independent active Faraday optical frequency standards, results show the frequency linewidth reaches 281(23) Hz, which is 19000 times smaller than the natural linewidth of the Cesium 852 nm transition line\[32\]. Besides, we also implemented an active Faraday optical frequency standard in good-cavity regime named as ultra-long Faraday laser \[34\], based on the Faraday anomalous dispersion atomic filter with ultra-narrow bandwidth of 26 MHz at Rubidium 780 nm transition line and the ultralong fiber extended cavity of 800 m with extremely small free-spectrum-range of 0.125 MHz. The stability around $10^{-10}$ level has been measured at sampling time from 0.01 s to 1 s. These progresses have indicated good performances of active optical frequency standards and may promote further research for better atomic optical clocks.

II. FOUR-LEVEL ACTIVE OPTICAL FREQUENCY STANDARDS

The scheme uses four atomic energy levels for optical pumping and lasing, which could reduce influences of pumping laser on lasing levels and emitted light frequency. Several atomic species have been proposed \[13,14,25,26\] and stimulated emission of cesium four-level active optical frequency standards have been realized \[28-31\]. The relevant cesium atomic energy levels are shown in Fig. 1, in which atoms are excited to $7^2S_{1/2}$ from the ground state $6^2P_{3/2}$ by 459 nm pumping laser. The 1469.9 nm output light is stimulated emission of radiation built up between $7^2S_{1/2}$ and $6^2P_{3/2}$ with weak optical feedback from cavity.

The experimental setup is shown schematically in Fig. 2. The cesium vapor cell was placed in the cavity composed of a concave mirror and a plane mirror. The former cavity mirror was coated with 459 nm and 1469.9 nm high-reflection coating and the latter with 459 nm high-reflection and the reflectivity of 1469.9 nm is 80%. Therefore, the cavity bandwidth was calculated to be 383.3 MHz for 1469.9 nm while gain bandwidth 7.38MHz. The ratio between cavity bandwidth and gain bandwidth is 51.94, which means the whole system worked in bad-cavity regime. The 1529 nm laser was used to calibrate the cavity.

When the 459 nm pumping laser was applied, the population inversion \[31\] and lasing between $7^2S_{1/2}$ and $6^2P_{3/2}$ can be realized. The 1469.9 nm output light power when changing 459 nm pumping laser power clearly shows a threshold of 0.8 mW represented by the 459 nm pumping laser power. The lasing of the stimulated emission can only be realized above the threshold.

The main characteristics of the cesium four-level active optical frequency standard were studied. First of all, the suppression of cavity pulling effect was examined using a Fabry-Perot interferometer (FPI) with a bandwidth of 2.7 MHz. We changed the mode frequency of the cavity by applying voltage on the piezoelectric element adhered to the cavity mirror. The results showed that the 1469.9 nm output light frequency was changed 6.8 MHz when cavity mode frequency was tuned 281.6 MHz away. Therefore, the cavity pulling effect was reduced to 2.4% compared with ordinary laser. The advantage of suppressed cavity pulling can reduce influences of vibrational noises on active optical frequency standards. In this way, the cavity inevitable thermal Brownian-motion noise...
frequency was scanned referenced to cesium $^6S_{1/2}$ and $^7P_{3/2}$ saturated absorption spectrum shown as the upper lines. The output 1469.9 nm light power was recorded and normalized as the lower lines in Fig. 4 (a) and Fig. 4 (b). The result indicated that the optimized at $^6S_{1/2}$ (F=4) and $^7P_{3/2}$ (F′=5) transition of pumping laser. The hyperfine state of $^7S_{1/2}$ (F′=4) was proved to be upper level of 1469.9 nm Cs active optical frequency standard. The gain bandwidth of the 1469.9 nm transition can be deduced from the velocity spread of Cesium atoms at $^7S_{1/2}$ state, which is eventually determined by the saturation effect due to the 455 nm pumping laser power. Therefore, the standing wave of pumping laser power should be reduced to decrease the saturation broadening of the gain linewidth at 1469.9 nm.

III.  ACTIVE FARADAY OPTICAL FREQUENCY STANDARDS

A. Active Faraday optical frequency standards in bad-cavity regime

Active Faraday optical frequency standards [32] utilize Faraday atomic filters [33-42] as frequency references when working in bad-cavity regime. The output light frequency is determined by alkali atomic transition line of Faraday atomic filter and the output stimulated emission light power can be increased by adopting Ti: sapphire, dye and semiconductor materials as gain medium since the quantum reference of frequency standard and the stimulated emission of gain medium are spatially separated. This unique approach of active optical clock opens the door for various atomic transitions and alternative gain media to optical clocks.

The Faraday atomic filters with narrow bandwidth have been studied in our group since 2011 [36-42]. The experimental setup of ultranarrow bandwidth Faraday atomic filter based on alkali atoms is shown in Fig. 5. It is realized by velocity-selective optical pumping of alkali vapor in a bias magnetic field. The Faraday atomic filter is composed of two polarization-orthogonal Glan-Taylor prisms, an alkali vapor cell, a pumping diode laser with circular polarization and a bias magnetic field. The pumping laser is frequency stabilized to alkali atomic saturation absorption spectroscopy. The bandwidth of the Faraday atomic filter could reach 25 MHz with a transmission of 18% [42], which is adequately applied in active Faraday optical frequency standards.

Fig. 5. (Color online). Experimental setup of active Faraday optical clock using ultranarrow bandwidth Faraday atomic filter [42].

When a Faraday atomic filter was placed inside extended cavity of laser diode with anti-reflection coating, stimulated emission of active Faraday optical frequency standard can be realized by adjusting the extended cavity mirror. Fig. 6 shows the output light power varying with injection current of laser diode. When increasing the injection current, the refractive index of semiconductor is changed. The whole cavity length, which is the sum of the extended cavity length and the optical length of semiconductor, is changed accordingly. Therefore, there are multiple thresholds of the injection current. The maximum output light power can reach 75 μW.

Fig. 6. The varization of output light power while increasing the injection current of laser diode [32].

We have also checked cavity pulling effect in active Faraday optical frequency standard by beat the output light with pumping laser. It was found that when changing the
extended cavity frequency within 350 MHz, the center frequency of the beat signal varied within a range of 40 MHz. Therefore, the cavity pulling coefficient is 11% which is reduced compared with common laser. That means the current system works in bad-cavity regime.

The output light frequency linewidth was measured by the heterodyne signal between two independent identical setups. The beat signal is shown in Fig. 7. The swept frequency range is 20 kHz within 200 ms and RBW 100 Hz. The data were fitted by a Gaussian profile, which indicates FWHM of the beat signal is 398 Hz. The linewidth for each setup can be deduced by the measured FWHM divided by √2, which is 281 Hz. The result is 19000 times smaller than the natural linewidth of the Cesium 852 nm transition line and two orders better than frequency-stabilized laser with atomic filter working in good-cavity regime or interference filter. It indicates that active Faraday optical frequency standard working in bad-cavity regime can reduce influence of vibrations on frequency stability.

![Fig. 7. The output light frequency linewidth of active Faraday optical frequency standard [32].](image)

As for the disadvantageous influence from thermal motion of atoms, laser cooled and trapped atom in magic wavelength optical lattice is a promising solution for high-performance active Faraday optical clock.

We can use the 689 nm transition of strontium atoms trapped by magic wavelength optical lattice, which has a natural linewidth of 7.6 kHz. With the existence of homogeneous magnetic field with appropriate intensity, we can realize ultra-narrow linewidth Faraday optical filter at 689 nm transition. If we use semiconductor diode or Ti:Sapphire as gain medium, by weak feedback of laser resonator at bad-cavity regime, we can realize Faraday active optical clock based on cold atoms. This scheme is shown in Fig. 8.

This research is on-going, and cold atom Faraday optical clock is expected to provide much higher performance.

**B. Ultra-long Faraday laser**

Besides of the configuration of active Faraday optical frequency standard at bad cavity regime, we developed another configuration of optical frequency standard based on Faraday optical filter, named *ultra-long Faraday laser*. This configuration combines ultranarrow FADOF and ultralong fiber cavity, see Fig. 9.

The ultranarrow FADOF (the black box in figure 9) with bandwidth of around 20 MHz will limit the lasing frequency from ARLD. The ultralong cavity provides ultra-small free spectral range of around 100 kHz or less. After mode competition, only one mode at the center of FADOF pass-band survives, which forms the active lasing frequency.

Since the FSR is determined by $FSR = \frac{c}{2nL}$, where $n$ is the refractive index of cavity medium, $L$ is cavity length. For optical fiber cavity, the cavity length can be immensely increased without loss of cavity stability. If we extend cavity length to 100 km, the FSR is about 1 kHz. Since the output frequency drift is expected to be no more than one FSR, we can expect remarkable performance of the ultralong Faraday laser.

The lasing output signal is ‘locked’ to FADOF pass-band center, which corresponds to a certain atomic transition line. The stabilized laser output when using 150 m long fiber cavity signal is shown in Fig. 10.

![Fig. 9. Schematics of ultralong Faraday laser](image)
optical frequency standards. We believe that this configuration provides a new path to compact traditional optical frequency standards' configurations. We have shown that the Allan deviation of fractional frequency with 800 m fiber can reach around $6 \times 10^{-10}$ as plotted in Fig. 11.

Since this output frequency is determined by FADOF pass-band, which is corresponding to quantum transition of atoms. We can treat the ultralong Faraday laser as an optical frequency standard. The Allan deviation is calculated, preliminary results shown that the Allan deviation of fractional frequency with 800 m fiber can reach around $6 \times 10^{-10}$ as plotted in Fig. 11.

The ultralong Faraday laser is simple compared with traditional optical frequency standards’ configurations. We believe that this configuration provides a new path to compact optical frequency standards.

IV. CONCLUSION

We proposed the concept of active optical clock ten years ago. In this paper, after a simple review, we have mainly presented the most recent experimental progresses of active optical frequency standards in Peking University, including 4-level Cesium active optical frequency standards and active Faraday optical frequency standards.

This work is supported by NSF of China under No. 91436210, and International Science & Technology Cooperation Program of China under No. 2010DFR10900.

REFERENCES


[27] Y. Wang, D. Wang, T. Zhang et al., “Realization of population inversion between 7S_{1/2} and 6P_{1/2} levels of cesium for four-level active optical clock,” Science China, Physics, Mechanics and Astronomy, 2013, 56, 1107-1110.


Precise Cascade Synchronization of Two Digitally Tuned Space Clocks to UTC (GPS)

H. Wang and G. H. Iyanu
Photonics Technology Department
The Aerospace Corporation
El Segundo, CA 90245, USA
he.wang@aero.org

Abstract—We report the results of a three-clock, cascade synchronization experiment, in which a crystal oscillator clock is slaved to a Rb atomic clock while the Rb clock is independently slaved to a Cs reference clock. The Cs clock in the experiment is calibrated against a GPS receiver. With such a cascade scheme and a real-time linear regression algorithm, we are able to precisely synchronize the digitally-tuned space clock with the Cs clock within 50 ns with an averaged residual frequency offset of 5 × 10⁻¹³. This cascade synchronization technique has also significantly reduced the two space clocks’ long-term frequency drift to 1.7 × 10⁻¹⁴ per day and 2.5 × 10⁻¹³ per day for the Rb clock and the crystal oscillator clock, respectively.

Keywords—precision timekeeping; cascade synchronization; atomic clocks, crystal oscillator clocks.

I. INTRODUCTION

A satellite constellation usually has multiple time and frequency reference clocks on board. Some satellites may carry atomic clocks and others have crystal oscillator clocks on board. For applications such as satellite communication networks, all on-orbit clocks in the constellation are required to be precisely synchronized to UTC through the reference clocks in the ground control system. In order to efficiently synchronize all on-orbit clocks, the satellite timekeeping system could be designed such that the ground control only manages one on-orbit clock, usually an atomic clock as the Master clock, and other on-orbit clocks are slaved to the Master clock through the constellation crosslink. Such a cascade timekeeping architecture is particularly useful if the satellite system has crystal oscillator clocks on board. In general, a crystal oscillator clock has poorer long term frequency stability than an atomic clock [1] and is more sensitive to space environmental effects [2]. Therefore it is important to know how well a crystal oscillator clock can be slaved to a Master atomic clock through satellite crosslink and how well the Master clock can be synchronized to the ground reference clock through the ground to satellite link. In order to answer those questions, we have constructed a unique hardware-software spacecraft atomic clock flight simulation and test station at The Aerospace Corporation [3], which is capable of simulating and studying the algorithms and synchronization performance of a multiple-clock timekeeping system.

II. EXPERIMENT

In order to help satellite operators and system engineers understand and resolve flight clock issues, we built a spacecraft atomic clock flight simulation and test facility as previously described in [3, 4]. Fig. 1 shows the block diagram of the simulation and test station.

Briefly, the test station has two subsystems: a simulated space segment and a simulated ground control segment of a satellite system. The space segment operates two digitally-tuned rubidium master oscillators (RMOs) housed in a vacuum chamber (RMO-15 and RMO-3). A personal computer (PC1) simulates the on-board computer of the satellites. A second computer (PC2) and a Cs clock

Fig. 1. The flight clock simulation and test station. RMO stands for Rubidium Master Oscillator and XMO for Crystal Master Oscillator. SPI refers to Serial Peripheral Interface that generates the digital clock tuning words. PC1 simulates the on-board computer and PC2 performs the timekeeping function of the ground control system. The Cs reference clock is calibrated against a GPS receiver (not shown in the diagram).
III. ALGORITHM

A clock’s accumulated time offset is the integral of the fractional frequency as given in (1). For a linear frequency drift, the clock’s time error is a quadratic function of time.

\[ \Delta t(\tau) = \int_0^\tau f(t) \, dt = f_0 \tau + \frac{1}{2} D \tau^2 \pm \sqrt{2} \tau \sigma_y(\tau) \]  

In (1), \( f_0 \) is an initial frequency offset and \( D \) is the clock’s frequency drift rate. \( \sigma_y \) is the Allan deviation of the clock’s frequency, which represents the clock’s frequency noise contribution to the accumulated time offset. On the other hand, a clock’s instantaneous fractional frequency offset \( \Delta f(t) \) is the first derivative of the time offset function as in (2). In real satellite timekeeping operation, the relative frequency offset of two on-orbit clocks is estimated by measuring the signal arrival time differences (or time errors) between the two clocks. The slope of the measured time offset as a function of time can also be computed using a simple linear regression (or linear least-squares fit) [10] as given in (3). This computed slope gives the estimated frequency offset between the two clocks. Specifically for this work, this is the “space” clock’s frequency offset \( \Delta f'_{corr} \) to be corrected.

\[ \Delta f(t) = \frac{d\Delta t(\tau)}{dt} \]  

\[ \Delta f'_{corr} = \frac{\sum_{m=N}^{m=N} \left[ t_m \times \Delta t_m \right] - \frac{1}{N} \left[ \sum_{m=N}^{m=N} t_m \times \sum_{m=N}^{m=N} \Delta t_m \right]}{\sum_{m=N}^{m=N} t_m^2 - \frac{1}{N} \left[ \sum_{m=N}^{m=N} t_m \right]^2} \]  

In (3), \( \Delta t_m \) is the measured time offset at time \( t_m \). \( N \) is the number of data points included in the linear regression computation. The detailed flow diagram of the algorithm used in this experiment has been discussed in [4]. The same algorithm is utilized to slave the crystal oscillator clock XMO-15 to the Rb atomic clock RMO-3 and to synchronize RMO-3 to the Cs reference clock as well. The relative time offsets between XMO-15 and RMO-3 as well as between RMO-3 and the Cs clock are derived from the real-time phase measurement of the clocks. A recursive exponential smoothing filter [6] as given in (4) is applied to eliminate the fast changing noise in the time offset data.

\[ \Delta t'_k = \lambda \Delta t_k + (1 - \lambda) \Delta t'_{k-1} \]  

where \( \lambda = 2^{-n}, n \) is an integer. In this experiment, \( n = 9 \) is used.

With the real-time linear regression algorithm, a clock update is initiated when the relative time offset exceeds a preset threshold \( \Delta t_{th} \). Immediately a linear regression calculation is started using the last \( N \) data points to determine the slope of the time offset curve. The computed slope, which is the estimated frequency error of the XMO-15 or RMO-3 just before a new clock update, is converted to the clock’s frequency tuning word command to be sent back to the “space” clocks to update their frequency. Upon completion of the frequency correction, the clock’s time is reset and a new autonomous time period begins. This real-time linear regression timekeeping algorithm is implemented with a C++ programming code.

IV. RESULTS

Using the real-time linear regression algorithm described above, we performed a three clock, Slave-Master-Cs timekeeping experiment, in which the crystal oscillator clock XMO-15 is slaved to the atomic clock RMO-3 while RMO-3 is independently synchronized to the Cs reference clock. As shown in Fig. 1, in a very real sense this is an “autonomy” experiment with the two independent clock synchronization loops closed. The reported results below are derived from experiments that continuously ran for 11 days without interruptions.

The real-time linear regression algorithm has kept both “space” clocks precisely synchronized to the “ground” Cs reference clock in a cascaded way. Fig. 2 shows the measured output frequency of RMO-3 relative to the Cs reference clock. Fig. 3 is the recorded output frequency of XMO-15 also relative to the Cs clock. The time offset thresholds used in this experiment are \( |\Delta t_{th}| = 50 \text{ ns} \) for RMO-3 and 100 ns for XMO-15. The data sampling rate is 1 Hz.

To determine the effective long-term frequency drifts of those two “space” clocks after being cascade synchronized to the Cs clock, we performed linear least-squares fits to the measured frequency data and the results are given in Table I. In Table I, we also list the frequency drift rates when those two clocks operate in free-running mode (unslaved). It is clearly demonstrated that the clock’s frequency drift rates are significantly reduced by 35 times and 44 times for RMO-3 and XMO-15, respectively. Table II gives the measured effective
Fig. 2. Measured output frequency as a function of time for RMO-3, which is synchronized to the Cs reference clock. The red line is a linear least-squares fit of the frequency data. The time offset threshold used to slave RMO-3 to the Cs clock is $|\Delta t_{\text{th}}| = 50$ ns. The data sampling rate is 1 Hz.

Fig. 3. Measured output frequency as a function of time for XMO-15, which is slaved to the Rb atomic clock RMO-3. The blue line is a linear least-squares fit of the frequency data. The time offset threshold used to slave XMO-15 to RMO-3 is $|\Delta t_{\text{th}}| = 100$ ns. The data sampling rate is 1 Hz.

<table>
<thead>
<tr>
<th>Frequency Drift (per day) relative to Cs clock</th>
<th>RMO-3 (Rb atomic clock)</th>
<th>XMO-15 (Crystal oscillator clock)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free running [4]</td>
<td>$6 \times 10^{-13}$</td>
<td>$1.1 \times 10^{-11}$</td>
</tr>
<tr>
<td>Slave-Master-Cs scheme (This work)</td>
<td>$1.7 \times 10^{-14}$</td>
<td>$-2.5 \times 10^{-13}$</td>
</tr>
</tbody>
</table>

Fig. 4. Recorded time offset of RMO-3 as a function of time relative to the Cs clock with a time offset threshold $|\Delta t_{\text{th}}| = 50$ ns.
synchronization experiment on RMO-3 with time offset threshold $|\Delta t_{th}| = 10$ ns, 50 ns, 100 ns and 500 ns. In Fig. 5 we plot the measured syntonization accuracy of RMO-3 for the above four time offset thresholds. In general, tighter thresholds result in better clock syntonization. However, once the residual frequency offset reaches the digital tuning word resolution limit as in the vicinity of $|\Delta t_{th}| = 50$ ns, further reduction of the threshold will no longer improve the syntonization accuracy. Fig. 6 reveals the similar effect of the time offset threshold on the effective clock frequency drift rates measured in the experiment. It appears that a time error threshold near $|\Delta t_{th}| = 50$ ns results in the best result under our experimental conditions. This “optimal” threshold value might be limited by the clock phase measurement errors in the experiment. To demonstrate how the synchronization threshold affects the time interval between two consecutive clock updates, we depict the averaged clock update time intervals in the experiment as a function of the threshold values in Fig. 7. A linear fit indicates that the time interval increases in average by 6.3 hours as the time offset threshold increases by 100 ns.

To further demonstrate the effectiveness of the Slave-Master-Cs cascade timekeeping scheme and the real-time linear regression algorithm, we have derived the Allan variances of the slaved RMO-3 and XMO-15 and compare them in Fig. 8 with the unslaved RMO and XMO [4]. The measured Allan deviation confirms that both the slaved RMO-3 and XMO-15 have significantly improved their long-term frequency stability for averaging times $\tau > 10,000$ seconds. In this region, random walk noise dominates in Rb atomic clocks and in crystal oscillator clocks as well.

Fig. 5. Measured clock syntonization accuracy (averaged residual fractional frequency offset between RMO-3 and the Cs clock) for time offset thresholds $|\Delta t_{th}| = 10$ ns, 50 ns, 100 ns and 500 ns.

Fig. 6. Measured residual frequency drift rates of RMO-3 relative to the Cs clock for time offset thresholds $|\Delta t_{th}| = 10$ ns, 50 ns, 100 ns and 500 ns.

V. CONCLUSION

In order to help satellite operators understand and resolve satellite constellation timekeeping issues, we have constructed a unique spacecraft atomic clock flight simulation and test facility at The Aerospace Corporation, and exercised two space flight RMO clocks to demonstrate its capabilities in studying satellite constellation timekeeping algorithms and operations. In this work, we have demonstrated a Slave-Master-Cs cascade clock synchronization scheme with a real-time linear regression algorithm. Our results validate the concept that a slaved crystal oscillator clock (XMO) can perform well with a long-term frequency stability similar to an Rb atomic clock.
(RMO) in a satellite constellation. Specifically, we have achieved an effective long-term frequency drift of $1.7 \times 10^{-14}$ per day for the Rb clock RMO-3, which is synchronized to a GPS-receiver calibrated Cs reference clock. More significantly, using the cascade timekeeping scheme the long-term frequency drift of the crystal oscillator clock XMO-15 is reduced from $1.1 \times 10^{-11}$ per day to $2.5 \times 10^{-13}$ per day. In addition, we prove that the real-time linear regression algorithm developed in this work can efficiently detect and correct the clock’s frequency drifts and frequency jumps on real time. The reported results are useful references for system engineers and satellite operators in developing satellite constellation timekeeping architectures and in resolving satellite timekeeping operational issues. In our future work, we plan to extend the simulation experiment to GNSS clocks.

ACKNOWLEDGMENT

We acknowledge contributions by D. L. Caponi in programming the algorithm and helpful discussions with J. C. Camparo.

REFERENCES


Robust Clock Ensemble for Time and Frequency Reference System

Qinghua Wang, Fabien Droz, Pascal Rochat
Orolia Switzerland SA (Spectratime)
Neuchâtel, Switzerland
qinghua@spectratime.com, droz@spectratime.com, rochat@spectratime.com

Abstract—A robust clock ensemble is proposed for the time and frequency reference system to improve the robustness and performance of the system. Studies on the feasibility of hardware and algorithm approaches have been conducted. All clocks in the ensemble are locked in phase and frequency via the steering loop. The system performs corrections on the master clock in function of weighted averaging of clocks to generate one ensemble output, and the clock fault detection and compensation is implemented in real time with minimum three clocks powered. As the design has been demonstrated on an elegant breadboard of the Robust On-board Frequency Reference Subsystem, this concept is proposed for the next-generation of Precise Timing Facility. Simulation results have demonstrated its capability and simplicity to provide a smooth and reliable timing or frequency output even in presence of clock feared events.

Keywords—clock ensemble; time and frequency system; robust

1. INTRODUCTION

A novel concept of a Robust Clock Ensemble (RCE) for the Time and Frequency Reference System was firstly proposed by Spectratime (SpT) in 2009 [1]. Over the last years SpT has conducted studies on the feasibility of hardware and algorithm approaches. With minimum three clocks powered, all clocks are kept in phase and frequency via the steering loop. The system performs corrections on the master clock in function of weighted averaging of clocks named as ONe CLock Ensemble (ONCLE). A simple approach of the real-time Clock Fault Detection and Compensation (CFDC) based on low-pass recursive filters is implemented. This allows reliable detection of the clock feared events (including clock hard failure, phase jump, frequency jump and White Frequency Noise (WFN) level increase), as well as jump correction, failed clock removal, healthy clock inclusion, and clock switchover. This clock ensemble generates a frequency output signal with improved robustness and performances.

The feasibility concept has been demonstrated under the European GNSS Evolutions Program in the frame of ESA contract [2]. SpT with the support of GMV has developed an Elegant BreadBoard (EBB) of the Robust On-board Frequency Reference Subsystem (FRS) as an improved version of the on-board Clock Monitoring and Control Unit (CMCU) in present Galileo system. The self-standing unit is able to receive and process a number of on-board clock signals, from Rubidium Atomic Frequency Standards (RAFS) or Passive Hydrogen Maser (PHM), and generate an output signal with improved availability and robustness. Tested under various clock degradation scenarios, this unit has demonstrated its capability to autonomously detect, correct, isolate and remove (or include) the identified clock with very minor impact on the output signal.

The Precise Timing Facility (PTF) is a key element of the Ground Mission Segment to generate the Galileo System Time (GST) [3]-[4]. The Active Hydrogen Maser (AHM) provides the physical realization of GST, insuring the extremely high short-term stability required for the navigation purpose.

With the cooperation of Consorzio Torino Time, SpT developed the Backup Active Hydrogen Maser (AHM) Steering algorithm to keep the backup AHM in phase to the master via the phase stepper [5]-[6]. Among several algorithms having been implemented in PTF monitoring function, SpT designed the Clock Anomaly Detection algorithm on the basis of the phase comparison measurements in sliding window to identify anomalies in GST(MC) and clocks. When the anomaly occurs, a warning is raised. The automatic clock switching is not allowed unless the loss of the nominal AHM signal.

In order to ensure a fully continuous and performance improved GST, real-time clock anomaly detection and automatic fault compensation or clock switch-over, a simple RCE solution is proposed for the next generation of PTF. An ensemble of at least three Active Hydrogen Masers (AHM) will insure the GST with advantages of robustness, full redundancy and improved performance. In particular the RCE will avoid the discontinuity in frequency and phase of the output signal to GST; thus it will allow a reliable satellite clock modeling which is essential to support the navigation time keeping performance as the key purpose of the PTF.

A mathematic model has been developed to simulate the expected RCE output versus four AHM inputs. The principal function of Phase-Locked Loop (PLL) based on the Femtostepper, the phase comparator and the Proportional-Integral (PI) filter, as well as functions of ONCLE and CFDC, have been established in the model. All algorithms can be implemented into a very simple industrial micro-controller chip. The simulation under various test cases has demonstrated its capability to provide a smooth timing output even in presence of clock feared events.
This paper will describe the algorithms for the RCE solution and provide simulation results in various cases.

II. CONCEPT AND SIMULATION OF ROBUST CLOCK ENSEMBLE (RCE)

The proposed RCE hardware and algorithm architecture supports 4 AHM inputs (minimum 3 AHM). The main functional blocks are PLL, ONCLE, CFDC and RCE output. The overall block diagram is presented in Fig. 1. A mathematic model has been developed to simulate the expected RCE output versus AHM inputs.

The RCE provides key features:

- All clocks are kept in phase and frequency to the master clock via the PLL controlling the frequency of the Femtosteppers.
- The frequency of the master clock is corrected by the ensemble, and one RCE output signal is generated with improved frequency performance comparing with individual clocks.
- The real-time detection of the clock feared events (failure, frequency or phase jump, WFN increase …) allows the fast isolation or correction of unhealthy clock to improve the robust of the timing system.

Different algorithm solutions are possible to fulfill the functional requirements. The trade-off analysis drives us to rather consider a simple reliable solution capable to be implemented in a very simple micro-controller chip.

The detailed description of algorithms is give below.

A. Phase-Locked Loop (PLL) based on PI filter algorithm

Fig. 2 shows the architecture of the steering system, forming a basic PLL. The algorithm, only the PI filter, acquires the phase difference between the secondary and master clocks, and generates a steering frequency correction to be applied to the secondary clock via the Femtostepper.

The s-transfer function of 2\textsuperscript{nd} -order closed loop is:

\[
C(s) = \frac{2\xi\tau s + 1}{\tau^2 s^2 + 2\xi\tau s + 1}
\]

where \(\tau\) is the loop time constant: 1000s, which is selected as the tradeoff of the time offset and the frequency stability to allow smooth steering of secondary clocks; \(\xi\) is the damping factor as 1.

The FemtoStepper (i.e. phase stepper), has been developed to provide phase or frequency correction of AHM signals with a 100 fs high resolution and the negligible degradation of the AHM signal phase noise and short term stability due to the reduction of output jitter.

Fig. 3 demonstrates that the secondary clock AHM2 is steered in phase with the master clock AHM1.

B. One CLOCK Ensemble (ONCLE) algorithm

The ONCLE algorithm based on weighted or simple average of clocks is proposed for the ensemble generation. The mathematic model is given in Fig. 4.

With four clock inputs for demonstration purposes (although one from the master is always 0), \(e_{w1}\), the frequency correction to the master clock, can be written as below
C. Clock Fault Detection and Compensation (CFDC) algorithm

The CFDC aims to the reliable detection and compensation of clock feared events with simple approaches. Fig. 7 illustrates three functions handling main clock fault types (in three different colors), which are addressed below.

1) Phase jump detection and compensation

The phase measurements per 1s from the phase comparator give the straightforward indication on phase jump. Strategies such as least-squares linear fitting or standard deviation computation on sliding windows are not considered yet, in order to keep the algorithm as simple as possible. In fact, phase jumps will be reduced to a factor of \(\frac{1}{4}\) in RCE output by ONCLE averaging with 4 clocks in the ensemble.

Once a phase jump exceeds the pre-defined threshold, the ‘compensation’ function makes the phase correction by the level of one ‘threshold’ to the input with 1s delay, and iterate the comparison for the next correction.

As shown in Fig. 8, the phase jump of 30ps is detected immediately in 1s after the event occurrence, and is fully corrected during next 7s, with the threshold of 50 steps (corresponding 5ps, the background noise as sampling time of 1s).

2) Clock failure fast detection and clock removal

The clock failure is typically represented by the high frequency jump. Such failure is detected by the derivative of the phase measurements from the phase comparator. The conversion from phase to frequency allows the fast detection of the high frequency jump. Fig. 9 shows the detection information before and after the derivative. The figure indicates the observations up to 15s after the frequency jump of \(8\times10^{-12}\) occurrence at 25’000s, and the failure is detectable in 1s with the derivative.

C. Clock Fault Detection and Compensation (CFDC) algorithm

The CFDC aims to the reliable detection and compensation of clock feared events with simple approaches. Fig. 7 illustrates three functions handling main clock fault types (in three different colors), which are addressed below.

1) Phase jump detection and compensation

The phase measurements per 1s from the phase comparator give the straightforward indication on phase jump. Strategies such as least-squares linear fitting or standard deviation computation on sliding windows are not considered yet, in order to keep the algorithm as simple as possible. In fact, phase jumps will be reduced to a factor of \(\frac{1}{4}\) in RCE output by ONCLE averaging with 4 clocks in the ensemble.

Once a phase jump exceeds the pre-defined threshold, the ‘compensation’ function makes the phase correction by the level of one ‘threshold’ to the input with 1s delay, and iterate the comparison for the next correction.

As shown in Fig. 8, the phase jump of 30ps is detected immediately in 1s after the event occurrence, and is fully corrected during next 7s, with the threshold of 50 steps (corresponding 5ps, the background noise as sampling time of 1s).

2) Clock failure fast detection and clock removal

The clock failure is typically represented by the high frequency jump. Such failure is detected by the derivative of the phase measurements from the phase comparator. The conversion from phase to frequency allows the fast detection of the high frequency jump. Fig. 9 shows the detection information before and after the derivative. The figure indicates the observations up to 15s after the frequency jump of \(8\times10^{-12}\) occurrence at 25’000s, and the failure is detectable in 1s with the derivative.

C. Clock Fault Detection and Compensation (CFDC) algorithm

The CFDC aims to the reliable detection and compensation of clock feared events with simple approaches. Fig. 7 illustrates three functions handling main clock fault types (in three different colors), which are addressed below.

1) Phase jump detection and compensation

The phase measurements per 1s from the phase comparator give the straightforward indication on phase jump. Strategies such as least-squares linear fitting or standard deviation computation on sliding windows are not considered yet, in order to keep the algorithm as simple as possible. In fact, phase jumps will be reduced to a factor of \(\frac{1}{4}\) in RCE output by ONCLE averaging with 4 clocks in the ensemble.

Once a phase jump exceeds the pre-defined threshold, the ‘compensation’ function makes the phase correction by the level of one ‘threshold’ to the input with 1s delay, and iterate the comparison for the next correction.

As shown in Fig. 8, the phase jump of 30ps is detected immediately in 1s after the event occurrence, and is fully corrected during next 7s, with the threshold of 50 steps (corresponding 5ps, the background noise as sampling time of 1s).

2) Clock failure fast detection and clock removal

The clock failure is typically represented by the high frequency jump. Such failure is detected by the derivative of the phase measurements from the phase comparator. The conversion from phase to frequency allows the fast detection of the high frequency jump. Fig. 9 shows the detection information before and after the derivative. The figure indicates the observations up to 15s after the frequency jump of \(8\times10^{-12}\) occurrence at 25’000s, and the failure is detectable in 1s with the derivative.
When one clock is removed from the ensemble, the ONCLE correction will be calculated by the remaining clocks and adjusted by a discontinuity compensation factor $C_p$. Equation (2) becomes:

$$e_{w1} = \frac{-(e1+e2+e3)}{3} + C_p$$  \hspace{1cm} (3)

where $C_p$ is the difference between the last data of $e_{w1}$ (1) before the intervention and the first average with the remaining clocks.

Same discontinuity compensation principle applies with the healthy clock inclusion as the new RCE input.

Fig. 10 demonstrated that the RCE output is not affected by the clock failure (with frequency jump of $8e$-$12$) and clock removal.

This algorithm is also able to handle the failure and removal of the master clock, and switch the master clock to the next healthy clock, which has been demonstrated by measurement on the FRS EBB. The hardware detection of clock loss is also available.

3) Frequency jump detection and compensation

A simple and efficient approach of the frequency jump detection is based on the recursive filters with different filtering factors.

The difference equation of a 1st-order low-pass recursive filter (Infinite Impulse Response (IIR) Filter) is:

$$y(n) = \frac{1}{N} x(n) + \frac{N-1}{N} y(n-1)$$  \hspace{1cm} (4)

where $N$ is the filtering factor.

As shown in the functional blocks in orange color in Fig. 7, outputs from recursive filters provide directly ‘averaged’ frequencies in different time constants. The subtraction between two successive recursive filters acts as a high-pass filter and gives frequency anomaly detection output, which will be compared with pre-defined threshold. The frequency jump of $5e$-$13$ is detected in 70s after the jump occurrence with the threshold of 0.7 (Fig. 11).

The compensation mechanism is as simple as which for the phase jump compensation describe above, capable to cope with all types of frequency jumps.
This detection and compensation technique requests little memories and registers, and is quite efficient for the detection of frequency jumps in various dynamics, as have been demonstrated in similar studies for onboard clocks ensemble.

The clock anomaly as WFN increase (Allan deviation degradation) is represented by more noisy frequency or phase. Therefore the algorithm is capable to detect such anomaly. Depending on the evolution of the abnormal event, the algorithm will decide if this clock just needs to be corrected, weighted, or removed.

The anomaly such as phase spikes happened in a secondary clock is also investigated. Without any compensation technique, the impact on RCE output is negligible, thanks to mainly the PI filtering and further ONCLE averaging. As shown in Fig. 14, the maximum phase spike of 250ps with the duration of 10s in one clock is mitigated in RCE output to 1.5ps, a reduction factor of 1/180.

III. CONCLUSIONS

The feasibility concept of the RCE has been demonstrated via the hardware and algorithms which were developed and verified by measurement on the EBB of the robust onboard FRS.

The simulation results of the RCE with four AHMs demonstrate the capabilities to provide a robust timing output even in presence of clock feared events, with improved output performance. The continuity in frequency and phase of RCE output signal is essential for the navigation purpose of the PTF.

The algorithms used in the RCE concept are simple enough and require the low scale memory. The whole function can be accomplished by a very simple industrial micro-controller chip.

The advantages of the RCE solution allow it a prospective candidate for the upgrade of the PTF.

Fig. 14. Phase AHM2 (master), AHM3, AHM4, X1o (phase spikes of 20-250ps with duration of 10s) and RCE output

REFERENCES

A status report on time scale generation in PTB

Andreas Bauch, Egle Staliuniene
Physikalisch-Technische Bundesanstalt
Bundesallee 100, 38116 Braunschweig, Germany
Andreas.Bauch@ptb.de

Gihan Gomah
National Institute for Standards,
P. O. Box 136, Giza, code 12211, Egypt

Abstract—This paper deals with the local realization of Coordinated Universal Time UTC by PTB, named UTC(PTB), and PTB’s free atomic time scale TA(PTB). Since February 2010 UTC(PTB) has been derived from an active hydrogen maser steered in frequency by PTB’s primary fountain clocks. Since then and up to early 2015 the time difference UTC – UTC(PTB) was always less than 9 ns. From early on it was foreseen that the steering of UTC(PTB) could also be derived from the ensemble of caesium beam clocks operated in PTB. When doing so in practice, it appeared that the algorithm of combining the inputs of individual clocks was not perfectly chosen. The properties of the time scales as generated up to now as well as the new strategy for a time scale reflecting better PTB’s ensemble of thermal beam caesium clocks is presented in this paper.

Keywords—time scale algorithm, Coordinated Universal Time, International Atomic Time

I. INTRODUCTION

Since February 2010 UTC(PTB) has been realized using an active hydrogen maser as signal source whose frequency is steered via a commercial phase micro stepper. Steering is based primarily on the comparison between the maser involved and PTB’s primary fountain frequency standards CSF1 and CSF2 [1, 2]. In the long term, UTC(PTB) is steered towards UTC based on the data published in the BIPM Circular T. A brief description of the hardware and software involved is given in Section II. In Section III we discuss properties of UTC(PTB) and the “free atomic time scale” TA(PTB) that is generated in a similar way. UTC(PTB) serves as the basis for all of PTB’s time services and international time comparisons. It will also serve directly as the time and frequency reference for the upcoming ACES campaign [3].

From early on it was foreseen that the steering of UTC(PTB) could also be derived from the ensemble of caesium beam clocks operated in PTB in case that data from the fountain clocks were not available (and described as Option 2 in [2]). It turned out that CSF1 and CSF2 provided data with greatest reliability so that this option needed to be applied on very rare occasions only. Since several months a timescale generation “test bed” has been used to realize UTT(PTB), based on the inputs of five thermal beam clocks. Contrary to the expectations, a considerable difference between UTC(PTB) and UTT(PTB) was noted and motivated a study into the causes thereof. The results are presented in Section IV. The final Section contains first results obtained using the new algorithmic strategy.

II. REALIZATION OF UTC(PTB) 2010-2015

A. Overview

Conceptually, the realization of UTC(PTB) has continuously followed the strategy laid down in [2]. The fountain frequency standards CSF1 and CSF2 have served as primary reference all times. Technically, however, several changes have occurred since 2010. The line centre of the Ramsey feature recorded in each of the two fountains is now determined with the help of a frequency synthesis chains locked to one of PTB’s active hydrogen masers (AHM) and no physical output signal (5 MHz) from a fountain is generated. Instead, two data files that contain the offset of the maser frequency from the unperturbed transition frequency as realized in each fountain, represented as numerical values with an averaging time of one hour, are produced. To the best knowledge the frequency corrections affecting the line centre [1] are predicted and applied to the data in real time. A data file contains empty lines when a fountain is intentionally declared unhealthy or when a fault in the operations has been detected.

Second source is the comparison of the AHM to the ensemble of thermal beam clocks, the primary clocks CS1 and CS2 [4], and three commercial clocks of type 5071A using their one-pulse-per-second (1 PPS) output and a time interval counter.

B. Hardware

For the time scale generation, the output frequency of an AHM is steered using a phase micro stepper (PMS, Spectra Dynamics Inc., HROG-5) and its in-built divider to 1 PPS. The PMS allows the introduction of a fractional frequency offset between the input and the output 5-MHz signals with a resolution of $6 \times 10^{-17}$. The PMS noise contribution remains lower than that of the typical noise of an AHM as long as the relative offset introduced remains bounded to $2 \times 10^{-13}$ in magnitude which we try to obey in practice. At the end of 2014, four generation chains of that kind are active. UTC(PTB) and UTH(PTB) are generated using two different masers and the equipment has been installed in two different locations in the building. TA(PTB) shares the maser with UTH(PTB), but no alignment to an external time scale is aimed at, the process...
is explained in the next section. UTT(PTB) shares the AHM with UTC(PTB) and serves as the test-bed as explained later.

C. Software features

The PMS is commanded daily to shift the 5 MHz input frequency by the relative amount

\[ \delta f_{\text{steer}} = \delta f_{\text{ref}} + \delta f_{\text{rate}} + \delta f_{\text{offset}} \]  

(1)

Here \( \delta f_{\text{ref}} \) is intended to adjust the frequency of the AHM to the rate of the reference clock. The summand \( \delta f_{\text{rate}} \) reflects the reference clock rate with respect to TAI as taken from Circular T and the accompanying data file rTAI\(^5\). The third term \( \delta f_{\text{offset}} \) ensures the long-term steering of UTC(PTB) towards UTC and is calculated as \( \{\text{UTC} - \text{UTC(PTB)}\}_{\text{LRD}} / 60 \text{ d} \), where the time difference is that of the last reported day (LRD) in the most recent issue of the Circular T. Obviously, the two last summands in (1) change monthly, whereas the daily steering actually affects only the summand \( \delta f_{\text{ref}} \). In case that one of the fountains is the steering reference, \( \delta f_{\text{ref}} \) is calculated based on daily average values. In all cases that the thermal beam clocks serve as the reference, \( \delta f_{\text{ref}} \) is calculated as extrapolation over previous results obtained during a configurable number of days, subject of Section IV.

III. PROPERTIES OF UTC(PTB) AND TA(PTB)

We are limiting our data analysis to the period since UTT has been realized, starting in September 2013. In Fig. 1 the comparison of UTC(PTB) to the two time scales published by BIPM are shown. We note the somewhat larger excursions of UTCr and occasional alignment steps of UTCr to UTC as explained in [6]. These may have implications for the steering of a time scale but using UTCr as a reference is quite convenient due to its rapid availability. During the period shown in Fig. 1 the uncertainty in the time differences UTC-UTC(PTB) were reduced from 1.8 ns to 0.8 ns due to new calibration results [5]. PTB has stated Calibration and Measurement Capabilities (CMC) for time scale comparison and frequency measurement. All are still based on the situation before 2010 and represent very conservative estimates [7] in view of the current realization of UTC(PTB). In the case of CMCs for frequency the traceability is based on the operation of primary clocks, not on the key comparison as it is the case for most other metrology institutes. In practice today, there is no signal any more from the fountains: one would always use UTC(PTB) rf signals as reference for frequency measurements. In Fig. 2 the deviation between the UTC(PTB) scale unit and the SI second as realized with CSF1 and CSF2 over a 4-months period is plotted, here shown as relative frequency difference. The mean relative offset is well below \( 1 \times 10^{-15} \) and thus close to the uncertainty \( u_B \) of the CSF1 and CSF2, respectively [1]. The frequency instability is depicted in Fig. 3. The values include the combined instability of two independent measurement processes, one relating the fountains to an active hydrogen maser, and the other relating the hydrogen maser to UTC(PTB), respectively. Such data have not been produced regularly, explaining the reduced time span of the data shown. The results comply favorably with the stated CMCs.

![Fig. 1. Time scale differences \( \Delta T = \text{UTC-UTC(PTB)} \) (O) and \( \text{UTCr-UTC(PTB)} \) (+) since September 2013.](image1)

![Fig. 2. Relative frequency difference between UTC(PTB) and the fountain frequency standards CSF1(blue) and CSF2 (pink), daily averages.](image2)

![Fig. 3. Frequency instability of the data of Fig. 2, expressed by the Allan Standard Deviation \( \sigma_\alpha(t) \).](image3)
TA(PTB) has been generated using similar equipment and software as UTC(PTB), except that the third summand in (1) is neglected. The scale unit of TA(PTB) thus represents more closely the SI second as realized with the CSF1 and CSF2, but in view of the foregoing the difference is of little practical relevance.

IV. THE TIME SCALE GENERATION TEST BED

A. Analysis of data up to end 2014: available clocks

A few words on the AHM are needed at this point, in order to better understand the following. Currently PTB operates 4 AHM, all manufactured by Vremya-Ch, Nizhny Novgorod, Russia, and designated H5, H6, H8, and H9. They were put into operations in 2004, 2006, 2009 and 2013, respectively, and all had since 2010 been used in the realization of UTC(PTB) for some time. Their performance values are nominally very similar, in practice this was not the case. This is not immediately visible in data as shown in Fig. 1, but should be noticeable in carrier-phase based GPS time comparisons involving UTC(PTB). When the steering shall be calculated based on averaged and extrapolated values with respect to thermal beam clocks in case of the UTT scale, the maser properties are much more relevant. In Fig. 4 the daily, non-weighted steering values used in the UTT generation is depicted, based on comparisons with CS1 (black), CS2 (red), and three commercial clocks of type 5071A (green, yellow, blue, representing serial numbers 128, 415, and 1078, respectively). This color coding remains valid in the following figures as well. We note three distinctive periods. Up to MJD 56760 the newest maser, H9, was employed, showing a remarkably small linear frequency drift (and frequency instability at $\tau = 1$ day, not visible here). Due to a component failure a switch to H6 was needed, which initially had a large frequency offset from nominal. This was corrected around 56780, and H6 has been employed until February 2015. For better visibility, the third period is shown in better detail in Fig. 5.

![Fig. 4. $\delta f_{\text{Steer}}$ calculated based on the comparisons of AHM with the thermal beam clocks, CS1 (black), CS2 (red), clocks of type 5071A (green, yellow, blue, representing serial numbers 128, 415, and 1078, respectively).](image)

B. Analysis of data up to end 2010: clock instabilities

For a characterization of the clocks we use the relative frequency instability $\sigma_f(\tau)$ calculated from the data shown in Fig. 5. In addition three plots representing the instability of three AHM with respect to PTB’s fountains – either CSF1 or CSF2, whatever available - are overlaid in Fig. 6. The 5 caesium clocks show their intrinsic white frequency noise level at $\tau = 1$ day. For long averaging the instability of the AHM H6 with its prevailing frequency drift (symbol +) gets dominant. The oldest maser (H5, symbol Δ) has a remarkably low frequency drift, but we had to exclude it from the time scale generation because of intermittent maser signal breakdowns, accompanied with frequency outliers. The other maser (H8, symbol ∇) is much less performant. This distinctly different maser performance requires in principle different strategies for the calculation of the daily values $\delta f_{\text{Steer}}$. Based on data taken in 2009 we had decided that $\delta f_{\text{Steer}}$ for day N was to be calculated based on a linear extrapolation over the last 25 days for all clocks except CS2 for which the interval was 16 days, and this practice had never been questioned.

![Fig. 5. $\delta f_{\text{Steer}}(1)$ calculated based on the comparisons of AHM H6 with the thermal beam clocks since mid May 2014, color code as in Fig. 4.](image)

![Fig. 6. Relative frequency instability in comparisons between AHM H6 and the five caesium clocks (color code as in Fig. 4), and between the fountains CSF1 and CSF2 to AHM H6 (+), H5 (Δ) and H8 (∇).](image)
C. Analysis of data up to end 2010: combining $\delta f_{\text{Steer}}$

When calculating the individual $\delta f_{\text{Steer}}$ we correct the result for the predicted rates of the individual clocks with respect to TAI as reported by BIPM [5, 2]. Otherwise the time scale would start deviating from UTC, unless counteracted by the third summand in (1). If the clock rates would not be corrected for, a more sophisticated software, handling changes in the clock composition from day to day, would be needed. The individual $\delta f_{\text{Steer}}$ values are finally combined using statistical weights. For simplicity we used the values reported in monthly wTAI files by BIPM [5], normalizing the sum of the weights to one. The wTAI clock weights reported between July and December 2014 and normalized to one are shown in Fig. 7. A very dominant weight was given to CS1. This fact combined with the periodic excursions in the rate of CS1 that can be seen in Fig. 5 explains a good part of the UTT deviations from UTC and UTCr, depicted in Fig. 8. The gap in the data around MJD 56820 was caused by a fault in the data storage procedure. In the following section we discuss the approach taken to avoid the large excursions to happen in the future.

![Fig. 7](image-url) Statistical weights of PTB’s caesium clocks in the calculation of TAI in the second half of 2014, normalized to one, color code as in Fig. 4.

![Fig. 8](image-url) Time scale comparison UTCr-UTC(PTB) (blue), UTC(PTB)-UTT(PTB) (red) and UTCr-UTT(PTB) (green).

V. Update of UTT generation procedures

In view of the fact that the PTB clock ensemble is essentially static, there are just two knobs to adjust in the algorithms used in the realization of UTT, the prediction period for determining $\delta f_{\text{Steer}}$ and the weights for their combination. As the determination of $\delta f_{\text{Steer}}$ for day N is based on a linear prediction over past data, the linear maser drift is compensated anyway, and the time constant (prediction period) should be decided looking at the instability plots generated after removal of the linear frequency drift. Having removed the linear drift means that the Cs-clock data essentially reflect the intrinsic properties of each clock, assuming that their own drift is negligible compared to that of the maser. The maser vs fountain data in turn reflect the maser instability apart from its linear drift [1]. Fig. 9 demonstrates that for the “good masers” the instability curves do not cross up to quite long averaging times. The CS2 instability is for all $\tau$-values lower than that of CS1, and that was not properly reflected in the statistical weights attributed to the clocks hitherto. For CS1 and CS2 we thus set a common prediction period of 30 days since January 2015. The frequency instability plots of the three commercial clocks are of course slightly different, we concluded, however, that averaging for longer than about 20 days is not meaningful as the instability does not drop further by a significant amount. It is obvious that all these prediction intervals would have to be adapted when a switch from one maser to another would be needed in the future. That had not been done all the years since 2010.

![Fig. 9](image-url) Relative frequency instability in comparisons between AHM H6 and the five caesium clocks (color code as in Fig. 4), and between the fountains CSF1 and CSF2 to AHM H6 (X), H5 (A) and H8 (V), linear frequency drift removed beforehand individually in all data sets.

The statistical weights used hitherto had been determined from a global analysis made within ALGOS and could change from month to month simply as the ensemble of clocks changed. A straightforward option would be to base the weights on the performance of each clock with respect to the average of clocks in the past. In view of the small number of clocks we decided to take a more cautious approach, calculating the weights as $1/\sigma^2$, $\sigma$ being the standard deviation.
of the last 6 monthly rates of the clocks published in rTAI. In
principle, if equal weights would be attributed, the value was
0.2, so we decided to limit the maximum statistical weight to
0.4. The sum of weights is fixed to one. In Fig. 10 the weights
calculated using these rules for the months July 2014 to
(including) March 2015 are depicted in Fig. 10. Comparison to
Fig. 7 reveals the clear differences.

![Graph showing the weights calculated for the months July 2014 to March 2015.]

Application of the new procedures was complicated by the
fact that one of the commercial caesium clocks came close to
its end-of-life failure and had to be replaced by a new unit that
had not been in operation for an extended period. In Fig. 10
one can note the decline in weight of the old clock (green
trace) and the large weight obtained by the new clock, SN
2987 in March 2015 (magenta). Maser H6 needed to be
removed from service as its hydrogen source is getting empty.
So a clear assessment whether the new algorithms for the
realization of UTT really provide the anticipated
improvements cannot be given at the time of writing this paper
as the new weights have been applied only after 12 March
2015 when the last Circular T had been published.

VI. CONCLUSION

The realization of PTB’s reference time scale UTC(PTB) has
been based on steering a AHM with respect to comparison
results to the two fountain frequency standards CSF1 and
CSF2. The quality of UTC(PTB) was competitive and the time
scale could be used in demanding applications, such as the
performance monitoring of the Galileo space clocks
throughout 2014, see Fig. 8 in [8]. In order to be on the safe
side, we aim to establish a backup realization with similar
performance based on the conventional caesium clocks of
PTB. Some deficiencies in the previous procedures were
detected and remedied. We are confident that deviations from
UTC can be limited to ±10 ns or better in the future based on
the strategy presented in Section VI.

ACKNOWLEDGMENT

G.G. gratefully acknowledges the support of the Egyptian
Government funding her hands-on training at PTB. The
quality of the PTB time scales is highly dependent on the team
operating the fountain clocks and the team managing the daily
operations of the hardware and software involved, and the
authors are grateful for their support.

REFERENCES

evaluation of the atomic caesium fountain CSF1 of PTB”, Metrologia,
vol. 38, no. 4, pp. 343–352, 2001; and V. Gerginov, N. Nemitz, S.
evaluation of the caesium fountain clock PTB-CSF2”, Metrologia, vol.
of UTC(PTB) as a fountain-clock based time scale”, Metrologia, vol. 49,
S43-S54 (2005).
folders cirt, rates, weights.
element of the Galileo operational phase”, Proc. of the Joint Meeting of
the 2015 IEEE IFCS and of the 29th EFTF, Denver, Colorado, April
2015.
Abstract—In this work we will present the Makkah timescale generation, its setup and changes over time. As a special challenge the Makkah Time Scale is located in two labs, which are connected with a zero delay system via a 200m glass fibre. Having a new type of time interval counters, we compared measurements of performances and capabilities.

I. THE MAKKAH TIMESCALE

A. History

The Makka Time Centre is located in the famous Makkah Clock Tower. It is about 450m high -out of 600m total building height- in the Kingdom of Saudi Arabia. This means the MTC lab is located just above the north clock face. The project started with a turnkey system from Symmetricom (now MicroSemi) consisting of five HP5071 high performance cesium standards. Systems were installed in 2011, and the data feeding to the BIPM started in November 2012.

In this system one of the cesiums was the master. The steering was done through the GPS signal. The time signal is distributed through two NTP servers - Symmetricom SyncServer S350- to the four tower clocks, the technical subsystems in the tower and the telescope stations of a worldwide network for Crescent observation. In addition, there are negotiations with the sacred mosque to get them connected via a dedicated line. Giving the current setup and the ambitious futuristic plans, the system encountered some limitations and an upgrade is planned. The new upgrade will be staged in multiple phases, with a steady and continuous increase in overall performance measures.

B. Phase Two

In the first step we replaced the old core system completely by a new one, which gives directly the uncertainties of a cesium standard. Namely we took our master time signal straight from one of the cesium standards (Cs-1). The standards and the GPS receiver - calibrated TTS-4 GNSS receiver- are measured against our master signal. The results are published via monthly (UTC) and daily (UTCr) clock files to the BIPM. The TTS-4 sends its measurements directly.

Through this we gained a much better uncertainty but we lost the steering mechanism. Since November 2014
C. Phase Three

The second step which is currently under progress, is to get the steering back in the game without losing the achieved uncertainties. To achieve the goal the number of our Cesiums increased from five to eleven. But, the six new standards are located in a different floor as part of an exhibition. This made it necessary to connect a temperature compensating zero delay system from TimeTech that transfers the master time signal via a 200m fibre optics cable to the 2nd lab where their time differences are measured with a multi channel TIC from TimeTech. The results are then sent back via a dedicated 200m TP cable to the controller. A second MCTIC is already installed next to the First Step SR620-Switch-combination in the main lab, measuring the first five Cesiums in parallel. The next thing will be to include the micro stepper from SpectraDynamics and to implement an algorithm for the steering.

II. TIC SR620 vs. MCTIC 10409

The aim of this examination is to test the multichannel time interval counter from TimeTech (MCTIC 10409) against the well known Stanford Research Systems TIC (SR620).

A. First Setup

In a first attempt we used two of our cesiums (Cs-1 and Cs-2) to measure the time interval between their 1PPS signals. We used the Cs-1 master signal as start and the 1PPS from Cs-2 as the stop signal. For interconnection different distribution amplifiers were used and for making as few changes as possible to our running
system also the Cytec signal switch was used. But the latter was constantly set to input port 0, i.e. our Cs-1 master signal.

By sending five copies of the stop stop signal to different channels of the TimeTech we checked the differences between the different channels and also if it makes a difference which slot is used. Finally we got one measurement series from the SR620 and five series from the MCTIC comparing the same start and stop signals and each over a period of one minute, i.e. 60 measurement values.

After normalizing each channel by subtracting its individual mean value, we get the following standard deviations:

<table>
<thead>
<tr>
<th>Table I.</th>
<th>CS-1 &amp; CS-2, 1 MIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device</td>
<td>Start</td>
</tr>
<tr>
<td>Stanford Research</td>
<td>A</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
</tr>
</tbody>
</table>

So this shows that the SR620 seems to be more accurate than the TimeTech by nearly a factor of 2. The graph shows two peaks on the TimeTech (at 23m57s and at 24m27s) in all channels simultaneously while the Stanford did not show these. Because it is hard to tell if these are caused by differences of the time signal sources or by the measurement devices, we decided to change the setup.

**B. Second Setup**

In the second setup we reduced the complexity by using one time signal as start and the same signal with additional delays as stop signal for all counters. So we simply connected one of the FS735’s outputs with the SDI distribution amplifiers input and removed Cs-2 from the setup. Now the difference of the start and stop signal for each channel was only defined by the cables used to make the individual connections.
After normalizing again each channel, we get the following standard deviations:

<table>
<thead>
<tr>
<th>Device</th>
<th>Start</th>
<th>Stop</th>
<th>σ [ps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford Research</td>
<td>A</td>
<td>B</td>
<td>5.9</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.1 Ch.2</td>
<td>12.6</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.1 Ch.3</td>
<td>15.7</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.1 Ch.4</td>
<td>14.2</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.2 Ch.1</td>
<td>13.1</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.2 Ch.2</td>
<td>12.1</td>
</tr>
</tbody>
</table>

Obviously it is not justified to take the mean values as replacement for the (unknown) real value. In the new setup the mean value of each channel differs from the real value only by an (unmeasured) constant, which is given by the delays of the cables used and the devices passed through. Also the graph now shows no more irregularities and it can be assumed that all variations are now only caused by the measurement devices themselves.

After having found a good setup we went on and increased the number of measurements to get more profound results. This time we let the experiment run for one hour resulting in 3600 measurement values per channel. From this we get the following standard deviations:

<table>
<thead>
<tr>
<th>Device</th>
<th>Start</th>
<th>Stop</th>
<th>σ [ps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford Research</td>
<td>A</td>
<td>B</td>
<td>5.8</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.1 Ch.2</td>
<td>13.2</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.1 Ch.3</td>
<td>14.6</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.1 Ch.4</td>
<td>14.0</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.2 Ch.1</td>
<td>13.7</td>
</tr>
<tr>
<td>TimeTech</td>
<td>Sl.1 Ch.1</td>
<td>Sl.2 Ch.2</td>
<td>14.0</td>
</tr>
</tbody>
</table>

The Stanford Research TIC SR620 is more accurate than the TimeTech MCTIC 10409 by a factor of approximately two. The former has a confidence interval of about ±12 ps, while the latter has about ±30 ps. For our laboratory both devices are acceptable compared to our general contemporary performance. Because of the advantage of being able to measure all of our standards in parallel every second, we will stay for the time being with the TimeTech devices for the permanent performance observation and measurement result generation.

ACKNOWLEDGMENT

The authors would like to thank especially Armin Söring for his continues support from the very first steps of the Makkah Time Centre until now, the Timekeeping community for being there when needed and especially the TÜBİTAK/UME team for the great improvement achieved through their support in our second phase system.
Balanced Low-Loss 2-IDT Double Mode SAW Filter with Narrowed Passband and Improved Selectivity

Doberstein Sergei
PJSC «Omskiy Nauchno Issledovatelskiy Institut Priborostroenia» (ONIIP)
Omsk, Russia
E-mail: info@oniip.ru

Abstract—This paper presents the new balanced low-loss DMS filter with the narrowed passband and improved selectivity on 42° YX LiTaO₃. The filter is realized as two-transducer scheme. A narrow passband of the filter is obtained by choosing a definite gap length between the input and output IDTs. Also the both long weighted input IDT and reflectors are used. A shape of the frequency response of the filter is determined by the weighting function of the input IDT, number of the electrodes in output IDT and reflectors, relationship between the electrode pitches in the input IDT, output IDT and reflectors. An optimization of the mentioned parameters with a computer simulation using equivalent circuit model allows to get a specified high selectivity of the filter with a fractional bandwidth of 0.5-0.6% under condition that the insertion losses are low and input and output impedances are close to the specified real values. To decrease the input and output impedances the parallel connection of the two filters in the different acoustic tracks is used. To improve the selectivity of the filter the cascaded connection and phase weighting the input IDT are employed. The 300 MHz sample of the balanced SAW filter has shown an insertion loss of 3 dB, 2-dB bandwidth of 1.5 MHz, stopband attenuation of 70 dB at ±12 MHz offsets from a center frequency in a 75-Ω balanced system. The filter did not require the matching elements and housed in the 9.1x7.1x2 mm SMD package.

Keywords—SAW filter; double mode; narrowband

I. INTRODUCTION

There are many applications and today’s market demands for the narrowband SAW filters with a fractional bandwidth of less than 1% having low loss and high selectivity. Recently the low loss 3-IDT double mode SAW (DMS) filter with fractional bandwidth Δf/f₀ = 0.96% and selectivity of 60 dB was developed [1]. But the further narrow passband on the retention of the low loss and high selectivity is required for front-end of the VHF receivers for example. This paper presents the new balanced low-loss DMS filter with the narrowed passband and improved selectivity on 42° YX LiTaO₃. The filter is realized as a two-transducer scheme on the longitudinal first and second resonance modes (Fig. 1) [2]. A passband width of DMS filter depends on a frequency difference between these modes and is determined by the gap length S₂ between the input and output IDTs. The frequency difference between the first and second resonance modes in the two-transducer filter is always less than frequency difference between the first and third resonance modes in the three-transducer filter, all thing being equal [2, 3]. Also the frequency responses of the input weighted IDT, output IDT and reflectors affect on the passband width. Then the both long weighted IDT and reflectors are used for narrowing the passband [1, 4]. A shape of the frequency response of the filter depends on the weighting the input IDT, frequency responses of the output IDT and reflectors. This shape is determined by the weighting function of the input IDT, number of the electrodes in output IDT and reflectors, relationship between the electrode pitches in the input IDT (P₁IDT), output IDT (P₀IDT) and reflectors (Pref) (Fig. 1).

Fig. 1. Balanced 2-IDT DMS filter.

Fig. 2. Equivalent circuit of the balanced 2-IDT DMS filter.
Fig. 3. Simulated frequency response of the 2-IDT DMS filter with long input weighted IDT.

The balance operation of the filter is made by symmetrical connection of the input IDT and output IDT to the loads because these IDTs have not a common grounded busbar. To decrease the input and output impedances the parallel connection of the two filters in the different acoustic tracks is used [5, 6]. To improve the selectivity of the filter the cascaded connection and phase weighting [7] the input IDT are employed.

II. SIMULATED AND EXPERIMENTAL RESULTS

The constructional and topological optimization of the narrow 2-IDT filter was provided with a computer simulation using an equivalent circuit model [6]. An equivalent circuit of the balanced 2-IDT DMS filter is shown on Fig. 2. Here \( P^{(1)} \), \( P^{(2)} \) are the mixed matrixes of the input and output IDTs, \( Z_0 \) is a characteristic impedance of the medium between IDTs and reflectors, \( V_0 \) is a SAW velocity, \( R \) is equivalent impedance for a reflector.

An optimization of the mentioned parameters of the filter topology (Fig. 1) allows to get the specified selectivity of the filter with a fractional bandwidth \( \Delta f/f_0 = 0.5\text{-}0.6\% \) under the condition that the insertion losses are low and input and output impedances are close to the specified real values. In this case, the passband width is considerably narrowed as opposed to known methods [1, 2, 4, 6]. The optimal parameters of the filter topology are shown in Table. Here \( \lambda \) is SAW wavelength on the center frequency of the filter. The gap length between IDTs \( S_2 \), gap length between IDT and reflectors \( S_1 \) are gap length between ends of the adjacent electrodes of the IDTs and reflectors.

The optimization was realized by successive approximations. The experimental samples were fabricated in the intervening stages. Frequency responses, insertion losses, filter input and output impedances were carefully measured, and the calculations and filter topology were corrected.

Fig. 3 shows a simulated frequency response of the filter on 42°YX LiTaO₃ in balanced 100-Ω system. The filter has shown a 1-dB fractional bandwidth \( \Delta f/f_0 = 0.6\% \) with minimal insertion loss and ripple, stopband attenuation of around 40 dB at ±12 MHz offset from a center frequency of 300 MHz. The measured frequency response of the filter on 42°YX LiTaO₃ is shown on the Fig. 4. The filter has been symmetrically connected to a balanced 100-Ω system. Measurement were carried out by the network analyzer and balanced transformers. The losses of the balanced transformers were eliminated from the measured insertion losses of the filter. The input phase weighted IDT had 163 electrodes. 23.5 finger pairs were used in each transverse section of the input IDT for self-matching. A 0.62 μm electrode thickness (h) was chosen for obtaining the value of \( h/\lambda = 4.4 \% \). This is typical condition of the minimum insertion loss for the DMS filter [2].

At a center frequency of 300 MHz the filter has shown an insertion loss of 1.5 dB, 1-dB bandwidth of 1.8 MHz with a low ripple of 0.5 dB and stopband attenuation of 40 dB at ±12 MHz offset from the center frequency. A comparison between Fig. 3 and Fig. 4 shows that a good agreement is available between the simulated and measured frequency responses.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap between IDTs</td>
<td>S₂</td>
<td>0.5λ</td>
</tr>
<tr>
<td>Gap between IDT and reflector</td>
<td>S₁</td>
<td>0.25λ</td>
</tr>
<tr>
<td>Number of the electrodes in the reflector</td>
<td>N_ref</td>
<td>120</td>
</tr>
<tr>
<td>Number of the electrodes in the input IDT</td>
<td>N_in</td>
<td>163</td>
</tr>
<tr>
<td>Number of the electrodes in the output IDT</td>
<td>N_out</td>
<td>47</td>
</tr>
<tr>
<td>Relation between the electrode pitches in the output IDT and input IDT</td>
<td>P_out/P_in</td>
<td>1.0021</td>
</tr>
<tr>
<td>Relation between the electrode pitches in the reflectors and output IDT</td>
<td>P_ref/P_out</td>
<td>1.01</td>
</tr>
<tr>
<td>Overlapped aperture</td>
<td>W</td>
<td>40λ</td>
</tr>
</tbody>
</table>
Fig. 5. Measured input impedance characteristic of the 2-IDT DMS filter with long input weighted IDT.

The measured input and output impedance characteristics of the filter are shown in Fig. 5, 6 respectively. As will be seen from Fig. 5, 6 at the center frequency of 300 MHz the input and output impedances of the filter are close to real value of 100 $\Omega$.

An application of a shorted input phase weighted IDT leads to widening the passband of the filter. Fig. 7 shows the measured frequency response of the filter with the input phase weighted IDT having the 83 electrodes. As will be seen from Fig. 7 at center frequency of 246 MHz filter have an insertion loss of less than 2 dB, 1-dB bandwidth of 3.3 MHz ($\Delta f/f_0=1.34\%$), stopband attenuation over 40 dB at ±11 MHz offset from the center frequency.

The measured frequency response of the filter with the long input phase weighted IDT having the 163 electrodes and with the same electrode pitch in the input IDT, output IDT and reflectors is shown in Fig. 8. In a 100-Ω balanced system the filter has shown an insertion loss of 5 dB, 2-dB bandwidth of 1.5 MHz ($\Delta f/f_0=0.53\%$) with a large ripple of 1 dB, stopband attenuation of about 30 dB at ±12 MHz offset from the center frequency. The shape factor of the frequency response is inadequate. To reduce an insertion loss, passband ripple, to improve the shape factor of the frequency response we must change the relationship between electrode pitches in the input IDT, output IDT and reflectors. Then we shall obtain the frequency response of the filter (Fig. 4) after the topology optimization according to Table.

To improve the selectivity of the filter we used cascading two filters. To decrease the input and output impedances we connected in parallel the two filters in the different acoustic tracks. The measured frequency responses of the cascade of two filters with the parallel connection are shown in Fig. 9 in the wide frequency range and in Fig. 10 in the passband.

Fig. 7. Measured frequency response of the 2-IDT DMS filter with shortened input phase weighted IDT.

Fig. 8. Measured frequency response of the 2-IDT DMS filter with long input weighted IDT before topology optimization.
In a 75-Ω balanced system cascaded filter has shown an insertion loss of 3 dB, 2-dB bandwidth of 1.5 MHz ($\Delta f/f_0=0.5\%$) with a low ripple of 0.5 dB, stopband attenuation of 70 dB at ±12 MHz offset from the center frequency. The measured input and output impedance characteristics of the filter are shown in Fig.11, 12 respectively. As will be seen from Fig. 11, 12 at the center frequency of 300 MHz the input and output impedances of the filter are close to real value of 75 Ω. The filter did not require the matching elements and housed in the 9.1x7.1x2mm SMD package.

### III. CONCLUSION

We have developed the new low-loss DMS filter with the narrowed passband and improved selectivity than its known prototypes [1, 2, 4, 6]. This filter have a balance operation, fractional bandpass width of 0.5% and it will be widely used as IF and front-end filter in the telecommunications equipment.

### REFERENCES


Switchable and Tunable Resonators with Barium Strontium Titanate on GaN/Sapphire Substrates

T.S. Kalkur¹, Milad Hmeda¹, Almonir Mansour¹, Pamir Alpay², Nick Sbockey³ and G.S. Tompa³
¹Department of Electrical and Computer Engineering, University of Colorado Colorado Springs, Colorado Springs, CO 80918-7150, USA
²Department of Materials Science and Engineering and Institute of Material Science, University of Connecticut, Storrs, CT 06269, USA
³Structured Material Industries, Inc., Piscataway, NJ 08854, USA

Abstract—A solidly mounted tunable barium strontium titanate (BST) based resonator was fabricated on a GaN/Sapphire substrate using a metalorganic solution deposition (MOSD) technique. An acoustic Bragg reflector was first formed on the GaN/sapphire substrate consisting of alternating layers of silicon dioxide and tantalum oxide deposited using a-spin on technique. Lower and upper electrodes were fabricated using sputter deposited platinum. The resonant frequency of the resonator could be tuned from 5.17 GHz to 5.20 GHz by applying a voltage of 8 V, resulting in tunability of about 0.6%. The quality factor of the resonator was found to depend on the applied voltage, with a maximum quality factor of 216 observed for an applied bias voltage of 8 V. The effective electromechanical coupling coefficient (kt²) of the resonator was found to be 13.1% at 8 V.

Keywords—FBAR; BST; GaN; MOSD; Tunability; Resonator.

I. INTRODUCTION

The excitement about gallium nitride (GaN) stems from its unique material and electronic properties. GaN devices offer five key characteristics: high dielectric strength, high operating temperature, high current density, high speed switching and low on-resistance[1]. These characteristics are due to the properties of GaN, which, compared to silicon, offer ten times higher electrical breakdown characteristics, three times the band gap, and exceptional carrier mobility. As present-day semiconductors used in wireless radio frequency (RF) electronics reach their limits, the communications industry faces an imminent need to develop new technologies. GaN has emerged as an in-demand semiconductor technology for future radar, electronic warfare and communications technologies. GaN is becoming very important in the implementation in high frequency, high power integrated RF circuits. Most of the efforts in GaN technology are directed towards the fabrication of high power transistors for power amplifiers. For implementation in RF circuits, a variety of RF building blocks are needed such as amplifiers, oscillators and filters. Resonators play an important role in the fabrication of RF circuits such as filters and low phase noise oscillators [3-5]. In the literature, a few papers have been reported in fabricating MEMS resonators on GaN. In this paper, we report results for a tunable bulk acoustic resonator implemented on GaN/Sapphire wafers with Barium Strontium Titanate (BST) films [6-12].

II. MATERIALS AND METHODS

The substrate was a commercially purchased 2 inch diameter GaN/sapphire template wafer, consisting of 3.5 μm of (0001) GaN on (0001) sapphire. Prior to processing, the wafer was cleaned ultrasonically using acetone and isopropyl alcohol, then blown dry in nitrogen gas. The wafer was then dehydrated by baking in a vacuum oven. An acoustic Bragg reflector was prepared on the GaN surface, in order to acoustically isolate the resonator from the damping effect of the substrate. The Bragg reflector consisted of six alternating layers of silicon dioxide and tantalum oxide, as shown in Figure 1a. The commercial silicon dioxide precursor (Allied Signal Corporation) was applied to the wafer surface and spun at 2000 rpm for one minute. The wafer was baked at a temperature of 180 °C for 2 minutes to remove the organics, then densified at a temperature 250 °C for 5 minutes. Another layer of Silicon dioxide was deposited using the same procedure. The wafer was then annealed at a temperature of 800 °C for 5 minutes in oxygen. Ellipsometer characterization showed that the deposited silicon dioxide thickness was about 265 nm. A similar process was used to deposit the tantalum oxide layers, using a commercial spin-on precursor (Kujundo Chemicals). The measured thickness of the resulting tantalum oxide layer was about 180 nm. This process was repeated in order to obtain 3 alternating layer pairs of silicon dioxide and tantalum oxide.

The bottom electrode consisted of a 20 nm titanium adhesion layer and 100 nm of Pt deposited by DC magnetron sputtering. The bottom electrodes for the resonators were patterned using standard photolithographic techniques and ion-milling. The MOSD BST layer was deposited using a spin speed of 2000 rpm for 30 seconds, followed by the same two-step baking procedure used for the silicon oxide and tantalum oxide layers. The wafer was then annealed in a furnace at a temperature of 800 °C for 5 minutes in oxygen. The thickness of the BST layer was found to be 50 nm. Two additional BST layers were deposited and the wafer was annealed at 800 °C for 30 minutes in oxygen. The top electrode consisting of 100 nm Pt was deposited by DC magnetron sputtering. The top electrode was patterned by standard photolithographic technique and ion milling. The final structure of the resonator device is shown in Fig 1a. Fig 1b shows the micrograph of the fabricated BST resonator on GaN/Sapphire substrate. The wavelength of the resonating waveform is given by (2).
where $V_n$ is the acoustic velocity, $f$ is the frequency, $C_n$ is the elastic coefficient, and $\rho_n$ is the density.

RF characterization of the resonator device was performed using a Cascade Microtech ground-signal-ground (GSG) probe and Agilent Vector Network Analyzer. The low frequency characterization was performed with an HP multi-frequency LCR meter.

III. RESULTS AND DISCUSSION

Fig. 2 shows the variation of the reflection parameter $S_{11}$ with applied voltage. With increasing DC bias voltage, the resonance frequency was found to decrease and the $S_{11}$ notch depth was found to increase. The resonance frequency was shifted from 5.20 GHz at 2 V to 5.17 GHz at 8 V.

The maximum notch depth in $S_{11}$ obtained was 23 dB for an applied bias voltage of 8 V. The variation of the ‘Q’ factor with frequency is shown in Fig. 3. With increasing bias voltage, the Q Factor was found to increase. The maximum ‘Q’ factor obtained was 216 at an applied bias of 8 V. This is comparable to the ‘Q’ factor reported for BST films deposited by pulsed laser ablation and RF sputtering. This shows that a simple and inexpensive MOD process can be used to implement tunable/switchable BST Resonators on GaN.
Fig. 4 shows the Smith chart plot of frequency response of the resonator at 8 V.

Fig. 4 shows the Smith chart plot of impedance response of the resonator in the frequency range of 5 GHz to 6 GHz for an applied DC bias of 8 V. The intersection of the impedance circle with the real axis gives the series and parallel resonance frequency of the resonator. The electromechanical coupling coefficient of the resonator is given by (3).

\[
K_t^2 = \frac{1}{2} \cdot \frac{L}{C} \cdot \frac{\tan \left( \frac{\pi}{2} \left( \frac{f_p - f_s}{f_p} \right) \right)}{\frac{1}{2} \left( \frac{f_p - f_s}{f_p} \right)}
\]  

(3)

Fig. 5 shows the variation of series resonance frequency \(f_s\) and parallel resonance frequency \(f_p\) with applied bias for BST film based resonators with FBARs on GaN/Sapphire substrates. With increase in bias voltage from 3V to 10V, parallel resonance frequency \(f_p\) was found to increase from 5.278GHz to 5.290GHz and the tunability is about 0.23%. Further increase in voltage resulted in decrease in parallel resonance frequency to 5.278GHz. The resonance frequency \(f_s\) remained almost constant at 5.01GHz for applied bias voltage from 3V to 10V.

The variation of electromechanical coefficient \(K_t^2\) with applied bias is shown in Fig. 6. The electromechanical coupling coefficient was found to increase from 0.120 to 0.131 with increase in bias voltage from 3V to 8V. The electromechanical coefficient \(K_t^2\) decreases to 0.0125 by applying bias of 10V for these resonators. In the case of BST resonators on silicon, the electromechanical coefficient was found to increase with applied bias [6]. The reason for decrease in electromechanical coefficient with applied voltage above 8 V for BST resonators on GaN/sapphire substrates is not clear at this point and will be communicated in future publications.

Fig. 5 shows the variation of series \((f_s)\) and parallel \((f_p)\) with applied bias for BST based resonators on GaN/sapphire.

Fig. 6 shows the variation of Electromechanical Coupling Coefficient \(K_t^2\) with applied bias for BST based resonators on GaN/Sapphire.

Fig. 7 and Fig. 8 show the real and imaginary impedance of the resonator at bias voltages from 1 to 10V for BST FBARS on GaN/Sapphire substrates.

Fig. 7 Real impedance of resonator at port 1 with variation in frequency for BST based resonators on GaN/Sapphire.
IV. CONCLUSIONS

We have fabricated a solidly mounted BST based tunable resonator on a GaN/Sapphire substrate with a tantalum oxide/silicon oxide acoustic impedance layer. The complete device was processed at a temperature of 800 °C. The reflection parameter S11 was reduced to -23 dB under an 8 V DC bias, which indicates strong voltage induced piezoelectric effect of the MOSD deposited BST film. The parallel resonant frequency of the resonator is found to increase with increasing DC bias. The electromechanical coupling coefficient at a bias voltage of 8 V is about 13.1% (0.0131) and the tunability of the resonator is about 0.6%. The quality factor Q of the resonator was found to be dependent on the applied bias voltage, with the maximum observed quality factor equal to 216. This Q value is comparable to those reported for sputter deposited BST films.

ACKNOWLEDGEMENT

We gratefully acknowledge the financial support through Phase II STTR Contract No. W911NF-13-C-0029b from the Department of Defense, US Army Research Office.

REFERENCES


The resonant transformation of acoustic waves in "Al/AlN/Mo/(100) diamond" piezoelectric structure has been studied by means of 2D FEM simulation. The nature of arising acoustical modes has been defined by the detail investigation of fields of elastic displacements. It was shown that the appearance of strongest spurious resonant peaks at frequencies above the bulk acoustic overtones is concerned with Lamb-like modes within diamond substrate, excited on the critical frequencies. In these points Lamb mode transforms into standing plate wave. Experimental and calculated data are in good accordance.

**Keywords** — diamond; AlN, HBAR; UHF; Lamb-like modes; 2D FEM simulation

**I. INTRODUCTION**

High-overtone Bulk Acoustic Resonator (HBAR) is a prospective acoustoelectronic device of sandwich piezoelectric structure operating up to microwave frequencies. A number of substrate materials as fused quartz and silicon [1], sapphire [1, 2], YAG [3], and diamond [4] have been already applied. HBAR based on synthetic diamond substrate has showed the excellent UHF resonant properties [4]. But one can observe the appearance of the additional resonant peaks which depends on the structure, dimensions and topology of thin film piezoelectric transducer (TFPT) as well as a diamond substrate thickness. As a result a deterioration of HBAR's fundamental resonant properties nearby the overtone's frequency takes place. The nature of such undesirable effect should be investigated for a more detail in order to design an improved HBAR at a given frequency band. Recall that such effect has been earlier observed for HBAR incorporating other crystalline substrates [2], and especially for FBAR and SMR [5 - 8].

The main objective of this paper is the FEM modeling and comparison with experimental data on resonant properties obtained for the diamond-based HBAR.

**II. HBAR PREPARATION**

As the substrates there were taken the thin plates of IIa synthetic diamond single crystal with [100] orientation grown by HPHT method at FSBI TESNCM (Troitsk, Russia). All the specimens were double-sized polished, faces were prepared with better than 5° deviation from [100] crystalline direction, and were polished to obtain the surface roughness $R_a \leq 10$ nm. The typical HBAR had sandwich structure shown on Fig. 1.

The lateral dimensions diamond substrates were close to $4 \times 4$ mm$^2$, but the lateral size of HBAR's active zone was varied from 100 up to 300 μm.

In order to control the depth of damage layer in diamond single crystal arising as a result of polishing, the observation of Kikuchi lines during experiments on Electron Back-Scattering Diffraction produced in Moscow Institute of Physics and Technology has resulted that damage layer depth for all studied diamond crystals is lower than 30 nm [9].

For experimental investigation the layered piezoelectric structure as "Al/AlN/Mo/(100) diamond" (164 nm/624 nm/169 nm/392 μm) has been applied. On the Fig. 2 one can see a map of top surface arrangement including 9 HBARs which differ by topology and sizes of top electrodes.
III. EXPERIMENT AND COMPUTER SIMULATION

HBAR’s microwave acoustic properties were tested by E5071C Network Analyzer (Agilent Tech.) and M-150 Multipurpose Probing System (CASCADE Microtech.). The initial experimental data were obtained by measurement of $S_{11}$ reflection coefficient as a function of frequency. Then such HBAR parameters as $Z_{11}$ and $Z_{11r}$ impedances were calculated. The “pure” impedance $|Z_{11r}|$ was obtained by the relation $|Z_{11r}| = |Z_{11}| - |Z'_1|$, where $|Z'_1|$ impedance was measured outside the acoustic resonance band. Taking $|Z_{11r}|$ frequency dependence one can obtain the quality factor $Q$ in a wide frequency range by the relation:

$$Q = \frac{f_r}{\Delta f},$$

where $f_r$ is resonant frequency, and bandwidth $\Delta f$ was measured at -3 dB level. For more details see [10].

Frequency dependences of $|Z_{11r}|$ impedance shown on Figs. 3-6 were selected for some resonant bands. Besides the bulk acoustic wave overtones, spurious resonant peaks are presented too. As one can see, the strongest spurious resonances were observed above the resonant frequencies of HBAR’s overtones.

Modeling calculation of frequency dependence of HBAR parameters was based on 2D FEM simulation. As an object of computer simulation the piezoelectric layered structure “Al/AlN/Mo/(100) diamond” (164 nm/624 nm/169 nm/392 μm) has been taken into account, because experimental data on resonant properties of HBAR based on this structure were available within a broad frequency band from 1 up to 9 GHz.

First, the ideal analogs of FBAR and HBAR, based on TFPT “Al/AlN/Mo” and “Al/AlN/Mo/(100) diamond” structure with the same layer thicknesses and differing the infinite lateral dimensions all the layers have been studied.

Second, the simulation of FBAR and HBAR, taking into account a number of top and bottom electrodes, and AlN with finite lateral dimensions corresponding to experimental ones has been executed. FBAR’s simulation was of secondary importance and had a purpose to clarify the peculiarities of resonant processes in HBAR.

The data on material constants and acoustical properties for Al, AlN, Mo, and diamond were taken from [10-12]. It was supposed that the acoustic attenuation for Al, AlN, and Mo obeyed to the Akhieser’s law: $\alpha \sim f^2$ at all the frequencies, but, starting from ~ 1 GHz the acoustic attenuation for single crystalline diamond should be changed as $\alpha \sim f$. The last is defined by the transition to Landau-Rumer regime as was shown in our paper [13].

The fields of elastic displacements as the functions of frequency have been analyzed by means of cross section of investigated specimens in a real 2D scale where the Y axis coincided with vertical direction, and X axis was directed along horizontal one. The conclusions about a type of elastic mode have been validated by the studying of the typical elastic displacement patterns which more clear image was observed at resonant frequencies. There were founded such modes as longitudinal (L) bulk acoustic wave (BAW), associated with HBAR’s overtones, and standing Lamb-like modes in lateral directions.

IV. RESULTS AND DISCUSSION

As the main origin of spurious satellites observed on amplitude frequency response of FBAR and HBAR, one can qualify the excitation of symmetric Lamb-like acoustic modes in the layered piezoelectric structure. Physical reason of such effect can be explain by the appearance of inhomogeneous mechanical and electrical boundary conditions as a consequence of finite lateral dimensions of electrodes and piezoelectric film.

It is well-known [14] that the normal acoustic waves are propagating in the lateral plate directions like the waveguide modes. So there are the points of critical frequencies and wave numbers starting from which the appearing of $m^{th}$ order Lamb-like mode will be take place. Lamb wave has a complex inhomogeneous structure including longitudinal and shear components. For example, the conditions of the occurrence for symmetric Lamb waves should be written as:

$$h = \frac{1}{2}\lambda_{L}^{(c)}, \frac{3}{2}\lambda_{S}^{(c)}, \frac{5}{2}\lambda_{L}^{(c)}, \ldots,$$

$$h = \lambda_{S}^{(c)}, 2\lambda_{S}^{(c)}, 3\lambda_{S}^{(c)}, \ldots,$$

where $h$ is the substrate thickness, $\lambda_{L}^{(c)}$ and $\lambda_{S}^{(c)}$ are the critical wave lengths for the longitudinal and shear components of Lamb wave, respectively. These values should be associated with bulk longitudinal and shear acoustic waves propagating along the Y axis, i.e. normally to the Lamb wave propagation direction. Note, that the relation (2) is exactly fulfilled only for the thin plate under the condition of free surfaces. So, the upper and bottom rows in the relation (2) are presented the standing longitudinal and shear plate waves, respectively. Taking into account relations (2), one can obtain the conditions on the critical frequencies:

$$f_L^{(c)} = \frac{V_L}{2h}; \frac{3V_L}{2h}; \frac{5V_L}{2h}; \ldots,$$

$$f_S^{(c)} = \frac{V_S}{h}; \frac{2V_S}{h}; \frac{3V_S}{h}; \ldots,$$

where $V_L$ and $V_S$ are the phase velocities of longitudinal and shear (S) bulk acoustic waves, respectively. Evidently, that the possibility of Lamb wave existence is given by the condition: $f \geq f^{(c)} (\lambda \leq \lambda^{(c)})$.

Fig. 3 represents experimental and calculated frequency dependences of $|Z_{11}|$ impedance of HBAR #3b “Al/AlN/Mo/(100) diamond” close to 1 GHz, as well as the visualization of elastic displacement fields nearby spurious satellite peak. Note that a good agreement between experimental and calculated data has been obtained: $Q_{exp} = 23000$, and $Q_{calc} = 22000$. Analyzing the Fig. 3a, b, one can see the longitudinal BAW overtone as well as the spurious satellite peak. The last has a complex structure (Fig 3c, d): there are two types of standing Lamb-like modes, one of which (S-type) is situated within the area under top electrode and other (L-type) is localized under the area of AlN thin film. Here one can easy to measure that within the thickness of a
substrate is stowed a number of half-lengths for \( L \)-type mode which is equal to \( m = 45 \), and for \( S \)-type mode this value is equal to \( n = 62 \). Now we have an interesting case of the degenerating of resonant frequency for these modes. Taking into account the relations (3), one can obtain the formula:

\[
mV_L = nV_S. \tag{4}
\]

Substituting the phase velocities \( V_L = 17542 \text{ m/s} \), and \( V_S = 12826 \text{ m/s} \) for \([100]\) single crystalline diamond into (4), one can demonstrate the correctness of this relation.

Drawing the attention on the top of Fig. 3d, one can see the spatial modulation of Lamb waves forming within the diamond substrate, by the Lamb-like waves propagating along the lateral directions within the AlN thin film. But these waves had the really small amplitudes and haven’t been found out in the experiment (Fig. 3a).

Fig. 4 represents experimental and calculated results on the HBAR \#3 close to 4.5 GHz. As one can see, a good agreement between experimental and calculated and calculated has been obtained: \( Q_{\text{exp}} \approx 10000 \), and \( Q_{\text{calc}} \approx 12300 \); distance between BAW overtone and spurious peak is equal to \( \Delta f_{\text{exp}} = 0.712 \), and \( \Delta f_{\text{calc}} = 0.73 \text{ MHz} \). Distribution of elastic displacements (Fig. 4c) corresponds to symmetrical Lamb standing wave with mainly vertical (\( L \)) displacements unlike to the case shown on Fig. 3c. Ultrasonic beams of Lamb standing wave are situated within the space under top electrode.

Fig. 5 represents experimental and calculated results on HBAR \#3c close to 8.0 GHz. Note, that quality factor remains a relatively high: \( Q_{\text{exp}} \approx 7500 \), and \( Q_{\text{calc}} \approx 9500 \). In a qualitative sense distribution of elastic displacements (Fig. 5c) corresponds to the case shown on Fig. 4c, but magnitude of displacement decreases significantly.

The analysis of spurious peaks close to 4.5 GHz obtained for HBARS with different lateral dimensions of top electrode showed that the shift of resonant frequencies of spurious peaks depends on the width of electrode as a function \( \Delta f \sim 1/w \), i.e. spurious peak will be closer to BAW overtone when the width will increase (Fig. 6). In this example all the types of spurious peaks are the same as Lamb standing wave associated with vertical polarization.

![Fig. 3.](image)

Fig. 3. Experimental (a) and calculated (b) frequency dependences of \( |Z_{11}| \) impedance of HBAR “Al/AlN/Mo/(100) diamond” close to 1 GHz; the total (c) and partial (d) patterns for the fields of elastic displacements nearby the spurious resonant peak (1.007 GHz). Visualization of elastic displacement directions is provided by arrows. In accordance with lateral dimensions of the HBAR \#3 (see Fig. 2) the width of top Al electrode was taken as 170 \( \mu \text{m} \), and width of bottom electrode is equal to 400 \( \mu \text{m} \) as well as AlN thin film. The thickness of diamond substrate is equal to 392 \( \mu \text{m} \).
Fig. 4. Experimental (a) and calculated (b) frequency dependences of $|Z_{11}|$ impedance of HBAR “Al/AlN/Mo/(100) diamond” close to 4.5 GHz; the total (c) and partial (d) patterns for the fields of elastic displacements nearby the spurious resonant peak (4.50225 GHz). In accordance with lateral dimensions of the HBAR #3 (see Fig. 2) the width of top Al electrode was taken as 250 μm, and width of bottom electrode is equal to 400 μm as well as AlN thin film. The thickness of diamond substrate is equal to 392 μm.

Fig. 5. Experimental (a) and calculated (b) frequency dependences of $|Z_{11}|$ impedance of HBAR “Al/AlN/Mo/(100) diamond” close to 8.0 GHz; the total (c) and partial (d) patterns for the fields of elastic displacements nearby the spurious resonant peak (8.0279 GHz). In accordance with lateral dimensions of the HBAR #3c (see Fig. 2) the width of top Al electrode was taken as 100 μm, and width of bottom electrode is equal to 400 μm as well as AlN thin film. The thickness of diamond substrate is equal to 392 μm.
Fig. 6. Calculated frequency dependences of $|Z_{11}|$ impedance of HBAR “Al/AlN/Mo/(100) diamond” as well as visualization of spurious peaks for a number of top electrode widths: (a) and (b) 170 μm; (c) and (d) 200 μm; (e) and (f) 250 μm; (g) and (h) 300 μm. The width of bottom electrode is equal to 400 μm as well as AlN thin film. The thickness of diamond substrate is equal to 392 μm.
V. SUMMARY

1. Only the pure extensional longitudinal mode exists under the excitation of ideal analogs of FBAR and HBAR.
2. The origin of spurious resonant peaks should be associated with finite lateral dimensions of thin films of HBAR’s structure.
3. Symmetrical Lamb-like waves excited within the AlN thin film were observed as a space modulation of HBAR’s resonant modes. Note, that the amplitudes of Lamb waves decreased considerably at higher frequencies. In comparison with FBAR the most part of such modes in the HBAR will be suppressed.
4. The arising of strongest spurious resonant peaks at frequencies above the bulk acoustic overtones is concerned with Lamb-like modes within diamond substrate, excited on the critical frequencies. In these points Lamb mode transforms into standing plate wave, which is spatially limited as a consequence of boundary conditions associated with mass loading and electrical shorting on the surface of piezoelectric film.
5. At higher frequencies up to 8 GHz the excitation of spurious peaks is almost completely suppressed, but the diamond-based HBAR’s quality factor remains a relatively high: $Q \sim 7500$.
6. It was shown that the resonant frequencies of spurious peaks depend on the width of electrode as a function $f_{spur} \sim 1/w$.

ACKNOWLEDGMENT

This work has been supported by the Ministry of Education and Science of the Russian Federation (№ 14.574.21.0074, unique ID project RFMEFI57414X0074) with the use of Shared-Use Equipment Center “Research of Nanostructured, Carbon and Superhard Materials” (FSBI TISNCM).

REFERENCES

An Analysis of Thickness-shear Vibrations of an Annular Plate with the Mindlin Plate Equations

Ji Wang, Hui Chen, Tingfeng Ma, Jianke Du, Lijun Yi
School of Mechanical Engineering & Mechanics
Ningbo University, Ningbo, Zhejiang 315211, CHINA
Email: wangji@nbu.edu.cn

Yook-Kong Yong
Department of Civil and Environmental Engineering,
Rutgers University, 623 Bowser Road,
NJ 08854, USA

Abstract—The Mindlin plate equations with the consideration of thickness-shear deformation as an independent variable have been used for the analysis of vibrations of quartz crystal resonators of both rectangular and circular types. The Mindlin or Lee plate theories that treat thickness-shear deformation as an independent higher-order vibration mode in a coupled system of two-dimensional variables are the choice of theory for analysis. For circular plates, we derived the Mindlin plate equations in a systematic manner as demonstrated by Mindlin and others and obtained the truncated two-dimensional equations of closely coupled modes in polar coordinates. We simplified the equations for vibration modes in the vicinity of fundamental thickness-shear frequency and validated the equations and method. To explore newer structures of quartz crystal resonators, we utilized the Mindlin plate equations for the analysis of annular plates with fixed inner and free outer edges for frequency spectra. The detailed analysis of vibrations of circular plates for the normalized frequency versus dimensional parameters provide references for optimal selection of parameters based on the principle of strong thickness-shear mode and minimal presence of other modes to enhance energy trapping through maintaining the strong and pure thickness-shear vibrations insensitive to some complication factors such as thermal and initial stresses.

Keywords—annular plate; vibration; frequency; shear; resonator

1. INTRODUCTION

High frequency vibrations of plates have been traditionally analyzed with Mindlin [1-5] and Lee plate theories [6-7], which are based on the expansion of displacements and electrical potential of the linear theory of piezoelectricity in power and trigonometric series. Mindlin plate theory, including effects of additional plate rigidity due to shear deformation and plate inertia due to rotations, is the foundation of analysis of quartz crystal resonators, and tremendous efforts have been made to improve the equations and obtain accurate solutions in past decades. For most quartz resonators, shapes of the crystal plates are either rectangular or circular, posing challenges in the analysis with such considerations of complications. The vibrations of rectangular crystal plates based on the Mindlin plate theory have been studied by earlier investigators in different aspects including the analysis of straight-crested waves. From earlier work on vibrations of rectangular crystal plates in Cartesian coordinates with the corrected Mindlin plate equations, we are familiar with the dispersion relations, frequency spectra, and mode shapes, which are playing important roles in the resonator design process. For circular plates, a theoretical analysis of vibrations is more mathematically challenging than rectangular plates and results are also limited.

The Mindlin and Lee plate equations with the consideration of thickness-shear deformation as independent variables have been used for the analysis of vibrations of quartz crystal resonators of both rectangular and circular types. The objective of such analysis is always on identifying optimal parameters of plates and resonator structures to enhance the strong and pure vibrations of the thickness-shear mode, which is the functioning mode of thickness-shear type resonators of both AT- and SC-cut quartz crystal. This goal can only be sufficiently achieved through accurate analysis of vibrations of coupled modes in a finite crystal plates with the consideration of closely clustered modes and complication factors such as electrodes and mounting supports which can affect the frequency and stability under loadings like impact and temperature. The Mindlin or Lee plate theories that treat thickness-shear deformation as an independent higher-order vibration mode in a coupled system of two-dimensional variables are the choice of theory for analysis [8-12].

For circular plates, we had derived the Mindlin plate equations in a systematic manners as demonstrated by Mindlin and others and obtained the truncated two-dimensional equations of closely coupled modes in a finite circular plate. Furthermore, we simplified the equations for modes in the vicinity of fundamental thickness-shear frequency and validated the equations and method from earlier studies [13-14]. In this paper, for the purpose of searching newer structures of quartz crystal resonators, we utilize the first-order Mindlin plate equations for the analysis of annular plates with clamped boundary conditions at the inner edge and free boundary conditions at the outer edge for frequency spectra with the finding that we can obtain similar spectra and vibration modes of a circular plate with free edges for isotropic materials. The radial thickness-shear, flexural, and the circumferential thickness-shear modes are considered in the analysis. The detailed analysis of vibrations of plates for the normalized frequency versus dimensional parameters in the vicinity of
fundamental thickness-shear frequency will provide guidelines for optimal selection of plate parameters based on the principle of strong thickness-shear mode and minimal presence of other modes to enhance energy trapping through maintaining the strong and pure thickness-shear vibrations insensitive to many complication factors such as thermal and initial stresses.

II. FIRST-ORDER PLATE EQUATIONS OF THICKNESS-SHEAR VIBRATIONS OF ANNULAR MINDLIN PLATES

We consider the isotropic, annular plate, for which the thickness coordinate is designated by \( z \), as shown in Fig. 1. Its inner radius, outer radius, thickness, and mass density are \( a_0 \), \( a \), \( 2b \) and \( \rho \), respectively. With the Mindlin plate equations in polar coordinates, we have the first-order stress equations of motion with coupling of the fundamental thickness-shear \( u_r^{(1)} \), flexural \( u_z^{(0)} \), and circumferential thickness-shear \( u_\theta^{(1)} \) [13-14]

\[
T_{\theta,\theta,d}^{(0)} + \frac{1}{r} T_{\theta,\theta}^{(1)} + \frac{T_{\theta}^{(0)}}{r} = 2\rho b \ddot{u}_\theta^{(0)},
\]

\[
T_{r,\theta,d}^{(0)} + \frac{1}{r} T_{\theta,\theta}^{(1)} + \frac{T_{\theta}^{(0)}}{r} - T_{\theta}^{(1)} = \frac{2b^3}{3} \rho \ddot{u}_r^{(1)},
\]

\[
T_{r,\theta,d}^{(0)} + \frac{1}{r} T_{\theta,\theta}^{(1)} + \frac{T_{\theta}^{(0)}}{r} - T_{\theta}^{(0)} = \frac{2b^3}{3} \rho \ddot{u}_\theta^{(0)},
\]

where \( T_{pq}^{(n)}(p=r,z,\theta; q=r,z,\theta; n=0,1) \) is the \( n \) th-order stress.

The constitutive equations for isotropic plates are

\[
T_{2,\theta,d}^{(0)} = 2b \mu k^2 \left( u_{\theta,d} + \frac{1}{r} u_{x,\theta} \right),
\]

\[
T_{2,\theta}^{(0)} = 2b \mu k^2 \left( u_{x,\theta} + u_\theta \right),
\]

\[
T_{1,\theta,d}^{(1)} = \frac{1}{r} \left( 2b^3 \mu \left[ \frac{1}{r} \left( u_{\theta,d} + u_r \right) + \nu u_{r,\theta} \right] \right),
\]

\[
T_{1,\theta}^{(1)} = \frac{1}{r} \left( 2b^3 \mu \left[ \frac{1}{r} \left( u_{\theta,d} + u_r \right) + \nu u_r \right] \right),
\]

where \( D = 2b^3 E / [3(1 - \nu^2)] \), and \( E \) and \( \nu \) are Young’s modulus and Poisson’s ratio, respectively, and \( k^2 = \pi^2 / 12 \). We can also use Lamé constants \( \lambda \) and \( \mu \) with these equations.

By inserting (2) into (1) and omitting the time factor \( \exp(i\omega t) \), the equations of motion for \( u_z^{(0)} \), \( u_r^{(1)} \), and \( u_\theta^{(1)} \) are

\[
k^2 \mu \left( \frac{\partial^2}{\partial r^2} + \frac{1}{r} \right) u_r^{(1)} + k^2 \mu u_\theta^{(1)} + \left[ k^2 \mu \left( \frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{1}{r} \frac{\partial}{\partial r} \right) + \rho \omega^2 \right] u_z^{(0)} = 0,
\]

\[
D \frac{\partial^2}{\partial r^2} + \frac{D}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{D(1 - \nu)}{r^2} \frac{\partial^2}{\partial \theta^2} - \left( \frac{D}{r^2} + 2k^2 \mu b - \frac{2}{3} \rho \omega^2 b^3 \right) u_r^{(1)} + \left[ \frac{D}{2r} \frac{\partial^2}{\partial \theta^2} - \frac{D(3 - \nu)}{2r} \frac{\partial}{\partial \theta} \right] u_\theta^{(0)} - 2k^2 \mu b \frac{\partial u_z^{(0)}}{\partial r} = 0,
\]

\[
D \frac{\partial^2}{\partial r^2} + \frac{D}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{D(1 + \nu)}{2r^2} \frac{\partial^2}{\partial \theta^2} + \frac{D(1 - \nu)}{2r} \frac{\partial}{\partial \theta} \right] u_\theta^{(1)} - \frac{2D(1 - \nu)}{r^2} + 2k^2 \mu b - \frac{2}{3} \rho \omega^2 b^3 \left[ \frac{\partial u_z^{(0)}}{\partial r} - 2k^2 \mu b \frac{\partial u_z^{(0)}}{\partial \theta} \right] = 0.
\]

Adding (3)_2 and (3)_3 together yields the governing equations of non-axisymmetric vibrations [10]

\[
\frac{\mu}{2\pi} \left[ \frac{\partial}{\partial \theta} u_r^{(1)} + \frac{3\lambda + 2\mu}{\lambda + 2\mu} q \left( \mathbf{v} \cdot u_r^{(1)} \right) \right] - k^2 \mu \left( u_r^{(1)} + \mathbf{v} \cdot u_r^{(0)} \right) = \frac{\rho b^3}{3} \ddot{u}_r^{(1)},
\]

\[
k^2 \mu \left( \mathbf{v} \cdot u_z^{(0)} + \mathbf{v} \cdot u_r^{(1)} \right) = \rho u_z^{(0)},
\]

where

\[
\mathbf{u}_r^{(1)} = u_r^{(1)} \mathbf{e}_r + u_\theta^{(1)} \mathbf{e}_\theta,
\]

\[
\mathbf{v} = \mathbf{e}_r \frac{\partial}{\partial r} + \mathbf{e}_\theta \frac{1}{r} \frac{\partial}{\partial \theta}.
\]

The boundary conditions for fixed inner and free outer edges are

\[
u_z^{(0)} = u_r^{(1)} = u_\theta^{(1)} = 0, \text{ at } r = a_0,
\]

\[
T_{rr}^{(0)} = T_{\theta\theta}^{(1)} = T_{\theta\theta}^{(0)} = 0, \text{ at } r = a.
\]

III. SOLUTIONS OF FREE VIBRATIONS

For non-axisymmetric vibrations of annular plates, we look for waves in the following form

\[
u_z^{(0)} = bA_1 J_1 (\delta r) \cos \theta + bB_1 Y_1 (\delta r) \cos \theta,
\]

\[
\mathbf{v} \cdot u_r^{(1)} = \frac{1}{b} A_2 J_1 (\delta r) \cos \theta + \frac{1}{b} B_2 Y_1 (\delta r) \cos \theta,
\]

where \( A_i(B_i) (i = 1, 2) \) and \( \delta \) are amplitudes and wavenumber, respectively. The normalized variables are utilized as

\[
X = \frac{\delta}{2b}, \Omega = \frac{\omega}{\omega_0}, \omega_0 = \frac{\pi}{2b} \sqrt{\frac{1}{\mu}},
\]

\[
\Delta = \frac{\mu}{2b}.\]
with \( \omega_0 \) as the fundamental thickness-shear vibration frequency.

The substitution of (7) and (8) into (4) yields

\[
3\pi^2\kappa^2X^2A_1 + \left[ \pi^2\Omega^2 - \frac{2\pi^2}{1-\nu^2}X^2 - 12\kappa^2 \right] A_2 = 0, \quad (9)
\]

and

\[
3\pi^2\kappa^2X^2B_1 + \left[ \pi^2\Omega^2 - \frac{2\pi^2}{1-\nu^2}X^2 - 12\kappa^2 \right] B_2 = 0, \quad (10)
\]

For nontrivial solution of \( A_1 \) and \( A_2 \), or \( B_1 \) and \( B_2 \), the determinant of the coefficient matrix of (9) and (10) must vanish, i.e.,

\[
\left| \begin{array}{cc}
3\pi^2\kappa^2X^2 & \pi^2\Omega^2 - \frac{2\pi^2}{1-\nu^2}X^2 - 12\kappa^2 \\
\pi^2(\Omega^2 - \kappa^2X^2) & 4\kappa^2 \\
\end{array} \right| = 0, \quad (11)
\]

Now we need to obtain the components of \( u_\theta^{(1)} \), which can be achieved by following the procedure developed by Lee et al. [10]. The components of \( u_\theta^{(1)} \) are

\[
u_\theta^{(1)} = \frac{4}{\pi^2(\Omega^2-1)} \left[ - \sum_{i=1}^{2} A_i \beta_i \delta_i b_i \left( \frac{\pi X_i r}{2 b} \right) + \frac{1}{r} \left( \frac{\beta_i}{r} \right) \right] \cos \theta 
\]

\[
+ \frac{4}{\pi^2(\Omega^2-1)} \left[ \sum_{i=1}^{2} B_i \beta_i \delta_i b_i Y_i \left( \frac{\pi X_i r}{2 b} \right) + \frac{1}{r} \left( \frac{\beta_i}{r} \right) \right] \cos \theta,
\]

\[
u_\theta^{(1)} = \frac{4}{\pi^2(\Omega^2-1)} \left[ \frac{1}{r} \sum_{i=1}^{2} A_i \beta_i l_i \left( \frac{\pi X_i r}{2 b} \right) - \frac{1}{r} \left( \frac{\beta_i}{r} \right) \right] \sin \theta 
\]

\[
+ \frac{4}{\pi^2(\Omega^2-1)} \left[ \frac{1}{r} \sum_{i=1}^{2} B_i \beta_i Y_i \left( \frac{\pi X_i r}{2 b} \right) - \frac{1}{r} \left( \frac{\beta_i}{r} \right) \right] \sin \theta,
\]

where

\[
\beta_i = \frac{8\alpha + 8\mu}{1+2\mu}, \quad \alpha_i = - \frac{4e^2}{\pi^2(\Omega^2-\kappa^2X^2)}, \quad i = 1, 2,
\]

\[
\delta_i^2 = \frac{\pi^2}{4be^2}, \quad \eta = \Omega^2 - 1.
\]

A substitution of stresses and displacements into the six boundary conditions at \( r = a_0 \), and \( r = \alpha \) in (6) yields to six linear equations for the six amplitudes \( A_i \) (\( B_i \)) (\( i = 1, 2, 3 \)). As usual, by setting the coefficient determinant of the boundary condition equations to vanish, frequency spectra and modes shapes can be calculated as the functions of parameters of the plate dimensions and material properties, as shown in equations of stresses and displacements. As usual, the frequency equation is given in transcendental functions and the solutions of roots or zero-points require accurate evaluation of Bessel functions. The algorithms for the numerical procedures are generally known and can be implemented with mathematical tools.

For clamped edge conditions at the inner \( (\alpha_0 = 2b) \) and free edge conditions at the outer boundaries of the annular plate, we obtained frequency spectra similar to a free circular plate as shown in Fig. 2. In order to identify the frequency features at the specific frequency at which the three vibration modes are coupled, we calculated mode shapes at various locations of the spectra in Fig. 2 labeled with A, B, and C. Figs. 3-5 shows the distribution of relative displacements at frequencies corresponding to points A, B, and C. The results show that the vibration modes depend on the frequency strongly. At point A, there is a strong coupling between all the vibration modes and the flexure \( u_\theta^{(1)} \) is strong. At point B, we can guess that it is the dominant thickness-shear mode \( u_\theta^{(1)} \), which has relatively large amplitude. Similarly, point C is another location with strong couplings of three modes and the transverse thickness-shear mode \( u_\theta^{(1)} \) is relatively strong. For resonator design, we always want to find the strong radial thickness-shear mode with less coupling to the flexural mode to achieve better resonance with less energy dissipation. The frequency spectra in Fig. 2 and the mode shapes in Figs. 3-5 are the needed information to find the right parameters and modes.

![Fig. 3. Distribution of displacements at frequency corresponding to point A.](image-url)
IV. CONCLUSIONS

We had utilized the first-order Mindlin plate equations for the analysis of the fundamental thickness-shear vibrations of an isotropic, annular plate, and successfully solved the coupled equations with the Bessel functions in a manner similar to the trigonometric functions in the analysis of vibrations of rectangular plates. With these validated equations and solutions experiences, we obtained frequency spectra and modes shapes of an annular plate with clamped boundary conditions at the inner edge and free boundary conditions at the outer edge. The results showed that there are frequencies reflecting dominant vibration modes and their strength of couplings. We can use the spectra to find ideal regions with the dominant thickness-shear vibrations for possible designs of quartz crystal resonators with different structures. The procedure outlined in this study can be used for the further search of novel configuration of quartz crystal resonators with new materials, structures, and features.

REFERENCES

Thicknees shear Vibration Frequencies of an Infinite Plate with a Generalized Material Property Grading along the Thickness

Ji Wang*, Wenliang Zhang, Dejin Huang, Tingfeng Ma, Jianke Du, Lijun Yi
Piezoelectric Device Laboratory, School of Mechanical Engineering & Mechanics, Ningbo University,
818 Fenghua Road, Ningbo, Zhejiang 315211, CHINA

*E-mail: wangji@nbu.edu.cn

Abstract—For quartz crystal resonators of thickness-shear type, the vibration frequency and mode shapes, which are key features of resonators in circuit applications, reflect the basic material and structural properties of the quartz plate and its variation with time under various factors such as erosive gases and liquids that can cause surface and internal damages and degradation of crystal blanks. The accumulated effects eventually will change the surface conditions in terms of elastic constants and stiffness and more importantly, the gradient of such properties along the thickness. This is a typical functionally graded materials (FGM) structure and has been studied extensively for structural applications under multiple loadings such as thermal and electromagnetic fields in recent years. For acoustic wave resonators, such studies are equally important and the wave propagation in FGM structures can be used in the evaluation and assessment of performance, reliability, and life of sensors based on acoustic waves such as the quartz crystal microbalances (QCM). Now we studied the thickness-shear vibrations of FGM plates with properties of AT-cut quartz crystal varying along the thickness in a general pattern represented by a trigonometric function with both sine and cosine functions of the thickness coordinate. The solutions are obtained by using Fourier expansion of the plate deformation. We also obtained the frequency changes of the fundamental and overtone modes which are strongly coupled for the evaluation of resonator structures with property variation or design to take advantages of FGM in novel applications.

Keywords—thickness-shear; frequency; vibration; functionally graded materials; resonator

I. INTRODUCTION

In the analysis and design of quartz crystal resonators, it has been known that the material is homogenous and uniform to simplify the design process. In reality, as one way to improve the performance of quartz crystal resonators, variation of material configuration has been adopted to change the properties of vibration modes to satisfy needs of resonator performance. The variation of quartz crystal plate configuration presents challenges to the analysis and fabrication of quartz crystal resonators due to complicated material properties, but the needs are justified from the significant improvement of device performances and efforts have been made through design and fabrication. It is known that the variation will reduce the couplings of typical vibration modes such as the thickness-shear (TSh) and flexure which appear in the functioning vibrations of quartz crystal resonators. The difficulties for accurate analysis of anisotropic plate vibrations are generally known with complications from anisotropic material, coupled modes, and accompanying boundary conditions. These complications have been there and many techniques primarily based on various simplifications and approximations have been existed for solutions which are required in resonator design. Further refinement and improvement of analysis are made with the consideration of more physical complications like irregular configurations of crystal blanks such as the commonly known beveling, which refers to the contour or thickness variations in edges of blanks to suppress strong couplings between thickness-shear and flexural modes. It has been found that the beveling in common configurations such as linear or quadratic variation of thickness can weaken couplings between thickness-shear and flexural modes significantly [1-2]. As a result, processing techniques based on grinding have been developed to take the advantage by making contour or beveling in edges of crystal blanks, but not necessarily in any specific form of design [3]. Since material grading is equivalent to thickness variation, we started the analysis of vibration frequencies of an FGM quartz crystal plate in the thickness-shear modes to explore possible applications in sensor technology [4].

There have been studies on wave propagations in FGM solids with various methods including approximations and discrete techniques [4-6]. There is no doubt that such analysis is important in establishing the correlation between performance changes and FGM patterns in wave propagation and high frequency vibrations of device structures. In the material processing part, FGM patterns can be realized through modern technologies such as radiation, laser, and chemical etching, and so on. We can make the partial materials to FGM materials so advantages can be taken and the equivalent effect of beveling can be achieved with minimal cost in fabrication. In this case, we have finally utilized the FGM for possible advantages rather than known applications in structural protection and enhancement [6]. Of course, applications of FGM in acoustic wave devices have been studied before for possible performance enhancement and novel fabrication techniques [7-9]. The essential nature of acoustic wave devices requires the analysis to consider wave propagations or high frequency vibrations in piezoelectric materials and solids as have been done for FGM plates and structures [7-16]. Such studies have been carried out before and there are positive leads to be followed up with more research for effective FGM to meet performance requirements of next generation acoustic wave devices. In this paper, we make further assumption on the material grading to represent a general pattern which is no longer symmetric regarding the middle plane, or the thickness. By following the earlier
approach based on the Fourier series expansion of the thickness-shear displacement which now has both symmetric and antisymmetric components, a more general example of vibrations of an infinite FGM plate is studied. The frequency equation of the FGM plate in the thickness-shear vibration mode is obtained for studies on effects of patterns and parameters of material grading.

II. PROBLEM AND FORMULATION

We start with an infinite plate of quartz crystal with modified material properties resembling to FGM. The plate is shown in Fig.1, and the thickness of plate is 2b.

![An infinite quartz crystal plate](image)

For a simplest problem of thickness-shear waves propagating in the elastic plate, the stress equation of motion is

\[
(c_{66}u_{1,2})_{,2} = \rho(x_2)\ddot{u}_1,
\]

(1)

where \(c_{66} \), \(\rho \), and \(u \) are elastic constant, density, and displacement of plate, respectively.

In case of \(c_{66} \) and \(\rho \) are constants, the simple solution is known as

\[
u = \begin{cases} 
A \sin \frac{r \pi x_2}{2b} e^{i\omega t}, & r \text{ odd,} \\
B \cos \frac{r \pi x_2}{2b} e^{i\omega t}, & r \text{ even,}
\end{cases}
\]

(2)

\[
\omega = \frac{n \pi}{2b} \sqrt{c_{66}/\rho}, \quad n = 1, 2, \cdots,
\]

(3)

where \(A(B)\) and \(\omega\) are amplitudes and vibration frequency, respectively.

In this study, we assume that quartz crystal plate has been treated and constants \(c_{66}\) and \(\rho\) are now functions of the thickness coordinate \(x_2\) in the form of

\[
(c_{66}\rho) = (C_{66}\rho_0) \theta(x_2), f(x_2) = \alpha + \beta_1 \cos \frac{x_2}{Nb} + \beta_2 \sin \frac{x_2}{Nb}
\]

(4)

where constants \(C_{66} \), \(\rho_0 \), \(\alpha \), \(\beta_1 \), and \(\beta_2 \) are known. Apparently, if integer \(N\) is large enough, we get material properties approach to uniform. With appropriate integer \(N\), we can study the effect of material property grading on the vibration frequency.

With the consideration of traction-free surfaces of the plate in Fig. 1, we can assume the thickness displacement as

\[
u = \left( \sum_{n=1,3,5,\cdots}^{\infty} A_n \sin \frac{n\pi}{2b} x_2 + \sum_{m=0,2,4,\cdots}^{\infty} B_m \cos \frac{m\pi}{2b} x_2 \right) e^{i\omega t},
\]

(6)

Then through the expansion of (1) we have

\[
c_{66,2}u_{,2} + c_{66}u_{,22} = \rho(x_2)\ddot{u},
\]

(7)

and by substituting (6) into (7) we obtain

\[
-C_{66} \left( \frac{\beta_1}{Nb} \sin \frac{x_2}{Nb} + \frac{\beta_2}{Nb} \cos \frac{x_2}{Nb} \right) \sum_{n=1,3,5,\cdots}^{\infty} \frac{n\pi}{2b} A_n \cos \frac{n\pi}{2b} x_2
\]

\[
-\frac{m\pi}{2b} B_m \sin \frac{m\pi}{2b} x_2
\]

\[
-C_{66} \left( \alpha + \beta_1 \cos \frac{x_2}{Nb} + \beta_2 \sin \frac{x_2}{Nb} \right) \sum_{n=1,3,5,\cdots}^{\infty} \frac{n\pi}{2b} A_n \sin \frac{n\pi}{2b} x_2
\]

\[
+\rho_0 \omega^2 \left( \alpha + \beta_1 \cos \frac{x_2}{Nb} + \beta_2 \sin \frac{x_2}{Nb} \right) \sum_{n=1,3,5,\cdots}^{\infty} \frac{n\pi}{2b} A_n \sin \frac{n\pi}{2b} x_2
\]

\[
-\frac{m\pi}{2b} B_m \sin \frac{m\pi}{2b} x_2
\]

\[
+\rho_0 \omega^2 \left( \alpha + \beta_1 \cos \frac{x_2}{Nb} + \beta_2 \sin \frac{x_2}{Nb} \right) \sum_{m=0,2,4,\cdots}^{\infty} \frac{m\pi}{2b} B_m \cos \frac{m\pi}{2b} x_2
\]

\[
+\rho_0 \omega^2 \left( \alpha + \beta_1 \cos \frac{x_2}{Nb} + \beta_2 \sin \frac{x_2}{Nb} \right) \sum_{m=0,2,4,\cdots}^{\infty} \frac{m\pi}{2b} B_m \sin \frac{m\pi}{2b} x_2 = 0.
\]

(8)

With \(f(x_2) = \alpha + \beta_1 \cos \frac{x_2}{Nb} + \beta_2 \sin \frac{x_2}{Nb}\) we also use

\[
f'(x_2) \cos \frac{m\pi}{2b} x_2 = f(x_2) \frac{f'(x_2)}{f(x_2)} \cos \frac{m\pi}{2b} x_2,
\]

(9)

\[
f'(x_2) \sin \frac{m\pi}{2b} x_2 = f(x_2) \frac{f'(x_2)}{f(x_2)} \sin \frac{m\pi}{2b} x_2.
\]

(10)

This actually is to obtain the Fourier expansion of functions in the form of

\[
f'(x_2) \cos \frac{m\pi}{2b} x_2 \sim \sum_{p=1,3,5,\cdots}^{\infty} A_{np} \sin \frac{p\pi}{2b} x_2 + \sum_{q=0,2,4,\cdots}^{\infty} B_{np} \cos \frac{q\pi}{2b} x_2,
\]

\[
\frac{f'(x_2)}{f(x_2)} \sin \frac{m\pi}{2b} x_2 \sim \sum_{r=1,3,5,\cdots}^{\infty} A_{mr} \sin \frac{r\pi}{2b} x_2 + \sum_{s=0,2,4,\cdots}^{\infty} B_{mr} \cos \frac{s\pi}{2b} x_2.
\]

(11)
where \( A'_{np}, B'_{nq}, A''_{mr}, B''_{ms} \) are coefficients in the form of

\[
A'_{np} = \frac{1}{b} \int_{-b}^{b} f'(x_2) \cos \frac{m \pi}{b} x_2 \sin \frac{p \pi}{b} x_2 \, dx_2, \quad p = 1, 3, 5, \ldots,
\]

\[
B'_{nq} = \frac{1}{2b} \int_{-b}^{b} f'(x_2) \cos \frac{m \pi}{b} x_2 \cos \frac{p \pi}{b} x_2 \, dx_2,
\]

\[
A''_{mr} = \frac{1}{b} \int_{-b}^{b} f''(x_2) \sin \frac{m \pi}{b} x_2 \sin \frac{r \pi}{b} x_2 \, dx_2, \quad r = 1, 3, 5, \ldots,
\]

\[
B''_{ms} = \frac{1}{b} \int_{-b}^{b} f''(x_2) \sin \frac{m \pi}{b} x_2 \cos \frac{s \pi}{b} x_2 \, dx_2, \quad s = 2, 4, \ldots.
\]

Clearly,

\[
A'_{np} = B''_{nq}.
\]  

(13)

Different notations are used for the convenience of writing and programming.

By substituting (11), (12) into (8), we have

\[
\sum_{n=1,3,5,\ldots}^{\infty} f_n(A, B) \sin \frac{n \pi x_2}{b} + \sum_{m=0,2,4,\ldots}^{\infty} g_m(A, B) \cos \frac{m \pi x_2}{b} = 0,
\]

(14)

where

\[
g_0(A, B) = -\frac{\pi C_{66}}{2b} (B'_{10} A_1 + 3B'_{30} A_3 + \cdots + nB'_{n0} A_n)
\]

\[
+ \frac{\pi C_{66}}{2b} (2B'_{20} B_2 + 4B'_{40} B_4 + \cdots
\]

\[
+ mB'_{m0} B_m) - \rho_0 \omega^2 B_0
\]

\[
f_1(A, B) = -\frac{\pi C_{66}}{2b} (A'_{11} A_1 + 3A'_{31} A_3 + \cdots + nA'_{n1} A_n)
\]

\[
+ \frac{\pi C_{66}}{2b} (2A'_{12} B_2 + 4A'_{32} B_4 + \cdots
\]

\[
+ mA'_{m1} B_m)
\]

\[
g_2(A, B) = -\frac{\pi C_{66}}{2b} (B'_{12} B_1 + 3B'_{32} B_3 + \cdots + nB'_{n2} B_n)
\]

\[
+ \frac{\pi C_{66}}{2b} (2B'_{22} B_2 + 4B'_{42} B_4 + \cdots
\]

\[
+ mB'_{m2} B_m)
\]

\[
+ 2^2 C_{66} \left( \frac{\pi}{2b} \right)^2 - \rho_0 \omega^2 \right) B_2,
\]

\[
f_3(A, B) = -\frac{\pi C_{66}}{2b} (A'_{13} A_1 + 3A'_{33} A_3 + \cdots + nA'_{n3} A_n)
\]

\[
+ \frac{\pi C_{66}}{2b} (2A'_{13} B_2 + 4A'_{33} B_4 + \cdots
\]

\[
+ mA'_{m3} B_m)
\]

\[
+ 3^2 \left( \frac{\pi}{2b} \right)^2 C_{66} - \rho_0 \omega^2 \right) A_3,
\]

\[
g_4(A, B) = -\frac{\pi C_{66}}{2b} (B'_{14} A_1 + 3B'_{34} A_3 + \cdots + nB'_{n4} A_n)
\]

\[
+ \frac{\pi C_{66}}{2b} (2B'_{24} B_2 + 4B'_{44} B_4 + \cdots
\]

\[
+ mB'_{m4} B_m) + \left[ 4^2 C_{66} \left( \frac{\pi}{2b} \right)^2 - \rho_0 \omega^2 \right] B_4,
\]

\[
\vdots
\]

\[
f_n(A, B) = -\frac{\pi C_{66}}{2b} (A'_{1n} A_1 + 3A'_{3n} A_3 + \cdots + nA'_{n} A_n)
\]

\[
+ \frac{\pi C_{66}}{2b} (2A'_{2n} B_2 + 4A'_{4n} B_4 + \cdots
\]

\[
+ mA'_{m} B_m) + \left[ n^2 \left( \frac{\pi}{2b} \right)^2 C_{66} - \rho_0 \omega^2 \right] A_n,
\]

\[
g_m(A, B) = -\frac{\pi C_{66}}{2b} (B'_{1m} A_1 + 3B'_{3m} A_3 + \cdots + nB'_{nm} A_n)
\]

\[
+ \frac{\pi C_{66}}{2b} (2B'_{2m} B_2 + 4B'_{4m} B_4 + \cdots
\]

\[
+ mB'_{mm} B_m) + \left[ m^2 C_{66} \left( \frac{\pi}{2b} \right)^2 - \rho_0 \omega^2 \right] B_m.
\]

(15)

For (14) to be true, we must have the coefficients vanish separately, i.e.

\[
f_n(A, B) = 0, \quad g_m(A, B) = 0, \quad n = 1, 3, 5, \ldots; m = 0, 2, 4, \ldots.
\]

(16)

With the normalized parameters

\[
\Omega = \frac{\omega}{\omega_1}, \quad \Omega^2 = \frac{\omega^2}{\omega_1^2} = \rho_0 \omega^2 \frac{4b^2}{\pi^2 C_{66}},
\]

where \( \omega_1 \) is the fundamental thickness-shear frequency of a plate with uniform properties given in (3).

Now we have a system of equations for amplitudes of displacements

\[
K \cdot S = 0,
\]

(18)

where \( K \) and \( S \) are matrices as

\[
K = \begin{bmatrix}
-\Omega^2 & -\frac{2b}{\pi} B'_{10} & \frac{2b}{\pi} B_{20} & \cdots & -\frac{2b}{\pi} nB'_{n0} & \frac{2b}{\pi} mB_{m0} \\
0 & -\frac{2b}{\pi} B'_{11} & \frac{2b}{\pi} A_{21} & \cdots & -\frac{2b}{\pi} nA_{n1} & \frac{2b}{\pi} mA_{m1} \\
0 & -\frac{2b}{\pi} B'_{12} & \frac{2b}{\pi} B_{22} & \cdots & -\frac{2b}{\pi} nB'_{n2} & \frac{2b}{\pi} mB_{m2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & -\frac{2b}{\pi} A'_{1m} & \frac{2b}{\pi} A_{2m} & \cdots & -\frac{2b}{\pi} nA'_{nm} & \frac{2b}{\pi} mA_{mm} \\
0 & -\frac{2b}{\pi} B'_{1m} & \frac{2b}{\pi} B_{2m} & \cdots & -\frac{2b}{\pi} nB'_{nm} & \frac{2b}{\pi} mB_{mm} \\
0 & -\frac{2b}{\pi} B'_{1n} & \frac{2b}{\pi} B_{2n} & \cdots & -\frac{2b}{\pi} nB'_{mn} & \frac{2b}{\pi} mB_{mn} \\
0 & -\frac{2b}{\pi} B'_{1m} & \frac{2b}{\pi} B_{2m} & \cdots & -\frac{2b}{\pi} nB'_{nm} & \frac{2b}{\pi} mB_{mm} \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots \\
\end{bmatrix}
\]

(19)
For free vibrations of an infinite FGM quartz crystal plate, frequencies can be obtained by setting the determinant of $K$ matrix to vanish through

$$|K| = 0.$$  

With all known parameters of material properties, we can evaluate (21) for the fundamental and overtone vibration frequencies. It is clear from (19) that the fundamental and overtone modes are now closely coupled. In other words, the FGM plates can no longer support pure modes of thickness-shear vibrations even it is infinite. In this case, our interests will be on the effects of FGM patterns on frequencies of the coupled thickness-shear modes.

### III. NUMERICAL RESULTS

With the frequency equation in (21), we can calculate vibration frequencies of an FGM plate with given property grading pattern and parameters. Since the frequency equation cannot be given in an explicit form, we have to perform the calculation with estimated number of modes, or order of the matrix, to obtain convergent solutions with given limit for the evaluation of frequency. It is clear that the fundamental and overtone modes are now closely coupled. In other words, the FGM plates can no longer support pure modes of thickness-shear vibrations even it is infinite. In this case, our interests will be on the effects of FGM patterns on frequencies of the coupled thickness-shear modes.

#### TABLE I. VIBRATION FREQUENCIES WITH THE ORDER OF DETERMINANT = 3

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.003534</td>
<td>1.000882</td>
<td>1.000392</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001288</td>
<td>2.00032</td>
<td>2.000142</td>
<td>2.00008</td>
<td>2.000051</td>
</tr>
</tbody>
</table>

#### TABLE II. VIBRATION FREQUENCIES WITH THE ORDER OF DETERMINANT = 5

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.003555</td>
<td>1.000886</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001792</td>
<td>2.000444</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001197</td>
<td>3.000296</td>
<td>3.000131</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000892</td>
<td>4.000222</td>
<td>4.000098</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000717</td>
<td>5.000176</td>
<td>5.000078</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

#### TABLE III. VIBRATION FREQUENCIES WITH THE ORDER OF DETERMINANT = 7

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.003559</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001788</td>
<td>2.000443</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001195</td>
<td>3.000296</td>
<td>3.000131</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000897</td>
<td>4.000222</td>
<td>4.000098</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000716</td>
<td>5.000177</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

#### TABLE IV. VIBRATION FREQUENCIES WITH THE ORDER OF DETERMINANT = 9

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00356</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001792</td>
<td>2.000444</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001197</td>
<td>3.000296</td>
<td>3.000131</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000892</td>
<td>4.000222</td>
<td>4.000098</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000717</td>
<td>5.000177</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

#### TABLE V. VIBRATION FREQUENCIES WITH THE ORDER OF DETERMINANT = 11

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00356</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001792</td>
<td>2.000444</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001197</td>
<td>3.000296</td>
<td>3.000131</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000897</td>
<td>4.000222</td>
<td>4.000098</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000716</td>
<td>5.000177</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

#### TABLE VI. VIBRATION FREQUENCIES WITH THE ORDER OF DETERMINANT = 13

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00356</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001792</td>
<td>2.000444</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001197</td>
<td>3.000296</td>
<td>3.000131</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000897</td>
<td>4.000222</td>
<td>4.000098</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000716</td>
<td>5.000177</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

#### TABLE VII. VIBRATION FREQUENCIES WITH THE ORDER OF DETERMINANT = 15

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.003561</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001794</td>
<td>2.000445</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001197</td>
<td>3.000296</td>
<td>3.000131</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000898</td>
<td>4.000222</td>
<td>4.000098</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000718</td>
<td>5.000178</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>
From tables above, we found that for specific $N=5, 10, 15, 20, 25$ tested, we can always obtain accurate solutions for the first five modes with $K$ of order 17. The results based on this calculation have been summarized in Table XI. Again, it is clear that the overtone frequencies are no longer multiples of the fundamental frequency. With $N$ takes a large value, we obtain the exact solutions of uniform plates.

### Table VIII. Vibration Frequencies with the Order of Determinant = 17

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.003561</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001795</td>
<td>2.000445</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001198</td>
<td>3.000296</td>
<td>3.000132</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000899</td>
<td>4.000222</td>
<td>4.000099</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000719</td>
<td>5.000178</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

### Table IX. Vibration Frequencies with the Order of Determinant = 19

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.003561</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001795</td>
<td>2.000445</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001198</td>
<td>3.000296</td>
<td>3.000132</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000899</td>
<td>4.000222</td>
<td>4.000099</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000719</td>
<td>5.000178</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

### Table X. Vibration Frequencies with the Order of Determinant = 21

<table>
<thead>
<tr>
<th>Mode no.</th>
<th>$N=5$</th>
<th>$N=10$</th>
<th>$N=15$</th>
<th>$N=20$</th>
<th>$N=25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.003561</td>
<td>1.000887</td>
<td>1.000394</td>
<td>1.000222</td>
<td>1.000142</td>
</tr>
<tr>
<td>2</td>
<td>2.001795</td>
<td>2.000445</td>
<td>2.000197</td>
<td>2.000111</td>
<td>2.000071</td>
</tr>
<tr>
<td>3</td>
<td>3.001198</td>
<td>3.000296</td>
<td>3.000132</td>
<td>3.000074</td>
<td>3.000047</td>
</tr>
<tr>
<td>4</td>
<td>4.000899</td>
<td>4.000222</td>
<td>4.000099</td>
<td>4.000055</td>
<td>4.000035</td>
</tr>
<tr>
<td>5</td>
<td>5.000719</td>
<td>5.000178</td>
<td>5.000079</td>
<td>5.000044</td>
<td>5.000028</td>
</tr>
</tbody>
</table>

As the continuation of our study on the thickness-shear vibrations of quartz crystal plates with material grading across the thickness, we extended our earlier results to more general variation of material properties represented by a trigonometric function with both symmetric and anti-symmetric elements included. The solutions are obtained by expanding the deformation into Fourier series and through forcing the coefficients of coupled terms to vanish in order to satisfy the equation of motion. Then, the frequencies from the coefficient matrix are obtained as the vibration frequencies of the thickness-shear modes including the fundamental one and the overtones. In numerical examples, only the asymmetric vibrations are obtained and the procedure of calculation is examined. The complications of vibration modes through the couplings of overtone modes are observed. Such results can be used in the analysis of frequency spectra of resonators with both surface and internal damages and corrosions to help the understanding of certain measurement results in applications like sensors for chemical and biological samples. Of course, it is also possible to create material variation patterns of quartz crystal resonators to enhance mode couplings so measurements can be performed with different frequencies for validations or enable multiple functions with one device.

**ACKNOWLEDGMENT**

This research is supported by the National Natural Science Foundation of China through grants 11372145 and 11372146.

**REFERENCES**


Wideband Ladder Filters Fully Covering Digital TV Band based on Shear Horizontal Plate Wave

Michio Kadota and Shuji Tanaka
Graduate School of Engineering, Tohoku University, Sendai, Miyagi, Japan
E-mail: mkadota@mems.mech.tohoku.ac.jp

Abstract— A cognitive radio terminal using vacant frequency bands of digital TV channels, i.e., TV white space, strongly requires a compact tunable filter covering a wide frequency range of the digital TV band. The authors proposed a new bandwidth- and frequency-tunable filter combining an ultra-wideband ladder filter and tunable band rejection filters. T-type and π-type ladder filters composed of ultra-wideband 0th shear horizontal mode (SH₀) plate wave resonators were fabricated on a submicron thick LiNbO₃ plate supported by a Si substrate. The wavelength ratio of the parallel/series resonators in the ladder filters was changed as a design parameter. As a result, ultra-large 6 dB bandwidths of 41 to 51%, which fully cover the digital TV band, were measured. The lowest peak insertion loss was low as 0.8 dB. The ultra-wideband ladder filter is useful for the above-mentioned new tunable filter for digital TV cognitive radio communication.

Keywords—ultra-wideband ladder filter; SH₀ plate wave, thin LiNbO₃ plate, cognitive radio, digital TV

I. INTRODUCTION

Currently, the wide spread of smartphones and other mobile terminals has led to the depletion of available frequency spectra. To address this problem, cognitive radio technology using a vacant frequency band (white space) of digital TV (DTV) channels is receiving a lot of attention and standardized as IEEE 802.11af [1][2]. Fig. 1 shows DTV channel structure in Japan, which is from 470 MHz to 710 MHz and composed of 40 channels with 6 MHz bandwidth (BW) for each. However, only a limited number of channels are actually used, depending on areas. For instance, only 12 channels colored in red and pink in Fig. 1 are used among 40 channels in Sendai, Japan and a neighbor city, while the other 28 channels are vacant, i.e., white space. Cognitive radio uses a single white space channel, two consecutive white space channels or four consecutive white space channels, which offers WiFi communication with a BW of 5 MHz, 11 MHz or 23 MHz, respectively. Therefore, a key device in a cognitive radio terminal is the tunable filter which can largely change the center frequency and BW to select an available TV white space.

The authors previously developed a monolithic BW-tunable surface acoustic wave (SAW) filter with BST varactors [3][4]. The tuning function was realized by connecting the varactors with SAW resonators in parallel or series. In this configuration, the tuning range is limited within the BW of the original resonator. To obtain a wider tuning range, a resonator with a larger BW is required. The authors fabricated ultra-wideband (22% BW) resonators in DTV band using 0th shear horizontal mode (SH₀) plate wave in a (0°, 120°, 0°) LiNbO₃ plate with large electromechanical coupling factor [5]. The potential frequency tuning range was estimated at ten-odd % by connecting different capacitors with the resonators [6]. This is almost the largest tuning range achievable by this tunable filter configuration, but not enough for DTV cognitive radio application.

The authors proposed another type of tunable filter, combining an ultra-wideband filter fully covering the DTV band and tunable band rejection filters [7][8]. The ultra-wide ladder filter and band rejection filters are numerically synthesized based on the measured frequency characteristics of the resonators, and the tunable filter was simulated [7]. A wider tuning range can be realized by restricting a wide passband using band rejection filters. For TV white space cognitive radio application, therefore, a bandpass filter fully covering the DTV band is primarily needed. In this study, T-type and π-type ladder filters using SH₀ plate wave were fabricated. The measured BW is as large as 41% to 51%, which is the largest value ever reported and fully covers the DTV band.

II. SH₀ PLATE WAVE RESONATOR

A large BW of the resonator is obtained using a piezoelectric substrate or film with a large electro-mechanical coupling factor (referred as coupling factor hereafter, k²). The largest effective coupling factor of a Love wave type of SAW is as large as 38%, but not yet enough to cover the DTV band [9][10]. To obtain a larger coupling factor, the authors studied SH₀ plate wave on a (0°, θ, 0°) LiNbO₃ plate [6]. This section reviews our recent results. Fig. 2 shows the coupling factors of SH₀ plate wave in a (0°, θ, 0°) LiNbO₃ plate of 0.12 thickness as a function of θ. The coupling factor larger than 50% is significant, indicating the potential for wide BW.
obtained with \( (0^\circ, 117.5-120^\circ, 0^\circ) \) LiNbO\(_3\) at a plate thickness of \( 0.1\lambda \). Fig. 3 shows the coupling factor of SH\(_0\) plate wave in \( (0^\circ, 120^\circ, 0^\circ) \) LiNbO\(_3\) as a function of the plate thickness. At a plate thickness close to 0.03\( \lambda \), the coupling factor reaches 55\%, which is the largest even compared with other plate waves, i.e. 0th symmetric mode (\( S_0 \)), 0th anti-symmetric mode (\( A_0 \)) and 1st anti-symmetric mode (\( A_1 \)) Lamb waves [6]. From a coupling factor point of view, thinner LiNbO\(_3\) is better, but there is a practical lower limit in the plate thickness due to mechanical strength and thickness uniformity.

Fig. 4 shows the frequency characteristics of two kinds of SH\(_0\) plate wave resonators in \( (0^\circ, 120^\circ, 0^\circ) \) LiNbO\(_3\) plate. An ultra-wide BW of 22\% is confirmed. They were fabricated on a \( (0^\circ, 120^\circ, 0^\circ) \) LiNbO\(_3\) plate of 0.5 \( \mu \)m \( (0.065\lambda \) for A and 0.081\( \lambda \) for D) thickness, which was self-suspended on a Si substrate. For both resonators, the apodization ratio of IDT is 0.081. 

III. SIMULATION

The frequency characteristic of a ladder filter composed of the SH\(_0\) resonators was investigated by simulation. The filter has T-type architecture shown in Fig. 5(a), where two resonators (D) with \( \lambda \) of 6.2 \( \mu \)m are in a series arm, and a resonator (A) with \( \lambda \) of 7.75 \( \mu \)m is in a parallel. The wavelength ratio (WR) of the parallel/series arm resonators, \( \lambda_p/\lambda_s \), is 1.25. The filter was numerically synthesized from the measured frequency characteristics of the SH\(_0\) resonators (Fig. 4). The frequency characteristic was calculated using a circuit illustrated in Fig. 5 (b). F-matrices, \( F_1(\lambda f) \), \( F_2(\lambda f) \) and \( F_3(\lambda f) \), were made from the measured impedances \( Z_A \) and \( Z_D \) of resonators A and D, respectively, at each frequency \( f \). The F matrix of the ladder filter is obtained from

\[
F \equiv \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \equiv F_1 F_2 F_3
\]

(1)

Insertion loss, \( 20\log(V_2/V_1) \), in Fig. 5 (b) is obtained from

\[
20\log \left( \frac{V_2}{V_1} \right) = 20\log \left( \frac{1}{R_1(F_{21} + F_{22}/R_2) + F_{11} + F_{12}/R_2} \right)
\]

(2)

Fig. 6 shows the frequency characteristic of the ladder filter simulated using equations (1) and (2). An ultra-wide BW of 37\% at 25 dB attenuation is confirmed. However, the BW is narrower than the DTV band, and the passband slightly mismatches with the DTV band. Another problem is large undulation in the pass band. Therefore, the frequency and impedance of each resonator have to be adjusted to improve the frequency characteristic.

Fig. 7 shows the frequency characteristic which was adjusted as the passband matches with the DTV band. The frequencies of resonators A and D are shifted higher by 35 MHz and 60 MHz from the measured data, respectively. Also, the impedances of both resonators are reduced to a quarter of the measured ones. Note that their impedances are better but might not be best. The adjustment of frequency is just realized by adjusting the wavelength of the IDT. After adjustment, the
WR of the parallel/series arm resonators is 1.28. The impedance can be adjusted by the aperture and the number of IDT pairs. The simulated filter characteristic shown in Fig. 7 has a low minimum peak insertion loss of 0.5 dB and an ultra-wide 6 dB bandwidth of 206 MHz (35%). The BW at an attenuation level of 25 dB is as wide as the DTV band, but that at 6 dB is not enough to cover the DTV band. Therefore, the BW at 6 dB is not enough to cover the DTV band. Full attenuation level of 25 dB is as wide as the DTV band, but that is realized at $\lambda_p = 7.56 \mu m$ and $\lambda_m = 5.53 \mu m$, i.e. WR = 1.37.

IV. FABRICATION OF LADDER FILTER

Based on the simulation, the T-type and $\pi$-type ladder filters composed of three SH$_0$ resonators were fabricated using a (0°, 120°, 0°) LiNbO$_3$ plate of 0.61 and 0.62 $\mu m$ thickness. The LiNbO$_3$ plate was bonded with a Si substrate using adhesive, and polished down to the designed thickness. IDTs were fabricated by lift-off process, and the Si wafer was etched from the backside by Bosch process to make the thinned LiNbO$_3$ plate self-suspended. Fig. 8 shows the fabricated ladder filters.

Table I shows the specifications of the ladder filters. The WRs of the parallel/series resonators are 1.32 to 1.47 for T-type and 1.32 for $\pi$-type. The parallel and series resonators have the same number of IDT finger pairs and apertures. The IDT is not apodized. The resulted impedances of the parallel and series arm resonators are about 1/4.6 and 1/9.2 of those of resonators A and D, respectively. Actually, the impedances (Z) are a bit smaller than designed (1/4 of $Z_A$ and 1/8 of $Z_D$), because the LiNbO$_3$ plate was thicker (0.61 and 0.62 $\mu m$) than expected (0.5 $\mu m$).

Fig. 9 shows the frequency characteristics of the T-type ladder filters with WRs of 1.32, 1.39 and 1.47. The center frequencies are a little lower than that of the DTV band (590 MHz), but they can be easily adjusted by changing the wavelengths of the IDT. Fig. 10 shows the 6 dB BW as a function of WR, where the circles and the triangle represent the T-type and $\pi$-type ladder filters, respectively. The BW can be adjusted by WR. The fractional bandwidth of DTV is 41% (470-710 MHz) in Japan and USA, and 51% (470-790 MHz) in Europe. The BW of about 45% is suitable for the DTV band in Fig. 10. A BW of 55%, which is needed to fully cover European DTV band with a margin, is expected at a WR of 1.55, although it has not been demonstrated yet in this study.
Fig. 11 shows the frequency characteristics of the T-type and π-type ladder filters with a WR of 1.39. The BW at 6 dB attenuation is 47% and 45% for the T-type and π-type ladder filter, respectively. The π-type ladder filter has a larger attenuation at the lower side pole, which corresponds to the resonance impedance of parallel arm resonator, than that of the T-type, because the π-type has two parallel arm resonators and the T-type has one. In terms of BW, the T-type ladder filter apparently a larger BW. This is almost due to the π-type ladder filter having spurious responses in around 720 MHz, because the cavity is far from an optimum location as shown in Fig. 8(b). The lowest peak insertion loss is as low as 0.8 dB. Passband ripples due to transverse mode are observed, but they can be suppressed using an apodized IDT, as demonstrated in [5].

The ultra-wideband ladder filter technology demonstrated in which is followed by tunable band rejection filter. The filter characteristics can be improved by increasing the number of resonators, changing the capacitance ratio of parallel and series resonators, and optimizing the design.

V. CONCLUSION

The authors proposed a new BW- and frequency-tunable filter combining an ultra-wideband ladder filter and band rejection filters for DTV cognitive radio application. In this study, the ultra-wideband ladder filters, which fully covered the DTV band, were designed, fabricated and evaluated. T-type and π-type ladder filters composed of three ultra-wideband SH0 plate wave resonators were fabricated using a polished (0º, 120º, 0º) LiNbO3 plate of 0.61 and 0.62 μm thickness. Ultra-wide 6 dB BWs of 41% to 51%, which fully covered the DTV BW of 51% in Europe as well as 41% in Japan and USA, were measured. The BW can be adjusted by the WR of the parallel/series resonators in the ladder filter. The lowest peak insertion loss was as low as 0.8 dB. The ladder filter demonstrated in this study has the widest passband that has been ever reported for acoustic wave filters.

REFERENCES

Enhancement of Effective Electromechanical Coupling Factor by Mass Loading in Layered SAW Device Structures

Gongbin Tang1,2, Tao Han1, Akihiko Teshigahara3, Takao Iwaki3, and Ken-ya Hashimoto2
1School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
2Graduate School of Engineering, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba, Chiba 263-8522, Japan,
3Research Laboratories, DENSO CORPORATION, 500-1 Minamiyama, Komenoki, Nisshin, Aichi 470-0111, Japan
E-mail: gongbin.tang@chiba-u.jp

Abstract—This paper describes drastic enhancement of $K_e^2$ by mass loading in layered SAW device structures such as the ScAlN film/Si substrate. It is shown that this phenomenon is obvious even when an amorphous SiO$_2$ film is deposited on the top surface for temperature compensation. This enhancement is caused by SAW energy confinement to the top surface of the ScAlN layer where the IDT is placed. This $K_e^2$ enhancement is also found when other electrode and/or substrate materials are employed.

Keywords—ScAlN film; layered structure; effective electromechanical coupling factor; mass loading; SiO$_2$ overlay

I. INTRODUCTION

Nowadays, radio frequency (RF) surface acoustic wave (SAW) devices are widely used in modern communication systems. With the growing demand for high operation frequency, low loss and wide bandwidth, many novel device structures are raised.

The effective electromechanical coupling factor $K_e^2$ is a key parameter to determine the realizable filter bandwidth. Recently, Scandium-doped AlN (ScAlN) films have drawn much attention due to the strong piezoelectricity[1-15]. The authors reported that large SAW velocity $V$, small propagation attenuation, and large $K_e^2$ of about 6.1% are simultaneously achievable when the film is combined with a high velocity base substrate such as single crystal diamond (SCD) and 6H-SiC for the Sezawa mode even for operation in the 3 GHz range [8-10].

In many cases, $K_e^2$ of the SAW is much lower than that of the bulk wave even when the same piezoelectric material is chosen. This is because the SAW field distribution does not match well to the electric field generated by the interdigital transducer (IDT), and thus $K_e^2$ can be sometimes enhanced when heavy electrodes are used and/or the IDT is inlaid in the layered substrate.

This effect is well known also for the bulk acoustic wave (BAW) devices[16], and is widely used in practical device design.

This paper describes drastic enhancement of $K_e^2$ by mass loading in layered SAW device structures such as the ScAlN film/Si substrate.

II. SIMULATION

At first, we choose the ScAlN/Si structure as an example, and it is shown that $K_e^2$ is considerably enhanced by the mass loading of the electrodes. This enhancement is caused by SAW energy confinement to the top surface of the ScAlN layer where the IDT is placed.

Then it is shown that such $K_e^2$ enhancement is quite often seen in SAW device structures composed of a thin piezoelectric film and a high velocity base substrate. It should be noted that this phenomenon is obvious even when an amorphous SiO$_2$ film is deposited uniformly on the top surface for temperature compensation.

Fig. 1 shows the device model of the Cu electrode/Sc$_{0.43}$Al$_{0.57}$N film/Si structure used in the simulation. The simulation was performed by the software SYNCL[17,18], which calculates the input admittance of the infinitely long IDT of the device structure as a function of the driving frequency $f$. Material constants for Sc$_{0.43}$Al$_{0.57}$N were taken from [10], while those of Si, Cu, Al and SiO$_2$ were taken from [19].

Fig.1. Model structure used for simulation.
effective SAW velocity $V$ and $K_e^2$ are estimated using the following equations:

$$V = f_r \lambda$$  \hspace{1cm} (1)$$

and

$$K_e^2 = \left( \frac{\pi f_r}{2 f_a} \right) \cot \left( \frac{\pi f_r}{2 f_a} \right)$$  \hspace{1cm} (2)$$

respectively.

III. INFLUENCE OF IDT MASS LOADING

At first, the SAW phase velocity $V$ and the $K_e^2$ are estimated for the uniform ScAlN/Si structure without taking the influence of the IDT thickness into account. The calculation was performed by the VCALL[17,18].

Fig. 2. SAW properties on the ScAlN/Si structure when the mass loading is not taken into account. (a) Variation of $V$ with $h_{\text{ScAlN}}/\lambda$, (b) Variation of $K_e^2$ with $h_{\text{ScAlN}}/\lambda$.

Fig. 2 shows estimated $V$ and $K_e^2$ as a function of the $h_{\text{ScAlN}}/\lambda$, where $h_{\text{ScAlN}}$ is the ScAlN thickness. Multiple propagation modes exist, and are labeled in the order of their phase velocities. It is seen that the second (Sezawa) mode exhibits relatively large $K_e^2$, which takes a maximum value of 2.9% at $h_{\text{ScAlN}}/\lambda = 0.62$, where $V$ is 5,470 m/s. The maximum value is much smaller than the value of 6.1% obtainable in the ScAlN/SCD structure.

This is due to difference in the elasticity between Si and SCD, which determines the SAW energy penetration into the base substrate.

It should be noted that $K_e^2$ of the fundamental (Rayleigh) mode is not so small. This fact implies that the electric field generated by the IDT does not fit well with the Sezawa mode, and thus the Rayleigh mode is efficiently excitable.

Next, $h_{\text{ScAlN}}$ is fixed at 0.62$\lambda$, and influence of the Cu electrode thickness $h_{\text{Cu}}$ is investigated. Fig. 3 shows variation of $V$ and $K_e^2$ in the Cu/ScAlN/Si structure as a function of $h_{\text{Cu}}/\lambda$. It is seen that $K_e^2$ for the Sezawa mode increases with $h_{\text{Cu}}/\lambda$, and takes a maximum value of 9.1% at $h_{\text{Cu}}/\lambda = 0.15$, which is more than three times larger than the value when $h_{\text{Cu}}/\lambda = 0$. In contrast, $K_e^2$ for the Rayleigh mode decreases with $h_{\text{Cu}}/\lambda$. This confirms variation of $K_e^2$ is due to variation of the SAW field distribution to fit with the electric field. Owing to the mass loading, $V$ is reduced, but the value is still large (4,883 m/s) at the thickness.

Fig. 3. SAW properties on the ScAlN/Si structure when the mass loading is taken into account. (a) Variation of $V$ with $h_{\text{Cu}}/\lambda$, (b) Variation of $K_e^2$ with $h_{\text{Cu}}/\lambda$. 

417
Fig. 4 shows change of $K_e^2$ with $h_{\text{ScAlN}}$ and $h_{\text{Cu}}$. The maximum $K_e^2$ is 9.5% at $h_{\text{Cu}}/\lambda \sim 0.15$ and $h_{\text{ScAlN}}/\lambda \sim 0.55$, where $V$ is 5,080 m/s.

This $K_e^2$ enhancement also occurs in various structures. Table I shows the maximum $K_e^2$ for some representative structures. It is seen that Cu offers larger $K_e^2$ enhancement than Al due to difference in the mass density.

IV. INFLUENCE OF SiO₂ LOADING

Because the $K_e^2$ enhancement described in this report is owed to the mass loading, addition of a uniform layer on the IDT is another choice. Here we investigate use of SiO₂ to the $K_e^2$ enhancement in addition to the temperature compensation. We can adjust the SiO₂ and electrode thicknesses to achieve large $K_e^2$ and temperature compensation simultaneously.

Fig. 5 shows variation of $V$ and $K_e^2$ for the Sezawa mode with $h_{\text{SiO₂}}$ on SiO₂/Cu/ScAlN/Si structure at 3 different electrode thicknesses while $h_{\text{ScAlN}}$ is fixed at 0.55λ. Here, the SiO₂ film is assumed to cover on not only the electrodes but also the gap between electrodes uniformly (see the inset in Fig. 5). It is seen that when $h_{\text{Cu}}$ is small, $K_e^2$ increases with $h_{\text{SiO₂}}$, and the maximum $K_e^2$ of 8.1% is achievable when both $h_{\text{SiO₂}}$ and $h_{\text{Cu}}$ are set properly. The maximum $K_e^2$ is scarcely dependent on $h_{\text{Cu}}$, and $h_{\text{SiO₂}}$ giving the maximum value becomes small with an increase in $h_{\text{Cu}}$.

The uniform SiO₂ deposition results in small reduction in the maximum $K_e^2$ because SiO₂ in the gap region increases the static capacitance of the IDT.

This $K_e^2$ enhancement is also found when other electrode and/or substrate materials are employed. For example, the Cu/ScAlN/SCD structure also offers the $K_e^2$ enhancement (from 5.5% to 9.8%). This enhancement is smaller than that of the Cu/ScAlN/Si structure (from 2.9% to 9.5%). However, the SAW velocity in this case (~5,600 m/s) is somewhat higher than the latter case (~5,000 m/s), which might be favorable for practical use.

V. CONCLUSION

This paper described drastic enhancement of $K_e^2$ by mass loading in layered SAW device structures such as the ScAlN film/Si substrate.

It was shown that this phenomenon is obvious even when

<table>
<thead>
<tr>
<th>Structure</th>
<th>$K_e^2$ [%]</th>
<th>$V$ [m/s]</th>
<th>$h_{\text{Cu}}/\lambda$</th>
<th>$h_{\text{ScAlN}}/\lambda$</th>
<th>$K_e^2$ w/o IDT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cu/ScAlN/SCD</td>
<td>9.8</td>
<td>5,782</td>
<td>0.11</td>
<td>0.71</td>
<td>5.5%</td>
</tr>
<tr>
<td>Al/ScAlN/SCD</td>
<td>9.0</td>
<td>5,869</td>
<td>0.19</td>
<td>0.57</td>
<td>2.8%</td>
</tr>
<tr>
<td>Cu/ScAlN/Sapphire</td>
<td>7.1</td>
<td>5,331</td>
<td>0.16</td>
<td>0.62</td>
<td>2.9%</td>
</tr>
<tr>
<td>Al/ScAlN/Sapphire</td>
<td>6.5</td>
<td>5,382</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu/ScAlN/Si</td>
<td>9.1</td>
<td>4,883</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Al/ScAlN/Si</td>
<td>7.7</td>
<td>4,876</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
an amorphous SiO$_2$ film is deposited on the top surface for temperature compensation. This enhancement is caused by SAW energy confinement to the top surface of the ScAlN layer where the IDT is placed.

ACKNOWLEDGEMENT
This work was partially supported by the Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science and the National Natural Science Foundation of China. GT acknowledges the support of the Japanese Government (MEXT) for the scholarship through the Super Global University Project.

REFERENCES
Second Order Temperature Compensated Piezoelectrically Driven 23 MHz Heavily Doped Silicon Resonators with ±10 ppm Temperature Stability

Antti Jaakkola, Panu Pekko, James Dekker, Mika Prunnila and Tuomas Pensala
VTT Technical Research Centre of Finland
Espoo, Finland
antti.jaakkola@vtt.fi

Abstract—We report quartz level temperature stability of piezoelectrically driven silicon MEMS resonators. Frequency stability of better than ±10 ppm is measured for 23 MHz extensional mode resonators over a temperature range of \( T = -40 \ldots +85 \) °C. The temperature compensation mechanism is entirely passive, relying on the tailored elastic properties of heavily doped silicon with a doping level of \( n > 10^{20} \text{cm}^{-3} \), and on an optimized resonator geometry. The result highlights the potential of silicon MEMS resonators to function as pin-to-pin compatible replacements for quartz crystals without any active temperature compensation.

I. INTRODUCTION

While silicon MEMS based solutions have well known advantages to offer to the timing and frequency control applications, the market is still dominated by quartz devices. The adoption of silicon MEMS resonator technology could be greatly enhanced if the devices could be made pin-to-pin compatible with quartz crystals through the combination of piezoelectric actuation and fully passive doping based temperature compensation. This paper reports progress towards this goal: we have designed and fabricated piezoelectrically driven 23-MHz silicon MEMS resonators, which have a ±10 ppm temperature stability corresponding to that of an AT cut quartz crystal, and whose electrical characteristics approach those of quartz at the same frequency.

Recent research has shown that heavy phosphorus doping \((10^{19} \text{cm}^{-3})\) of silicon can be used for reducing the thermal drift of a MEMS resonator frequency from over 3000 ppm to less than 300 ppm over the industrial temperature range [1]. It has been identified that further doping to carrier concentrations above \(10^{20} \text{cm}^{-3}\) has the potential to reduce the temperature dependency through its effect on the 2nd order temperature coefficient \(T CF_2\) [2]. Recently, we have verified this to be true with ultra heavily doped (UHD) capacitively coupled (bare silicon) resonators, illustrated in Fig 1. This data leads to two observations:

1) There is an optimal doping level which, together with a correct design, produces near zero \(T CF_2\), and yields quartz level temperature stability. The best experimen-
tally demonstrated level of stability is ±10 ppm for 
\[ T = -20 \ldots + 85^\circ C \].

2) The usually negative second order temperature coefficient \( TCF_2 \) can be made positive, up to \( TCF_2 \sim +15 \text{ppb/}^\circ C^2 \).

Observation 1 leads to attractive possibilities for realizing various types of passively temperature compensated bare silicon resonators, but observation 2 is the key for realizing piezoelectrically driven silicon MEMS resonators with quartz class temperature stability: Piezoelectric actuation requires addition of piezoelectric and metal (electrode) layers to the resonator device, and, practically all such materials (typically AlN + Mo/Al) have negative first- and second order temperature coefficients \( TCF_1 \) and \( TCF_2 \), respectively [3], [4], [5]. By balancing the positive contribution from the UHD silicon resonator body and the negative effect from the piezoelectric and metallic layers to both \( TCF_1 \) and \( TCF_2 \) by correct composition of the resonator, it is possible to reach quartz level frequency stability.

II. METHODS

A. Resonator design and fabrication

The resonators were fabricated using the VTT cavity-SOI based process platform, see Fig. 2. First, SOI wafers featuring ultra heavily doped (doping in excess of \( 10^{20} \text{cm}^{-3} \)) Si device layers, including pre-etched cavities were prepared (steps 1,2). Next, AlN was deposited and patterned (3) right onto the Si device layer acting as a substrate and as the bottom electrode for device operation. SiO\(_2\) was deposited on the wafer (4). AlN was located only on top of the resonator, while SiO\(_2\) was used as the insulator between the top and bottom electrodes elsewhere. Openings were etched to the SiO\(_2\) layer, one onto the AlN layer and another onto the place where the bottom electrode contact would be formed (5). Aluminum was deposited and patterned as the top electrode material (6). Deep reactive ion etching was used to define the device geometry and to release the resonator (7).

![Fabrication process](image)

Figure 2. Fabrication process. UHD-Si refers to ultra heavily doped silicon with n-type carrier concentration above \( 10^{20} \text{cm}^{-3} \).

B. Measurements

The frequency-vs-temperature curve measurements were performed on wafer level under atmospheric pressure on a Cascade Summit probe station using a HP 4294A impedance analyzer. The resonance frequencies were extracted by fitting the response of a BVD equivalent circuit to the measured admittance traces spanning the resonance peak (see Fig. 4). The wafer was held on a temperature-controlled chuck, whose temperature was varied from −40°C to +85°C with seven steps. A flow of dry air was used to prevent condensation of moisture on the non-packaged resonators. A total of \( \sim 30 \) resonators were characterized on a wafer.

III. RESULTS

A frequency stability better than ±10 ppm was measured for several devices over a temperature range of \( T = -40 \ldots + 85^\circ C \). The frequency-vs-temperature curves of three resonators are shown in Figure 3. Typical resonator performance parameters were: \( R_m \sim 100 \text{\ Ohm} \), \( C_0 \sim 11 \text{\ pF} \), \( Q \sim 4000 \), \( f_0 - f_s \sim 750 \text{\ ppm} \), and \( k^2 \sim 0.15\% \) — a frequency response of a resonator is shown in Fig. 4. The scatter between the \( f - vs - T \) curves of the set of \( \sim 30 \) resonators spanning the whole wafer are shown in Fig. 5.

![Frequency response of the resonator](image)

Figure 4. Frequency response of the resonator. \( Q \), \( C_0 \), \( R_m \) and \( f_0 \) were obtained by fitting a BVD equivalent circuit to the data.

IV. DISCUSSION

The result of ±10 ppm frequency stability shows that passively temperature compensated piezoelectrically coupled silicon MEMS resonators can reach a similar performance to AT cut quartz crystals. With further optimization (reduction of \( TCF_2 \)), there is a potential for even better stability.

It can be seen in Fig. 3 that the measured data points do not accurately lay on top of the quadratic fit for the \( f - vs - T \) curve. At maximum, a fit error of \( \sim 5 \text{\ ppm} \) is observed. The measurement was done in open air, which could cause instability of the resonance frequency at this level. Encapsulation of the devices is needed for performing more accurate measurements.
The performance parameters of two types of VTT MEMS resonators (HD-Si and UHD-Si) and a typical quartz crystal at the same frequency are compared in Table I. It can be seen that the electrical performance in terms of the equivalent series resistance (ESR) and shunt capacitance $C_0$ corresponds to that of quartz for the case of HD-Si resonators, however, in this case the frequency instability is an order-of-magnitude too high (see Fig. 1 as well). For the UHD-Si resonator of this work, frequency stability is sufficient, but the electrical performance parameters do not yet quite reach those of quartz, and thus complete pin-to-pin compatibility is not yet realized for these devices. Our further work includes reduction of ESR (through increased $Q$) by more optimized anchoring of the resonator. Reduction of the shunt capacitance will be assessed in particular by using thicker AlN layer on the resonator. The target is that the resonator can be driven with a standard oscillator IC intended for quartz crystals.

Figure 5 illustrates the scatter between the $f$-vs-$T$ curves as well as the initial accuracy error of resonators on a wafer. For approximately two thirds of the resonators, the frequency stability is within ±20 ppm. The distribution of the initial accuracy has a standard deviation of 1500 ppm. It is of paramount importance to develop an economically viable way of reducing the scatter of these properties. Both the temperature characteristics as well as the frequency of the resonator are functions of the Si/AlN/Al stack. Thus, selective addition or removal of material can be used for their fine tuning. Mapped ion beam trimming is a particularly promising technique for assessing this problem [6].

V. CONCLUSION

Piezoelectrically driven 23-MHz silicon MEMS resonators having a ±10 ppm temperature stability and electrical characteristics approaching those of quartz crystals were demonstrated. The presented result greatly improves the competitiveness of silicon based resonator technology in timing and frequency reference applications, and presents an attractive alternative to current silicon MEMS approaches using active (PLL-based) temperature compensation. The work suggests that eventual pin-to-pin compatibility between silicon MEMS resonators and quartz is within reach.

REFERENCES


Highly Tuneable X-Band Bragg Resonator - Initial Results

Pratik D Deshpande
Department of Electronics
University of York
York, United Kingdom
pd617@york.ac.uk

Simon J. Bale
Department of Electronics
University of York
York, United Kingdom
simon.bale@york.ac.uk

Mark Hough
Department of Electronics
University of York
York, United Kingdom
mark.hough@york.ac.uk

Jeremy Everard
Department of Electronics
University of York
York, United Kingdom
jeremy.everard@york.ac.uk

Abstract—This paper describes the design and measurement of a broad tuning aperiodic Bragg resonator at X-band. The resonator utilizes an aperiodic arrangement of non-$(\pi/4)$ low loss alumina plates ($E_r=9.75$, loss tangent of $\sim$1 to $2\times10^{-5}$) mounted in a cylindrical metal waveguide. The initial results demonstrate a spurious free tuning range of 100MHz. The insertion loss, $S_{21}$, varies from -5.6dB to -4dB while the unloaded Q varies from 40,000 to 60,000 over the tuning range. The loaded Q varied from 20,000 to 23,000. It is also possible to tune over a range of 400MHz with similar unloaded Qs up to 64,000, but on occasion the required $TE_{21}$ mode passes through several lower Q modes which therefore degrades the unloaded Q of the wanted mode.

Keywords—Bragg Resonator, Dielectric Resonator

I. INTRODUCTION

Broad tuning ultra-high Q cavities enable versatile ultra-low phase noise tuneable oscillators offering broadband electromechanical course tuning and electronic fine tuning. These could achieve noise floors below -200dBc at X band. Fixed frequency alumina based aperiodic cylindrical Bragg resonators, developed by this group, show unloaded Qs exceeding 200,000 [1] with some sapphire structures demonstrating unloaded Qs exceeding 700,000 [2].

Theoretical simulations predict that this resonator could offer broad electro-mechanical tuning if just the length of the centre section is varied [3]. This is because the Bragg mirrors produce very low loss high reflectivity over a broad frequency range exceeding 10% of the centre frequency.

Previous attempts at changing this length using concentric cylinders were unsuccessful due to energy loss from the wanted high Q mode. Further, to achieve low insertion loss the probes (or probe for a one port) have to be in the correct position in the cavity.

II. RESONATOR MODELLING

The performance of the Bragg resonator can be modelled by using ABCD parameters to describe the resonator as a cascaded set of waveguides [1]. The advantage of using ABCD matrices is that the response of cascaded sections is just the product of the matrices. This is achieved through the definition of the directions of the input and output currents.

The resonator model must now contain ABCD matrices for the air sections, dielectric sections and for the end wall to form the Bragg Resonator. The ABCD matrix for a lossy transmission line of length $l$ meters with complex propagation constant $\gamma$ and characteristic impedance $Z_o$ is shown in (1).

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} \cosh(yl) & Z_o\sinh(yl) \\ \frac{1}{Z_o}\sinh(yl) & \cosh(yl) \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

(1)

The complex propagation constant, $\gamma$, is defined as:

$$\gamma = \alpha + j\beta$$

(2)

Where $\alpha$ is the attenuation co-efficient ($Np/m^{-1}$) and $\beta$ is the phase constant ($rad \ m^{-1}$). The phase constant for the air and the dielectric sections can be found by using:

$$\beta = \sqrt{\frac{\omega^2\mu\varepsilon - (\frac{\chi_{mn}}{a})^2}{}}$$

(3)

where $\varepsilon$ is the permittivity of the material filling the guide, $\omega$ is the angular frequency and $a$ is the cavity radius. $\chi_{mn}$ represents the $n^{th}$ zero of the derivative of the Bessel function of the first kind of order $m$. In the case of the $TE_{21}$ mode the value of $\chi_{mn} \approx 3.8318$. The attenuation coefficients for the various sections of the Bragg resonator are now discussed.

A. Air Section:

The loss in the air sections is due to the conductive side wall losses. This can be modelled as shown in Eq. (4). This equation represents the attenuation coefficient, in units of $Np/m^{-1}$, for a transverse electric (TE) mode with circumferential mode number $m$ and radial mode number $n$ in a cylindrical waveguide of radius $a$ operating at frequency, $f$:

$$\alpha_{air} = \frac{R_s}{\eta \left(1-\frac{f_c}{f}\right)^2 + \frac{m^2}{(\frac{\chi_{mn}}{a})^2-m^2}}$$

(4)

where $\eta$ is the wave impedance for a plane wave inside an unbounded infinite medium with permittivity, $\varepsilon$ and permeability $\mu$.

B. Dielectric Section:

The total loss in the dielectric sections, $\alpha_{total}$, can be considered as the sum of the sidewall conducting loss, $\alpha_{air}$, and the dielectric losses, $\alpha_d$. 
The conductive side wall losses can be calculated using (4) but the loss in the dielectric must be treated differently. The attenuation due to the lossy dielectric, \( \alpha_d \), can be calculated from the complex propagation constant as shown in [4]. If the loss is small then the phase constant in the dielectric section can be assumed to be constant. The attenuation due to dielectric loss is given by equation (6):

\[
\alpha_d = \frac{\omega^2 \mu \tan \delta}{2 \sqrt{\omega^2 \mu - \left( \frac{\chi_{mn}}{a} \right)^2}}
\]  

\( \text{(6)} \)

C. End Wall:
The loss in the metal end walls of the cavity can be approximated by considering the complex propagation constant, \( \gamma \), and intrinsic wave impedance, \( \eta \), for a plane wave in a good conductor[5]. The ABCD parameters for the end wall section can be written as

\[
\begin{bmatrix}
V_1 \\
I_1
\end{bmatrix} = \begin{bmatrix}
1 & 0 \\
1/Z_S & 1
\end{bmatrix} \begin{bmatrix}
V_2 \\
I_2
\end{bmatrix}
\]

\( \text{(8)} \)

Where:

\[
Z_S = (1 + j) \sqrt{\frac{\omega \mu}{2 \sigma}}
\]  

\( \text{(9)} \)

Where \( \sigma \) is the electrical conductivity of cavity shield.

The loss tangent (\( \tan \delta \)) of the dielectric material determines the maximum unloaded quality factor. The attenuation in the air and the dielectric sections along with the wall losses degrade the unloaded quality factor. Hence, to maximize the quality factor of the resonator, it is critical that wall and dielectric losses are minimized. The side wall loss can be reduced by using a high conductivity metal such as copper, silver or silver plated Aluminium. Therefore \( \gamma_{\text{air}} \) and \( \gamma_{\text{dielectric}} \) [4] [5] [6] have to be calculated in order to model the individual ABCD matrices of the air and the dielectric sections. Table I summarizes the various constants used in the ABCD model.

### Table I. Bragg resonator simulation parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavity radius</td>
<td>( a )</td>
<td>60 mm</td>
</tr>
<tr>
<td>Dielectric permittivity</td>
<td>( \varepsilon_r )</td>
<td>9.75</td>
</tr>
<tr>
<td>Dielectric Loss Tangent</td>
<td>tan( \delta )</td>
<td>2 \times 10 (^{-1} )</td>
</tr>
<tr>
<td>Wall conductivity (using silver)</td>
<td>( \sigma )</td>
<td>6.173 \times 10 (^{-1} ) Sm (^{-1} )</td>
</tr>
<tr>
<td>Air-Attenuation coefficient</td>
<td>( \alpha_a )</td>
<td>1.09 \times 10 (^{-4} ) Npm (^{-1} )</td>
</tr>
<tr>
<td>Air-Phase constant</td>
<td>( \beta_a )</td>
<td>199.68 radm (^{-1} )</td>
</tr>
<tr>
<td>Dielectric - Attenuation coefficient</td>
<td>( \alpha_d )</td>
<td>6.61 \times 10 (^{-3} ) Npm (^{-1} )</td>
</tr>
<tr>
<td>Dielectric - Phase constant</td>
<td>( \beta_d )</td>
<td>651.18 radm (^{-1} )</td>
</tr>
</tbody>
</table>

### III. Simulations of a Periodic Resonator Using the ABCD Model

The parameters shown in Table I were used in the ABCD model and an S-parameter simulation was performed for a periodic Bragg resonator. In a periodic Bragg resonator each of the dielectric plates and air sections are one quarter of the guide wavelength (\( \lambda_g/4 \)) in thickness in order to maximize their reflectivity [7].

### IV. Simulations of the Aperiodic Resonator

Once the dimensions of the air sections and the dielectric sections for a periodic Bragg resonator were obtained, the model was then split in two.

The reflector section lengths were then optimized until the magnitude of the input reflection coefficient at port one (\( S_{11} \)) reached a maximum using the ABCD solver and a genetic algorithm. The phase response of the reflection was taken into account by adjusting the length of the centre section.

This then produced an aperiodic Bragg resonator. The lengths are now dependent on the losses and dispersion in each section as well as the frequency of operation. The dimensions of the individual sections are given in Table II.

### Table II. Dielectric and air section reflector thicknesses for an optimised 6 plate Bragg resonator. [1]

<table>
<thead>
<tr>
<th>Section Identifier</th>
<th>Material</th>
<th>Length (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Dielectric</td>
<td>1.512</td>
</tr>
<tr>
<td>L2</td>
<td>Air</td>
<td>11.023</td>
</tr>
<tr>
<td>L3</td>
<td>Dielectric</td>
<td>1.887</td>
</tr>
<tr>
<td>L4</td>
<td>Air</td>
<td>9.300</td>
</tr>
<tr>
<td>L5</td>
<td>Dielectric</td>
<td>2.253</td>
</tr>
<tr>
<td>L6</td>
<td>Air</td>
<td>8.060</td>
</tr>
<tr>
<td>LC</td>
<td>Air</td>
<td>17.033</td>
</tr>
</tbody>
</table>

### V. Design of Tunable Resonator

As stated earlier, tuning can be achieved by changing the length of the centre section, this is because the Bragg mirrors offer low loss high reflectivity over a broad frequency range, exceeding 10% of the centre frequency.

This is illustrated by tuning the length of the centre section by \( \pm 20\% \). Note that the frequency changes by twice this difference. Further, this model only considers the wanted \( \text{TE}_{011} \) mode.

The nominal length of the centre section (17.033mm) was tuned by \( \pm 4 \) mm in 1 mm increments and the ABCD model was used to simulate the new unloaded Q and centre frequency for the \( \text{TE}_{011} \) mode. The plot of change in frequency and Unloaded Q vs change in length is shown in Fig.1.
VI. DESIGN AND CONSTRUCTION OF THE TUNABLE CENTRE SECTION

The resonator developed in [1] was modified and a new centre section designed. The air waveguide dimensions of the centre sections were optimized to incorporate the thickness of the copper sheets which form the tuning bellows. The initial prototype of the resonator is shown in Fig. 2. The central section comprises an upper section, bottom section and a solid middle section for the probes with two bellows either side of the centre section.

![Fig. 2: Cross section view of a 6 plate Tuneable aperiodic Bragg Resonator](image)

The following method was used to construct the centre section:

Firstly, each bellows is made up of two large copper rings with etched solder release groves in order to control the position of the solder as shown in Fig.3.

This controls the exact position of the solder within the bellows and also prevents it from flowing into the cavity which may degrade the Q. These rings also have a number of tabs around the outer edges which are folded shut to ensure the bellows remain soldered during the later processing stages. As a number of different soldering operations have to be performed at different times, two different temperature solders were used to stop the solder from reflowing when a new joint was produced.

![Fig. 3: Copper sheets with etched solder release grooves in order to control the position of the solder](image)

A. Coupling Loops

To obtain the correct ratio of loaded to unloaded Q ($Q_l/Q_0$) and insertion loss for low noise oscillators [8], the probes need to be placed in the middle of the centre section close to the cavity wall.

The tunable centre section is shown in Fig 4. Micrometres were used to tune the length of the central section. In the final design electronic micrometers with piezo nano tips will be used.

![Fig. 4: Centre section with the micrometers and the loop probes to couple energy into and out of the cavity](image)

The complete assembly is shown in Fig 5.

![Fig. 5: Tuneable Bragg resonator with micrometers](image)

VII. CURRENT RESULTS

The micrometers were used to tune the frequency of the cavity over 500 MHz and the insertion loss and the loaded Q were measured on a network analyser. The highest unloaded Q obtained was 64,000. It was observed that the required TE011 mode.
mode passes through several low Q modes degrading quality factor at certain frequencies as shown in Fig.6.

![Fig. 6: Plot of insertion loss and unloaded Q vs Frequency with a 500MHz span](image)

A spurious free region was noted over a 100MHz tuning range. In this region the insertion loss, $S_{21}$, varies from -5.6dB to -4dB while the loaded Q varied from 20,000 to 23,000. The unloaded Q can be found by using (10):

$$Q_L = \frac{Q_L}{1-S_{21}}$$  \hspace{1cm} (10)

Where $Q_L$ is the loaded quality factor and $S_{21}$ is the insertion loss. Using equation (10) the unloaded Q varies from 40,000 to 60,000 over the tuning range of 100MHz. The plot of insertion loss and unloaded Q vs Frequency with the narrower 100MHz span is shown in Fig.7.

The unloaded Q is significantly lower than the numerical simulation results. This may be due to conductor losses in and around the bellows including losses in the solder as well as leakage and mode conversion due to the discontinuities in the structure.

![Fig. 7: Plot of insertion loss and unloaded Q vs Frequency with a narrow 100MHz span with no unwanted modes](image)

VIII. CONCLUSIONS AND FUTURE WORK

The initial results demonstrate a spurious free tuning range of 100 MHz. The insertion loss, $S_{21}$, varies from -5.6dB to -4dB while the unloaded Q varies from 40,000 to 60,000 over the tuning range. The loaded Q varied from 20,000 to 23,000. X-band oscillators with a fixed frequency Bragg resonator which have an unloaded Q of 210,000 are currently under investigation and noise floors less than $-200$dBc/Hz are expected. Also, residual phase noise measurements of the active components are being measured using a broadband cross correlation phase noise measurement System [9]. This system has a noise floor of approximately $-204$ dBc/Hz at L Band.

ACKNOWLEDGMENTS

We wish to thank Mr. John Clapham for his help in the construction of the main cavity.

REFERENCES

The effect of contour concentricity on the acceleration sensitivity of quartz crystal resonators

Morley, Peter E.
Vectron International
4914 Gray Road, Cincinnati, OH 45232, USA
pmorley@vectron.com

Abstract—Optimization of the acceleration sensitivity of quartz crystal resonators has been a challenging problem for resonator designers for decades. The structural symmetry of the resonator and mount combination has been shown in past work, both theoretical and practical, to have a strong influence on acceleration sensitivity, and specialized structures have been developed [1], [2], [3] that have greatly improved performance. However, with applications such as airborne radar systems, there is a persistent demand for further improvement.

The design of many of the practical high-stability resonator products that have a need for good acceleration sensitivity is also constrained by other attributes, such as high quality factor, and these constraints typically result in a low-frequency overtone device with a fully contoured resonator element design. In this paper, the effect of the concentricity of the contour shape on the quartz disk in contoured resonators is considered, and results are presented that demonstrate a strong correlation between the contour offset from the blank center and the acceleration sensitivity of the resonator. Methods are also described for measurement of the contour position relative to the perimeter of the disk.

Keywords—Quartz crystal, resonator, contoured, sensitivity, acceleration, g-sensitivity, optical measurement.

I. BACKGROUND

The acceleration sensitivity, sometimes called g-sensitivity, of crystal resonators has been widely discussed over the past few decades. The parameter is most important in applications that require good phase noise, but where the device is exposed to high vibration fields. A good example would be a frequency reference for a radar system in a helicopter. Theoretical work by Tiersten and Zhou and others essentially concluded that a quartz resonator with perfect spatial symmetry in both the resonator element and the mounting structure will exhibit zero g-sensitivity [5],[6],[7],[8].

Theoretical and practical work by Eernisse and colleagues [9],[10],[11],[12] proposed and implemented practical mount designs to approximate symmetric structures with the aim of achieving low acceleration-induced frequency shifts. They also looked at the use of carefully positioned masses deposited onto the surface of blanks to modify the position of the resonance mode and hence to improve g-sensitivity. This technique may be very useful for planar, higher frequency resonator designs, but in typical lower frequency, low phase-noise designs, the blank geometry is by necessity contoured, and in these cases incremental mass loading changes on the electrode surface have very little influence on the mode shape. Many other authors have discussed practical and theoretical considerations for achieving low acceleration sensitivity, including Kosinsky [13],[14] and Lee [15]. Practical designs were also developed in France from the 1970s onwards in the form of the highly complex BVA structures that utilize quartz bridges in the resonator elements as well as multiple quartz components and conductive structures to provide symmetrical support for the resonator [16],[3].

Haskell et al [1],[2],[4] introduced the patented Quad Relief Mount product or QRM. This uses a planar mount design configuration that is positioned to coincide with the central plane of the resonator element as shown in Fig 1. The outer ring is a rigid ceramic structure that is firmly attached to the crystal base and in between the ring and the blank is an essentially planar array structure that also provides a reduction in static stresses to the resonator. This design approach has achieved excellent performance, with g-sensitivity results below 10^{-10}/g in some cases, while still performing well for Q and phase noise. However, as is often the case with crystal parameters, there is typically a distribution in performance for g-sensitivity in each manufactured group, and this causes yield problems as well as unpredictability in production scheduling. The work reported here aimed to find the root cause of these anomalous results, with a focus on the contour concentricity.

II. EXPERIMENTAL APPROACH

The focus for this work was a typical QRM resonator type that is currently manufactured: a 10MHz 3rd overtone SC cut for ovenized oscillator application. The design uses a plano-
convex blank geometry with a contour of approximately 1.5 diopters on the convex side. Rather than deliberately manufacturing units with known asymmetries and then measuring them for acceleration sensitivity, the approach that was used was to select parts from past groups with a range of performance. The units were re-measured to verify g-sensitivity results and then inspected for manufacturing anomalies. Finally the blanks were removed from the mounts and the electrodes stripped to allow analysis of the contour.

III. CONTOUR OFFSET MEASUREMENT METHODS

The geometry of a spherically contoured surface is shown in Fig 2. Historically, because the machining processes in quartz crystal manufacture were derived from methods used in the optical lens industry, the radius of curvature is often specified in diopters. Strictly, this parameter is only defined for a medium with a known refractive index and, as described later, the refractive index is not well defined for crystalline quartz, so the index for crown glass of 1.525 is usually substituted, which results in the relationship $R = \frac{525}{D}$ where $R$ is measured in mm and $D$ is the diopter value.

![Fig 2 Contour geometry](image)

To derive the relationship between the radius of curvature $R$ and incremental thickness change $\varepsilon$ at an offset radius $r$, application of Pythagoras gives:

$$(R - \varepsilon)^2 + r^2 = R^2$$

And omitting the $\varepsilon^2$ term gives

$$\varepsilon \approx \frac{r^2}{2R}$$

There are various viable methods for measurement of the concentricity between the blank periphery and the contour surface which determines the mode position of the resonance. There are pros and cons for each technique depending on the geometry being observed, so the methods were evaluated for the particular blank geometry used in this product.

A. Pre-electroding

One potential option for measurement of the contour offset of a blank is based on the relationship between the electrode location relative to the blank geometry and the resulting motional parameters of the resonator. To determine the relationship between contour center position and the resulting motional capacitance $C_1$ of a resonator, a model was created using the structural mechanics module of Comsol Multiphysics. An example of a 3D plot showing displacement intensity for the 10 MHz 3rd overtone SC cut resonator used in this study is shown in Fig 3. The model was set up with varying contour offset, and the resulting relationship between $C_1$ and contour offset is shown in Fig 4 for various electrode diameters. Clearly this relationship could also be derived analytically, but Comsol has proven to be a very useful tool for this type of calculation.

![Fig 3 Contour map of resonator with offset contour](image)

To sort a group of contoured blanks using this method, they would first be checked for contour radius, since this is clearly also a parameter that strongly influences $C_1$. They would then be accurately plated with small circular electrodes in the center of the blanks, preferably with an electrode material that is easily removed, and then inserted into temporary mounts. The optimum size of the electrode depends on the design being analyzed. A simple motional parameter check would then provide the tool to select for good contour concentricity, after which the electrodes would be removed and then the parts re-processed into the final product.

![Fig 4 C1 vs. contour offset for various electrode diameters](image)

B. Profile measurement

Another method that can be considered to measure contour concentricity is a 1D or 2D profile measurement, preferably using a non-contact approach. Various methods have been used...
for non-contact contour measurement of blanks, such as a laser triangulation or a confocal type of depth gauge. A typical output plot from such a system is shown in Fig 5, which in this case includes a least-squares fit to a circular arc. In the standard process, this fitted curve is used to calculate contour radius. This plot illustrates that the blank is also beveled on the contoured side.

In the measurement shown here, the blank is inserted into a fixture with a flat upper surface and an accurately machined pocket in which the quartz disk is placed. The flat surface provides a reference line in the plot, as well as reference edges that indicate the perimeter of the blank. After mathematically compensating for the slope of the reference surface, the center of the fitted curve relative to the pocket edges should represent the contour offset.

In practice, this method cannot differentiate between contour offset and physical tilt of the blank (for example due to a particle under one side of the disk), identifying the blank edge is difficult, and to characterize a blank fully, multiple scans are required. So although it is a potentially useful method, the inherent sources of inaccuracy need to be considered.

C. Use of optical properties of quartz - birefringence

A property of quartz that can be useful for various measurement techniques throughout crystal manufacture is its anisotropic optical characteristic of birefringence, a property that is exhibited to a varying degree by all transparent media with non-cubic crystalline structures. Birefringence is characterized by a refractive index that depends on the propagation direction or polarization direction of light passing through it. The simplest form of birefringence is described as uniaxial, which means that rotation about one axis does not affect the passage of light passing through the medium. This single axis is called the optic axis, and light for which the polarization direction is perpendicular to the optic axis is called an ordinary ray, and it exhibits a refractive index of \( n_O \). Light with a polarization direction parallel to the optic axis is called an extraordinary ray, and its refractive index is denoted \( n_E \).

Quartz has three axes of two-fold symmetry and one axis of three-fold symmetry; this form of crystal structure is classed as having trigonal symmetry, and materials such as quartz with this form exhibit uniaxial birefringence. The axis of three-fold symmetry is conventionally denoted the Z-axis, and this is the optic axis for quartz. The two discrete refractive index values at various wavelengths through and beyond the visible range are shown in table I [17].

<table>
<thead>
<tr>
<th>( \lambda (\text{nm}) )</th>
<th>( n_O )</th>
<th>( n_E )</th>
</tr>
</thead>
<tbody>
<tr>
<td>231</td>
<td>1.6140</td>
<td>1.6256</td>
</tr>
<tr>
<td>340</td>
<td>1.5675</td>
<td>1.5774</td>
</tr>
<tr>
<td>394</td>
<td>1.5585</td>
<td>1.5681</td>
</tr>
<tr>
<td>434</td>
<td>1.5540</td>
<td>1.5634</td>
</tr>
<tr>
<td>508</td>
<td>1.5482</td>
<td>1.5575</td>
</tr>
<tr>
<td>589</td>
<td>1.5442</td>
<td>1.5534</td>
</tr>
<tr>
<td>768</td>
<td>1.5390</td>
<td>1.5479</td>
</tr>
<tr>
<td>833</td>
<td>1.5377</td>
<td>1.5466</td>
</tr>
<tr>
<td>991</td>
<td>1.5351</td>
<td>1.5439</td>
</tr>
<tr>
<td>1159</td>
<td>1.5328</td>
<td>1.5415</td>
</tr>
</tbody>
</table>

The measurement method uses full-spectrum white light in transmission through the sample, and two linear polarizing filters are placed above and below the blank being measured. The polarizing filters are set up with the polarization directions at right angles to each other so that the background is dark. The plate orientations of any of the rotated cuts that are typically used as resonators have components along both the optic axis and perpendicular to it, so light passing through the quartz will experience two distinct velocities, as defined by the two refractive indices. The resulting effect is a rotation in the polarization of the light that is a function of the blank thickness and the wavelength of the light, and this causes a range of colors to be observed in the transmitted light.

To quantify the effect, consider the graphic in Fig 6.

The separation distance \( \Delta \) between wavefronts of the two rays is given by:

\[
\Delta = c(t_E - t_O)
\]

where \( c \) is the speed of light in free space and \( t_E \) and \( t_O \) are the propagation times for the two rays, so
\[ \Delta = c \left( \frac{1}{v_E} - \frac{1}{v_O} \right) = h \left( \frac{c}{v_E} - \frac{c}{v_O} \right) \]

where \( h \) is the thickness of the medium and \( v_E \) and \( v_O \) are the velocities of the rays.

But \( c/v \) is the refractive index \( n \) of the material, so

\[ \Delta = h(n_E - n_O) \]

The absolute difference \( n_E - n_O \) between the two refractive indices is the definition of the birefringence of the material, and for quartz at a typical visible wavelength, \( \Delta \approx 0.009 h \)

The resulting image exhibits bright lines where the value of \( \Delta \) corresponds to integral wavelength multiples for the wavelength of the light being passed, but because of the range of wavelengths in the visible spectrum, a color pattern is observed, as shown in the Michel-Levy birefringence chart in Fig 7.

For the 10MHz blank in question, the resulting first ring radius would be about 1.3mm, and this may work well in a metrology system, but in practice when using this method the contrast is fairly poor in the center region where the required accuracy is highest. It works very well for blanks with steeper contours, and can be used on blanks without polished surfaces. A 10MHz fundamental blank with a contour of 10D would exhibit a ring with a radius of 0.4mm.

D. Newton’s rings using monochromatic light

Another simple method to view incremental thickness variations of a plate of a transparent medium is to expose the plate to monochromatic light from one side, normal to the plate. If both surfaces are specular, the light is reflected from the top and bottom surfaces, resulting in interference patterns known as Newton’s rings. The method is often used on convex lenses in conjunction with an optical flat, where the lens is placed such that there is an air gap between the flat and the part being observed, but, as in this case, it also works for thin transparent lenses where the reference surface is the other face of the lens.

In this measurement method, the thickness increments between light and dark bands are directly related to the wavelength of the light being used. In our configuration the light source was a low-pressure sodium lamp that has a spectral line pair at 590nm and 590.6nm. The thickness increment \( \varepsilon \) is given by:

\[ 2\varepsilon = \frac{\lambda}{n} \left( m - \frac{1}{2} \right) \]

where \( \lambda \) is the light wavelength in free space, \( n \) is the nominal refractive index of the medium and \( m \) is the ring number. The radius of the observed rings \( r_m \) is then given by:

\[ r_m = \sqrt{\frac{R\lambda}{n} \left( m - \frac{1}{2} \right)} \]

In the crystal design investigated in this study, the first ring would occur at a radius of about 0.25mm, which gives very good resolution of the contour location. The set up to measure blanks using this method employs a sodium lamp from which the light is directed through a beam splitter to the part being tested, and then the image is viewed through a low-power microscope using a Scienscope Smartcam; this system provides a very simple method for measurement of the distance between two circles. The main disadvantage of this technique is that it requires highly polished surfaces.

IV. RESULTS

A group of 50 finished, encapsulated QRM resonators was measured in three axes for g-sensitivity using the passive method described in 2003 [18]. They were selected from past groups based on good performance in other respects such as Q and C1, to avoid otherwise anomalous parts being included in the population. The packages were opened, the blanks removed and the gold electrodes were stripped with aqua regia, taking great care to retain the individual identity of each blank.

The group was then measured for contour concentricity in magnitude and direction, and the scatter plot between contour offset and the magnitude of the g-sensitivity vector is shown in Fig 8. Although both g-sensitivity and contour offset were measured as vector quantities, the best correlation was obtained between the magnitudes of the two parameters as shown. Some scatter can be expected in this plot because the magnitudes and not the directions are taken into account for either parameter. Also other mounting irregularities were not taken into account.
consideration, but nevertheless, the overall trend is very conclusive.

V. CONCLUSION

Several methods have been identified that can be used to measure the concentricity of the contour shape in a quartz resonator blank. An experiment was set up on a 10MHz 3rd overtone QRM product to compare the contour offset with the measured values of acceleration sensitivity. As expected for this product type, there is a clear correlation between the two parameters. The positive result has prompted a series of developments to improve contour machining processes, and has resulted in significant improvements in product quality and yield.

ACKNOWLEDGMENTS

The author would like to thank Rick Puccio at Quartzdyne for many very helpful discussions and pointers to obtain meaningful simulation results from the Comsol Multiphysics finite element analysis system, to Bharat Desai at Vectron Cincinnati for his meticulous work in preparing and maintaining the samples, and to Todd Palmer at Vectron Hudson for his support through the project.

REFERENCES


Anchor Loss Suppression using Butterfly-Shaped Plates for AlN Lamb Wave Resonators

Jie Zou1*, Chih-Ming Lin1, and Albert P. Pisano2

1Department of Mechanical Engineering
University of California at Berkeley
Berkeley, CA, USA
*E-mail: jiezou@berkeley.edu

2Department of Mechanical and Aerospace Engineering
Department of Electrical and Computer Engineering
University of California at San Diego
San Diego, CA, USA

Abstract—The use of butterfly-shaped thin plates, formed by reducing the tether-to-plate angle, can raised the quality factor (Q) of aluminum nitride (AlN) Lamb wave resonators (LWRs) by eliminating the anchor loss. The finite element analysis (FEA) simulation results show that the butterfly-shaped plate can efficiently keep the vibration far from the edges at the tether-to-plate plane, so that the acoustic wave leaky through the supporting tethers is reduced. Specifically, the rounded butterfly-shaped resonators show more efficient suppression in the anchor loss compared to the beveled butterfly-shaped resonators. The measured frequency response for a 863-MHz AlN LWR with 45° beveled tether-to-plate transition yields a Q of 1,979 which upwards 30% over a conventional rectangular resonator; another AlN LWR on the butterfly-shaped plate with rounded tether-to-plate transition yields a Q of 2,531, representing a 67% improvement.

Keywords—Lamb wave resonators; quality factor (Q); anchor loss; aluminum nitride (AlN); piezoelectric resonators; RF MEMS; butterfly-shaped plate

I. INTRODUCTION

The recent demand for highly-integrated and low-loss band-pass filters and oscillators for the on-chip radio frequency (RF) front-ends has led current research efforts towards making the CMOS-compatible and high-Q MEMS resonators [1], [2]. Among various microelectromechanical resonator technologies, aluminum nitride (AlN) Lamb wave resonators (LWRs) have attracted much interests since they show CMOS technologies, aluminum nitride (AlN) Lamb wave resonators (LWRs) have attracted much interests since they show CMOS compatibility, high frequencies (\(f_0\)), low motional impedances (\(R_m\)), and moderate coupling coefficient (\(k^2\)) [3]–[6]. However, the piezoelectric LWR usually shows a moderate Q due to the anchor loss [8]–[13] and the interfacial loss [14], [15] so an improvement in the Q of the piezoelectric AlN LWRs is highly desirable to further enable the low-loss filters and low-phase-noise oscillators.

There is a clear evidence that a large portion of mechanical energy dissipation via the support tethers of the LWR, and this kind of energy dissipation is usually called anchor loss or tether loss [8]–[13]. By suppressing the acoustic wave leakage through the tethers, the Q of the AlN LWR can be effectively increased [8]–[13].

In general, the anchor loss in the MEMS resonators can be minimized by various design approaches, such as placing the supporting tethers at the nodal locations of the resonance mode, increasing the acoustic impedance mismatch between the support anchors and vibrating structure, or designing the tether length with an odd multiple of a quarter-wavelength (\(\lambda/4\)) [1], [16]. Another way to reduce the anchor loss is to change the resonator geometry itself to concentrate the mechanical displacements far from the supporting tethers [8]–[13]. For example, the suspended biconvex edges of the AlN Lamb wave resonator were demonstrated to enable high efficient concentration of the displacement distributions in the resonance body and the Q was then boosted. However, some unwanted spurious modes are induced in the AlN plate due to the convex free edges [9].

Fig. 1(a) illustrates a conventional AlN LWR employing the orthogonal tether-to-plate transition and with one pair of interdigital transducer (IDT) and two straight suspended free edges as the acoustic wave reflectors. To effectively minimize the anchor loss and not sacrifice the other performance such as effective coupling, a AlN LWR utilizing a beveled butterfly-shaped plate is recently investigated to reduce anchor loss [10]. As shown in Fig. 1(b), a butterfly-shaped AlN LWR composed of the same IDT finger electrode configuration and straight free-edges, but with a beveled tether-to-plate transition has been demonstrated to enable lower mechanical displacement in the tethers and higher anchor Q (\(Q_{anchor}\)) [10]. For the first time, a rounded butterfly-shaped AlN LWR with a rounded tether-to-plate transition angle of 45°.
II. Resonator Design and Modeling

A. The PML-based Finite Element Analysis

Recently, a new finite element analysis (FEA) simulation approach using a perfectly matched layer (PML) technique is introduced to evaluate the anchor dissipation in the MEMS resonators [11], [17], [18]. The PML can attenuate the acoustic energy leaky via the tethers into the substrate and also perfectly match the rest of the domain so that there is no acoustic wave reflected from the substrate/PML interface. We employ the commercial FEA software, COMSOL Multiphysics®, and adopt the PML-based FEA approach to simulate the $Q$ anchor in the AlN LWRs based on different plate geometries.

Fig. 2 (a) indicates the position of the tether-to-plate plane that we investigate the displacement fields in the FEA model. A half AlN plate with one simple beam tether attaching to a substrate layer covered by PMLs and the mesh adopted in FEA.

B. The Displacement in the Tethers

The anchor loss is directly proportional to the displacement fields in the tether. In the conventional AlN LWR, there are relatively strong displacements delivered to the tether from the resonant plate, so that a large portion of mechanical energy is dissipated through the tethers, resulting in a low $Q_{\text{anchor}}$. In this work, two butterfly-shaped AlN plates are proposed to reduce the displacements in the tethers when the AlN resonator is in resonance. In order to study the effect of the plate shape on the displacements in the tethers, the LWRs on AlN plates utilizing the rectangular plate, the beveled butterfly-shaped plate, and the rounded butterfly-shaped plate are designed and compared as summarized and in Table I.

The resonance displacement profiles on the top edge of the tether-to-plate plane at the $S_0$ mode resonance for the three designs of AlN LWRs.

![Fig. 3. Displacement profiles at the $S_0$ mode resonance of the tether-to-plate plane for the AlN LWRs with the (a) orthogonal (b) beveled, and (c) rounded tether-to-plate transitions.](image)

![Fig. 4. Simulated displacement profiles on the top edge of the tether-to-plate plane at the $S_0$ mode resonance for the three designs of AlN LWRs.](image)

Table I: Geometric dimensions of the AlN Lamb wave resonators.

<table>
<thead>
<tr>
<th>Orthogonal</th>
<th>Beveled</th>
<th>Rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDT finger electrodes</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>IDT aperture</td>
<td>180 $\mu$m</td>
<td>180 $\mu$m</td>
</tr>
<tr>
<td>IDT electrode width</td>
<td>3 $\mu$m</td>
<td>3 $\mu$m</td>
</tr>
<tr>
<td>IDT electrode thickness</td>
<td>200 nm</td>
<td>200 nm</td>
</tr>
<tr>
<td>Tether-to-plate angle</td>
<td>90°</td>
<td>45°</td>
</tr>
<tr>
<td>Tether length</td>
<td>33 $\mu$m</td>
<td>33 $\mu$m</td>
</tr>
<tr>
<td>Tether width</td>
<td>8 $\mu$m</td>
<td>8 $\mu$m</td>
</tr>
<tr>
<td>AlN plate length</td>
<td>210 $\mu$m</td>
<td>252 $\mu$m</td>
</tr>
<tr>
<td>AlN plate width</td>
<td>78 $\mu$m</td>
<td>78 $\mu$m</td>
</tr>
<tr>
<td>AlN plate thickness</td>
<td>4.0 $\mu$m</td>
<td>4.0 $\mu$m</td>
</tr>
</tbody>
</table>

---

The displacement profiles indicate that the butterfly-shaped design can efficiently reduce the vibration displacement in the plane center so the mechanical energy dissipation through the tethers is reduced. The rounded butterfly-shaped plate shows a more efficient suppression on the displacement in the center of the tether-to-plate plane. More specifically, Fig. 4 depicts the total displacements on the top edge of the tether-to-plate plane. It is clear to observe that the beveled butterfly-shaped plate significantly reduces the displacement at the plane center, and the rounded one offers a better efficiency on the displacement suppression at the plane center where the tether locates.
C. The Quality Factor

As is well known, the mechanical $Q$ can be generally expressed as [10]:

$$Q = 2\pi \frac{E_{\text{stored}}}{E_{\text{dissipated}}}$$

(1)

where $E_{\text{stored}}$ is the vibration energy stored in the resonator and $E_{\text{dissipated}}$ denotes the energy dissipated per cycle of vibration, respectively. It is clear that to obtain a higher $Q$, it requires less energy dissipation. In this work, the anchor loss is assumed to be the main contributor of various energy dissipation sources so the total quality factor increases with a higher $Q_{\text{anchor}}$, which can be obtained by using following equation in the PML-based FEA simulation [11], [13], [17]:

$$Q_{\text{anchor}} = \frac{\text{Re}(\omega)}{2 \text{Im}(\omega)}$$

(2)

where $\omega$ is the eigen-frequency of the $S_0$ Lamb wave mode solved in the FEA simulation.

Fig. 5 presents the mode shapes of the $S_0$ Lamb wave at resonance in the AlN plates employing different tether-to-plate transition designs. As shown in Fig. 5 (a), there is obvious vibration displacement in the support tether of the orthogonal plate, indicating a large part of mechanical energy loss via the tether. The $Q_{\text{anchor}}$ equals 7,910 which is obtained from the PML-based FEA approach. As illustrated in Figs. 5(b) and (c), the displacement field in the supporting area is less than that in the rectangular plate. In addition, the rounded butterfly-shaped plate shows the most effective displacement suppression and the vibration displacement in the tether is the minimum. The simulated $Q_{\text{anchor}}$ of the $S_0$ mode in the butterfly-shaped AlN plate using the beveled transition is increased to 16,605, and the one utilizing the rounded transition is improved to 43,652. Interestingly, the butterfly shape employing the rounded tether-to-plate transition can effectively suppress the displacements occurring in the tether and further shows a 5.52× improvement in $Q_{\text{anchor}}$.

III. Fabrication Process

The micro-fabrication process used to make the resonators is a two-mask process [10]. First, a 300-nm-thick low-stress nitride (LSN) layer for electrical isolation was deposited on the high-resistivity silicon wafer. A highly $c$-axis oriented AlN film was reactively sputtered onto the LSN layer. A 20-nm-thick chromium (Cr) was sputtered on the AlN thin film as the adhesion layer and then a 180-nm-thick platinum (Pt) IDT electrodes layer was sputtered on the Cr layer and patterned using a lift-off process. A low temperature oxide (LTO) hard mask layer was deposited and the dry etching process was used to pattern the LTO layer and AlN thin film. Finally, the AlN LWRs were released using XeF$_2$-based isotropic dry etching of the Si substrate. Fig. 6 shows the scanning electron micrograph (SEM) images of the three AlN Lamb wave resonators on 4.0-$\mu$m-thick AlN plates.

IV. Experimental Results and Discussions

To diminish the micro-fabrication process variations, all the devices were fabricated on the same wafer and placed in the vicinity. They were tested in air at room temperature and $S_{11}$ parameters were extracted using an Agilent E5071B network analyzer. The measured $Q$ was extracted from the admittance plot by dividing the resonance frequency ($f_0$) by the 3dB bandwidth.

The measured $Q$’s of the LWRs based on the butterfly-shaped AlN plates are consistently higher than the rectangular resonators with the orthogonal tether-to-plate transition. Fig. 7
Fig. 7. Measured admittance spectra of the AlN LWRs using the orthogonal, beveled, and rounded tether-to-plate transitions.

presents one set of the one-port admittance response spectra for the LWRs using orthogonal, 45° beveled and rounded tether-to-plate transitions. The 863.3-MHz resonator on the butterfly-shaped plate with a beveled tether-to-plate angle of 45° yields a $Q$ of 1,979, upwards 30% over a conventional rectangular resonator with a $Q$ of 1,518. Moreover, the rounded butterfly-shaped AlN plate shows a measured $Q$ of 2,531, representing a 67% improvement in the $Q$. The employment of the butterfly-shaped AlN plates successfully suppresses the displacement in the supporting area while the LWRs are in resonance, offering reduction of the anchor loss and the improvement of the $Q$.

The effective coupling ($k^2_{\text{eff}}$) of the resonators on the butterfly-shaped plates are slightly lower than that in the conventional AlN LWR, but it is still sufficient for oscillators and narrowband filters. In addition, as it was predicted in the FEA simulations, both the beveled and rounded butterfly-shaped AlN plate neither introduces any other spurious mode nor shifts the resonance frequency $f_r$ significantly.

V. CONCLUSIONS

A new design approach to reduce the anchor loss in the AlN LWRs is presented in this work. The employment of the butterfly-shape AlN plates can sufficiently suppress the resonance displacement in the center of the tether-to-plate plane and the energy loss via the tethers is reduced. The beveled butterfly-shaped LWR with a tether-to-plate angle of 45° presents a $Q$ of 1,979, showing 30% increase over the conventional rectangular resonator. The LWR on the rounded butterfly-shaped AlN plate shows a measured $Q$ up to 2,531, representing a 67% improvement. The PML-based FEA model also confirms that the rounded butterfly-shaped AlN plate can more efficiently eliminate the anchor dissipation and boost the $Q_{\text{anchor}}$.

ACKNOWLEDGMENT

The authors would like to offer special thanks to the staff at the Berkeley Marvell Nanofabrication Laboratory.

REFERENCES

Ultra-Low Noise All Fiber Mode-Locked Laser

Yaolin Zhang, Quansheng Ren, Shuangyou Zhang, Dong Hou, and Jianye Zhao*
School of Electronics Engineering and Computer Science, Peking University, Beijing, China, 100871
zhaojianye@pku.edu.cn

Abstract — A new design of all fiber-based, mode-locked laser with ultra-low phase noise of -158 dBc/Hz at 1 MHz is reported. Both high repetition rate optical pulses and microwave with ultra-low phase noise were generated using an injection locked actively mode-locked fiber laser. The obtained microwave showed much lower phase noise than that of the equivalent commercial synthesizer. As the laser used is all fiber-based, our design is a cost-effective solution for low noise microwave, and high repetition rate optical pulse train generators.

Keywords—mode-locked laser; phase noise; timing jitter; frequency detuning

I. INTRODUCTION

High repetition rate optical pulse generators with low phase noise are becoming more and more popular in many applications such as high-speed communication, arbitrary waveform generation, high-speed and high-resolution optical analog-to-digital conversion, and photoelectronic neural computing [1-5]. Such applications require multi- or tens of gigahertz microwave with very low phase noise to achieve desired system performance that are hard to be implemented by pure electrical means. Mode-locked lasers are ideal optical pulse generators and are recognized as the purest microwave generators [6]. Passive mode-locked laser can generate optical pulse train with ultralow phase noise [7], but the repetition rate is limited by the cavity length. Actively mode-locked laser, however, can be mode-locked at much higher RF frequencies that are multiple of the reciprocal of the cavity roundtrip time, and the absolute phase noise performance of an actively mode-locked laser is usually limited by the RF driving source [8].

In the past two decades, many researches about increasing the repetition rate and reducing the phase noise of the actively mode-locked lasers such as harmonically mode-locking technique [9-10], short cavity technology [11, 12] and using very low-phase noise RF sources [8, 13] have been carried out. However, harmonically mode-locked Er-doped fiber lasers suffer from the unequal amplitude of the optical pulses [10]. Short cavity mode-locked lasers, although can be mode-locked at high repetition rate, but the noise performance deteriorates as the quality factor of cavity reduces when the cavity length shortens [12]. The previously reported ultralow-jitter actively mode-locked lasers based on an ultralow noise sapphire loaded cavity oscillator (SLCO) [8, 13] show excellent noise performance, but the SLCO is not cost effective, and it is reported that fiber based mode-locked lasers show better performance than semiconductor based mode-locked lasers [14], it would be meaningful to realize an ultralow phase noise microwave and high repetition rate optical pulse generator based on cascaded passively and actively mode-locked fiber lasers.

In this letter, we propose a new design of ultra-low noise and high repetition rate mode-locked fiber lasers, which utilizes a passively mode-locked fiber ring laser (PMLL) to generate very low phase noise optical pulses, and then extract high harmonic from them to drive an actively mode-locked fiber ring laser (AMLL). By this scheme, a ~1.117 GHz optical pulse train and microwave with -158 dBc/Hz of low phase noise at 1 MHz offset frequency were obtained. To our knowledge, this is the first time that a very cost effective way of generating ultralow noise, high repetition rate optical pulse train and microwave at the same time based on a cascaded fiber ring laser was explored.

II. EXPERIMENTAL SETUP

A schematic of the cascaded fiber ring laser is illustrated in Figure 1. The RF source was generated by the PMLL, which is an Er-doped fiber (EDF) ring laser based on nonlinear polarization rotation (NPR). The medium gain was provided by a 0.4 m highly Er-doped fiber. To ensure the facility of adjusting the repetition rate, we used spatial optical components including collimators, wave plates and polarizing
beam splitter (PBS). The cavity length can be easily adjusted. The PMLL operates at a repetition rate of ~74.5 MHz, corresponding to a cavity length of ~2.68 m. Stable mode-locking can be obtained by rotating the wave plates. ~3dBm power of the optical pulses was divided by the PBS and injected into a fast photodiode (PD) with 3.2 GHz bandwidth and responsivity of 0.8 A/W.

The AMLL was mode-locked at ~1.117 GHz via gain modulation of an intensity modulated electro-optical modulator (EOM, KG-DDMZ1510PS). The EOM has a bandwidth of 10 GHz and a half-wave voltage of 2.6 V. The cavity length of the AMLL was set at about ~14.8 m, corresponding to a fundamental repetition rate of ~13.46 MHz. A 1.2 m highly Er-doped fiber was used for laser pump and an isolator was used to guarantee unidirectional light propagation in the cavity. The EOM was biased at 0 V, and approximately 5 dBm of RF power from the LNA was applied to the EOM. Before the EOM, a fiber stretching based three-paddle polarization controller (PC, Thorlabs FPC561) was used to restrict the polarization state of the light entering the EOM, each paddle acted proximately as a λ/4, λ/2 and λ/4 wave plate respectively in the designed wavelength. The laser diode generates ~600mW maximum optical power at 980 nm and all of the power was injected into the cavity via a 980nm/1550nm wavelength division multiplexer. When the cavity is closed, stable mode-locking can be easily obtained by rotating the paddles of the PC. ~10% power of the optical pulses was coupled to high speed PD2 through a fiber coupler for signal detection and phase noise measurement.

III. EXPERIMENTAL RESULTS AND DISCUSSION

The frequency spectrum of the response signal from the PD1 is illustrated in Figure 2(a). It shows that the original signal converted by PD had a power of ~12 dBm at 1.117 GHz. This signal was then filtered and amplified to ~5 dBm, whose spectrum is illustrated in Figure 2(d), and was used to drive the AMLL. Here we choose the 15th harmonic of repetition rate of the PMLL because according to Figure 2(a), the power of the 15th harmonic was much higher than those of the higher order harmonics and could be easily amplified to achieve stable active mode-locking with only one low noise amplifier (LNA). All components of the PMLL were put in a thermal enclosure box to achieve stable mode-locking and to reduce the phase noise of the PMLL, especially at low offset frequencies.

Figure 2(a) shows the frequency spectrum of the PD1’s response signal, ranging from DC to 3 GHz. As the bandwidth of the PD is only 3.2 GHz, the response over 3 GHz will be much weakened. Nevertheless, we still observed over ~25 dBm RF power at the 40th harmonic component that is ~2.98 GHz. Such amplitude of RF power sources can be easily amplified to above 0 dBm with only one LNA, therefore, we are confident that the PMLL can provide much higher frequency low noise RF sources by adjusting the repetition rate of the PMLL and employing narrow band-pass filters and LNAs operate at higher frequency. Figure 2(b) shows the frequency spectrum of the PMLL at its repetition rate. It shows the power of signal is 50 dB higher than that of the noise and its line width is narrow. Figure 2(c) and Figure. 2(d) show the power of 1.117 GHz RF signal before and after amplified respectively, the signal-to-noise ratio is not as good as that of the repetition rate, but still has ~40 dB, and the noise is not increased significantly in the amplification. The results indicate that it is feasible to use a PMLL for low phase noise, high frequency microwave generating.

An average optical output power of ~8.5 dBm was observed, which was converted to the low noise RF source with power of 2.15 dBm at 1.117 GHz by PD2. Such high power level of RF can be directly used for phase noise measurement, which eliminates any necessity of using external components for amplification, and avoids the possible noise performance loss. The RF signal by PD2 is directly supplied to an advanced noise analyzer (R&S FSUP-26).

Figure 3 shows the absolute single side band phase noise of the microwave generated by the PMLL and the AMLL. The blue/solid line represents the phase noise of the amplified 15th harmonic of the PMLL, and the red/solid line represents the original noise performance of the microwave from the AMLL, detected by PD2. The amplified RF source from the PMLL exhibits ~15 dB/decade phase noise roll-off from 1 Hz to 1 kHz and ~20 dB/decade from 1 kHz to 100 kHz. After that, it reached the lowest noise floor around -150 dBc/Hz. We can see that when offset frequency is lower than 2 kHz, the low noise characteristic of the RF driving source is perfectly delivered to the AMLL. When offset frequency is greater than 2 kHz and lower than 400 kHz, the absolute phase noise is a little worsened. We believe the noise deterioration was caused by the EOM, which is only component has electric input in the loop. After that, the noise level drops quickly, and the lowest phase noise reached ~166 dBc/Hz, which is better than the previous reported ultralow noise mode-locked laser [8]. For comparison, a noise performance of a commercial frequency synthesizer (Agilent N5180) operating at the same frequency is also added, which represents the typical performance of electric RF oscillators available on the market. The figure shows that
despite the noise peaks at 100 kHz offset frequency, the noise performance of our scheme is superior to that of the electric oscillator.

As we used a fiber-based PMLL to generate a low noise RF source, the phase noise at low offset frequency can be improved with many proven techniques such as air flow control, thermal stabilization and frequency locking the repetition rate to an external low noise frequency source at tens of megahertz such as the temperature compensation x’tal oscillator, which is cheap and easily available on the market. Then, we can deliver the low noise to high frequencies though the mechanism of mode locking, rather than the electric phase locked loop technique. As a result, the least noise performance loss can be achieved.

The highest operating frequency of our scheme can be calculated with the highest repetition low noise fiber-ring-based PMLL reported multiplying the highest order of the harmonic that can be amplified to drive the AMLL. Since the power of the harmonics is highly dependent on the pulse shape when the bandwidth and responsivity of the PD are sufficiently high, the highest operating frequency is determined by the pulse shaping. The pulse shaping, however, can be controlled by adjusting the dispersion of cavity. Besides, recently developed new structure EDF-based fiber ring laser has achieved from multi-gigahertz to even terahertz mode-locking [16], there will be no obstacle for our scheme to achieve the state-of-art noise performance in the future.

IV. CONCLUSION

We have proposed a new scheme to generate ultra-low noise optical pulses and microwave at the same time. A 1.117 GHz microwave with -158 dBc/Hz of low noise at 1MHz offset frequency and 2 dBm of power was obtained without using of extra high frequency oscillators. To our knowledge, this is the first time that a low noise optical pulse train generator based on all fiber mode-locked laser had been explored. The repetition rate, limited by the experiment conditions, can be much improved by replacing high frequency components. By this scheme, both the low phase noise delivery by mode-locking mechanism of the PMLL and phase noise filtering effect of the AMLL can be easily utilized.

REFERENCES

Digitally Temperature Compensated SAW Oscillator Based on the New Excitation Circuit

Alexei N. Liashuk*, Sergey A. Zavyalov, Aleksandr N. Lepetaev, Anatoliy V. Kosykh, Igor V. Khomenko
Omsk State Technical University
The Chair of Radio Technical Devices and Diagnostic Systems
Omsk, Russian Federation
e-mail*: vostok3@front.ru

Abstract—New application of 4-pole SAW component with small losses for the realization of high-frequency voltage-controlled SAW oscillator is described. In contrast to the known circuit solutions of SAW resonators-based oscillators, a new use of radio frequency SAW filter allows to realize the schematic of Colpitts oscillator with noticeably large continuously frequency tuning range than that of counterparts. Proposed oscillator schematic is used to realize digital temperature compensated SAW generator.

Keywords—SAW resonator; SAW delay line; SAW filter; oscillator; excitation circuit; temperature compensation

I. INTRODUCTION

Typically, SAW elements-based oscillators are realized in one of the three ways. Depending on the mode of interaction of the SAW element with active element oscillators can be divided into:

- oscillators with SAW delay line included in the feedback – these are usually a tunable two-stage schemes (two or more active elements) for compensating large (approximately 10 dB or more) loss of wideband SAW delay line;
- oscillators with the quartz substrate-based SAW resonators used as 4-pole to form a positive feedback loop - oscillators are characterized by a low frequency tuning;
- oscillators with SAW resonator used as a 2-pole element – these are similar to the oscillator circuits with the piezoelectric bulk wave resonator (quartz resonator), characterized by the simplicity of circuit solutions, low frequency tuning.

Often SAW resonator-based oscillators are realized for fixed frequencies without being able to make any frequency tuning at all. An embodiment of such a generator is a single-input SAW resonator oscillator US 5721515.

Analysis of SAW oscillators embodiments showed that broadband SAW filter components with low losses in the passband, usually defined as RF filter for the input circuits of receivers or low-loss SAW filters, in relatively simple single-stage circuit configurations are not used [1].

II. THE OSCILLATOR REPRESENTATION

To analyze the properties of the oscillator we will use the representation of the oscillator with negative input resistance [2]. In contrast to the well-known representations, for example, the method proposed in the works by Alechno [3], supposing feedback loop break and the oscillator representation using transfer model described with harmonically linearized S-parameters, the proposed approach based on the extraction of active and passive parts eliminates the need to enter the unobvious "virtual ground".

No matter how many components there are in passive and active parts of the oscillator, passive and active parts of oscillator circuits are considered as linear 2-poles and described with certain characteristic parameters. In this view (Fig. 1), the properties of a passive oscillation circuit and an active circuit can be described by relying only on the ratio between the signals on the external pins, i.e. using a macro model. Properties of the oscillation circuit can easily be expressed in terms of its linear macro model.

![Fig. 1 2-pole oscillator representation](image)

For 2-pole representation active circuit can be completely characterized with complex impedance averaged over the first harmonic (for series resonant circuit) or averaged over the first harmonic of the complex conductivity (for parallel resonant circuit). The condition of self-excitation of the scheme in this case is the energy excess released by an active two-pole, above the losses energy in an oscillation circuit with low (zero), the amplitude of the oscillations.

In Fig. 2, in accordance with the model of 2-pole representation results of experimental investigation of active
part of Colpitts oscillator (Fig. 1) by measuring its negative resistance depending on the amplitude of the excitation is given.

Reactive component of the input negative resistance does not depend on the amplitude of the excitation and has a capacitive character.

To make oscillations start in the circuit, passive part impedance must be inductive with the value of real component of the impedance in the range approximately minus (50..125) ohms. Fig. 3 shows a new schematic of the passive part based on a broadband SAW filter component with small losses, which satisfies the requirements [4].

In order to guarantee frequency stability SAW filter should be selected (or designed) with linear phase response in the pass-band with the phase shift no more than $2\pi$. The pass-band value is a compromise between the requirements to spectral purity of the output oscillations and the necessary to provide frequency tuning.

The SAW filter is a transversal non-minimum phase device that allows to design independently frequency and phase responses of different complex shape, such as symmetrical response and linear phase (Fig. 4). The bandwidth and the amount of phase shift in the band determine the frequency stability and spectral purity of the output oscillation. Thus, in Fig. 4 (a) and 4 (c) shows the characteristics of a SAW device and unambiguous change of the linear phase characteristic in the pass band, while the SAW structure in Fig. 4 (c) provide greater frequency stability of the oscillator than the structure in Fig. 4 (a) because of the lower bandwidth. At the same time in the SAW oscillator based on structure with the parameters shown in Fig. 4 (b) the conditions of phase and amplitude balance can be satisfied at several frequencies.

Such ambiguity of the phase response in the filter pass-band as shown in Fig. 5 when loaded with voltage controlled LC-circuit (Fig. 3) can lead to an excitation at the parasitic (unwanted) frequency or failure to generate a desired frequency while tuning.

Fig. 6 The real and imaginary part of the excitation circuit impedance when SAW filter has characteristics shown in Fig. 5

To exclude the possibility to generate at spurious frequency or failure of generation in general, as noted above, it is necessary to ensure the unambiguity of phase response, that is
to guarantee in the pass-band the phase shift of not more than $2\pi$. Fig. 7 shows an example of such a phase response of the SAW filter, and Fig. 8 shows the measured real and imaginary parts of the impedance of the cascaded SAW filter and LC-circuit.

![Fig. 7 SAW filter with a unambiguous phase response (curve 1) in the passband filter (curve 2)](image)

It can be seen in Fig. 8 that in a wide frequency band there is only one distinct peak value of real and imaginary part of the impedance near which (frequency 171,95 MHz) oscillation is possible.

To confirm the effectiveness of the proposed solutions a model of the passive part of the oscillator with the 4-pole SAW component with small losses was developed. New excitation circuit characterized in that one pair of inputs is used for controlling the 4-pole complex impedance with inductive character, and the other pair of inputs - for connection to the active part of the Colpitts-type oscillator circuit. The proposed approach provides a continuous change of inductive character at least 30% of the SAW component bandwidth.

In contrast to the known circuit solutions of SAW resonators-based oscillators, a new use of radio frequency SAW filter allows to realise the schematic of Colpitts oscillator with noticeably large continuously frequency tuning range than that of counterparts [4].

### III. PRACTICAL IMPLEMENTATION

With the aim to provide optimal microprocessor resources usage, in the microprocessor compensation unit a tabular method for approximating the frequency-of-temperature dependence it is implemented when the array of DAC and ADC codes are formed using the experimental data obtained in the temperature interval $\Delta T$, and additional values are obtained using the approximation (Fig. 9).

![Optimal method of temperature compensation from the point of view of efficiency, complexity (difficulty) of implementation is a piecewise linear approximation with the static information processing. According to the method, compensation function is implemented in the form](image)

$$\bar{U}_d = U_{oj} \times \left( \frac{T_{j+1} - T_{oj}}{T_{j+1} - T_j} \right) + U_{oj+1} \times \left( \frac{T_{oj} - T_j}{T_{j+1} - T_j} \right),$$

where $U_{oj}$ and $U_{oj+1}$ - compensation voltage at temperatures $T_j$ and $T_{j+1}$; $T_{oj}$ - environment temperature; $T_{oj}$ and $T_{oj+1}$ temperatures, corresponding to the lower and upper boundaries of the temperature interval $\Delta T$.

On the basis of the proposed approach the temperature compensated generator (Fig. 9) with operating frequency of 201,6 MHz was designed. Temperature compensation increased frequency stability of SAW oscillator more than 200 times. Measured typical temperature instability (Fig. 10) was no more than $\pm 35 \times 10^{-6}$ in the temperature range (-55..+55)°C with calculated mean square error of temperature instability in the temperature range no more than $3.2 \times 10^{-6}$.
Experimental studies of the SAW filter-based oscillator phase noise were carried out with an operating frequency of 171.6 MHz and a comparative analysis with the analogues. The following values for the oscillator: -123 dBc/10 kHz. For comparison - MEMS generators SiT8209 at a frequency of 156.25 MHz: and HT-MM900A) at 100 MHz are -122 dBc/10 kHz.

The key factors that influence the oscillator performance can be identified as: the bandwidth of SAW component, SAW component topology (structure [parallel channels, the resonance], the number of electrodes in the IDT and multistrip couplers), also electrical characteristics of the LC-tank affecting the value of the reflection coefficient and the standing wave mode.

It is seen to be perspective to implement a broad frequency tuning generator using MEMS technology [6].

The investigation in the article had been carried out under the program "Research and design on priority directions of scientific-technological complex development of Russia 2014-2020". The unique identifier of the agreement RFMEFI57414X0033.

REFERENCES


Phase Group Characteristics and Phase Coincidence Detection Based Phase Noise Measurement Method

Dong Shaofeng, Zhou Wei, Hu Wei, Zhan Jinsong, Qin Hongbo
School of Electro-Mechanical Engineering
Xidian University,
Xi’an, Shaanxi, China
Email: {shaofengdong, zjs8012, qgb}@126.com, {wzhou, whu}@xidian.edu.cn

Abstract—A novel phase noise measurement method is proposed based on phase group processing and phase coincidence detection. Using the regularity variation of the phase difference between frequency signals, phase group processing can be made without the normalization of the frequency signals. The phase group characteristics, such as phase group synchronization, and the regularity of phase coincidence detection between different frequency signals are revealed in this paper. Based on the phase group characteristics and phase coincidence detection, two adjacent phase coincidence points of the reference frequency signal and the measured frequency signal are used to generate the measurement gate. The fluctuation of the gate time between multiple consecutive measurement reflects the close-in phase noise. And the edge feature of phase coincidence detection fuzzy area can be used to obtain the far-out phase noise. Experiment results show that this method has the advantages in simpler structure, wider measurement band, smaller additional noise.

Keywords—phase noise; phase group processing; phase coincidence detection;

I. INTRODUCTION

Phase noise is an important parameter in time-frequency measurement and control. The available phase noise measurement methods[1][2], such as phase discrimination method and high-speed A/D sampling method[3], are founded on the phase processing between two signals at the same frequency. Frequency normalization is needed for the phase noise measurement between two different frequency value signals, which increases the system complexity and introduces additional phase noise. By analysising the phase difference variation regularity between two different frequency signals, the paper reveals the phase group synchronization and phase coincidence detection fuzzy area between every two cyclical signals. With these phase group characteristics, a wide-band and fast-response phase noise measurement method is put forward.

II. PHASE GROUP CHARACTERISTIC BETWEEN DIFFERENT FREQUENCY SIGNALS

Phase group processing is accomplished and phase group synchronization phenomenon is discovered by researching on the characteristics that appear between two different frequencies such as the greatest common factor frequency \( f_{\text{max}} \) [4], the least common multiple period \( T_{\text{min}} \), and the equivalent phase comparison frequency (EPCF) [5].

![Fig. 1: Phase group variation rule between different frequency.](image)

It is verified from the relations between \( T_{\text{min}} \), \( f_1 \) and \( f_2 \) that \( T_{\text{min}} \) is just equal to \( AT_1 \) or \( BT_2 \), which means each one \( T_{\text{min}} \) starts at the phase coincidence point of \( f_1 \) and \( f_2 \), and ends at the next phase coincidence point of them. We define the regular phase difference variation as phase group synchronization based on taking the least common multiple period \( T_{\text{min}} \) as a phase difference group. The phase group variation rule between different frequency is shown as Fig.1. The time fluctuation of phase difference group between multiple consecutive measurement reflects the close-in phase noise[6]. The measurement gate can be generated based on phase group characteristic, which can eliminate \( \pm 1 \) count error greatly.

III. PHASE COINCIDENCE DETECTION FUZZY AREA BETWEEN DIFFERENT FREQUENCY SIGNALS

Phase coincidence detection is used to generate the measurement gate. Two adjacent phase coincidence points are used as the start and stop signals of the gate, which are detected between the measured frequency and the reference frequency. The precision of phase coincidence detection has a big influence on the time interval measurement accuracy. In practical measurement, the uncertainty of measurement gate caused by the phase coincidence detection error would result in the measurement error. Because of the existence of phase noise and trigger error, the limited resolution of phase coincidence detecting circuits, the phase coincidence error is inevitable and hard to eliminate by circuit design. There is a fuzzy region which consist of lots of narrow pulses differing in amplitude in
traditional phase coincidence detection. The experiment observed the phase coincidence detection between two different frequency signals is not a singular pulse[7], but a pulse group containing millions of pulse, shown as Fig. 2.

Fig. 2: Phase coincidence detection pulse group observing from the oscilloscope.

To measure the far-out phase noise, a short gate time, such as several milliseconds, is required. However so short a measurement gate is difficult to generate. In experiment we found that there is a discontinuity in phase coincidence detection pulse group, and this discontinuity is just caused by the far-out phase noise. Combining the discontinuity in phase coincidence detection pulse group counting and data processing, the far-out phase noise can be obtained.

IV. EXPERIMENT

Combing the fluctuate of gate time and the discontinuity of the phase coincidence detection pulse, the SSB phase noise of the measured frequency can be obtained by FFT. Fig.3 shows the SSB phase noise of 10MHz generated by HP8662.

Fig. 3: SSB Phase noise of 10MHz generated by HP8662.

Compared with the available phase noise measurement methods, the method has the advantages in simpler structure, wider measurement band, smaller additional noise.

V. CONCLUSION

The phase group characteristics, such as phase group synchronization, and the regularity of phase coincidence detection between different frequency signals are revealed in this paper. Based on the phase group characteristics and phase coincidence detection, two adjacent phase coincidence points of the reference frequency signal and the measured frequency signal are used to generate the measurement gate. The fluctuation of the gate time between multiple consecutive measurement reflects the close-in phase noise. And the edge feature of phase coincidence detection fuzzy area can be used to obtain the far-out phase noise.

REFERENCES

Precise Measurement of Complicated Frequency Signals

Lina Bai, Meina Xuan, Yuzhen Jin, Bo Ye, Zhenjian Cui, and Wei Zhou
Department of Measurement and Instrumentation
Xidian University
Xi’an, Shaanxi, People’s Republic of China
lnbai@mail.xidian.edu.cn

Abstract—Based on the border effect and relevant theories, it is found that measurement precision depends on resolution stability which is much higher than resolution itself. With the help of this new theory and the more sensitive discrimination of the border of a fuzzy area, it is possible to greatly improve the measurement precision. Since there are a lot of complicated frequency signals in quantum frequency standards, telecommunication, fundamental subjects and other fields, precise measurement is significantly important. In a great deal of measurement of complicated frequency signals, discrete fuzzy areas are common. In this paper, we put forward a feasible and high-precision frequency measurement scheme to make it easier to capture border information, which shows great advantages of border effect.

Keywords—border effect; complicated frequency; discrete fuzzy area

I. INTRODUCTION

The frequency-measuring method is based on phase-change regularities between complicated frequency signals. Because of the higher stability of measurement resolution itself, the comparison characteristics between the object to be measured and the standard signal can more sensitively reflect different detection results inside and outside fuzzy areas at the border. Measurement precision only depends on resolution stability which is much higher than resolution itself when the border of a fuzzy area is utilized for detection. We adopt the theory of phase coincidence detection and border effect to devise the method for high-precision frequency measurement[1]. Measurement results which are apparently higher than resolution of the circuit by three orders of magnitude are acquired[2].

There are a variety of complicated frequency signals in frequency standard technology, telecommunication, exploration of natural phenomena, and electronic engineering etc., such as the frequency signal of 1,420,405,750Hz of a hydrogen atom’s energy-level transition, 19.59425MHz in a hydrogen atomic clock, 9,192,631,770Hz of a cesium atom’s energy-level transition, 16.384MHz and 19.44MHz in telecommunication, the subcarrier of 4.43361875MHz in a color television, signals used in GPS (Global Positioning System) and so on[3]. When only one standard frequency signal is utilized to measure complicated signals in a wide range of frequency, the phase coincidence detection fuzzy areas are always discrete because of the complex relationship between them. Traditional technique is always to make the reference signal equal to the frequency to be measured through frequency transformation[4], or to adopt related high-precision time interval measurement methods such as analog interpolation, vernier method, time-digit transformation. These methods are difficult to implement, and the ultimate measurement precision is still limited, the frequently-obtained precision being at the level of $10^{-11}$/s[5]. If the border information of the signals to be measured can be directly captured accurately in discrete fuzzy areas, higher measuring precision can be obtained directly. High resolution measurement of complicated frequency signals can exert a significant effect on precision improvement and function extension. An in-depth research on border effect helps to improve not only the measurement precision, but also the measurement scope and response-time.

II. THE SCHEME OF HIGH-PRECISION FREQUENCY MEASUREMENT WITH BORDER EFFECT

In the actual measurement, we discover that when the reference signal is the same as the signal to be measured, or is a multiple of it with a tiny difference, the order of phase coincidence pulse sequence is regular. The change of phase differences between two signals is monotonous and this results in a concentrated fuzzy area, as shown in Fig.1. Considering some random factors, actually the coincidence pulses triggering the counting gate are those in phase coincidence pulse clusters which are the first to be able to trigger the counting gate and at the border of fuzzy areas. States of phase difference between the two signals at the opening position and closing position of the counting gate are nearly the same. Although error still exists in theory, due to the high stability of the circuit, error is offset because of similar states of opening and closing of the phase difference between the two signals. So, high precision of frequency measurement is realized. In this case, measuring stability depends on the stability of the circuit at the border of fuzzy area, which is the stability of the coincidence detection resolution.
When the frequency relationship between two signals is complex, the order of phase coincidence pulse sequence is irregular, forming discrete fuzzy areas, as shown in Fig. 2. Utilizing these irregular coincidence pulses to trigger the counting gate, along with some external random factors, causes a large difference between phase differences of opening the gate and closing the gate, so the ultimate measurement error is great.

Fig. 1 Characteristics of same frequency periodical signals

III. THE SCHEME OF MEASURING COMPLICATED FREQUENCY SIGNALS

For the processing of discrete fuzzy areas, usually the frequency conversion is adopted. Using a direct digital frequency synthesizer (DDS) and so on, we adjust the reference signal frequency according to the measured signal frequency in order to keep the change of the phase difference between the reference signal and the signal to be measured monotonous. Thus a concentrated fuzzy area of phase coincidence detection can be formed, which is used for high resolution frequency measurement. The method is simple and convenient, but noise is produced at the same time. What’s more, the circuit is complicated, and the cost is high.

Based on border effect, if the border information can be accurately captured and used to trigger the counting gate, then stability of the fuzzy area can be utilized to replace resolution of the fuzzy area, and thus higher precision can be obtained. Because of the obvious characteristics of concentrated fuzzy areas, high precision can be obtained easily at this time. The case in which discrete fuzzy areas resulted, the information detected is complicated and its characteristics vary greatly with the numerical change of the quantity to be measured. No matter what type of fuzzy areas they are, concentrated or discrete, their internal states are relatively stable, and the status of the border changes easily. Phase information states are very stable within the fuzzy area, but the phase coincidence between signals can either be detected or not detected at the border of a fuzzy area, and slight changes at the border fuzzy area reflect subtle characteristics of comparison signals. Using discrete fuzzy area's border to trigger the counting gate can also improve measurement precision.

According to the research, the block diagram of complicated frequency measurement is shown in Fig. 3. The phase coincidence detection circuit with \( f_r \) and \( f_t \) as the input signals outputs discrete information. All information which indicates phase difference between signals is shown in pulses when the information is less than the resolution of the coincidence detection, leading to a discrete fuzzy area. Because D triggers have a delay from the clock-end to the output-end and a difference related to the frequency to be measured, most of coincidence detection signals are filtered except the coincidence information at the border. The delay module shown in the block diagram can eliminate the metastable phenomenon during collection.

Fig. 2 Characteristics of complicated periodical signals

Some similar phase coincidence states at the border acquired by several D triggers can work as signals which open and close the actual counting gate. It can be sure that using these states can obtain higher measurement precision. That is to say, utilizing the border of the discrete fuzzy area to trigger the gate eliminates the quantization error in the measurement, greatly improving precision.

The counting gate is started when the first coincidence pulse at the border after the reference gate opens is detected; the counting gate is closed when the first coincidence pulse at
the border after the reference gate closes is detected. Numbers of signals $f_0$ and $f_x$ are recorded separately in the measurement gate and the value of $f_x$ can be determined by Eq. (1).

$$f_x = f_0 N_x / N_0 \quad (1)$$

IV. EXPERIMENTS AND RESULTS ANALYSIS

In experiment 1, OSA8607 outputs 10MHz as the reference signal and the frequency to be measured is synthesized by HP8662. The measurement gate is triggered by these coincidence pulses between the two signals. Experiments 2 and 3 are two improved schemes based on the method for frequency-measuring proposed in this paper. The frequency to be measured is also produced by HP8662 and the reference signal is 10MHz from OSA8607. The measuring results of these three experiments are shown in Table 1.

<table>
<thead>
<tr>
<th>Frequency to be measured</th>
<th>Measured frequency value (Hz)</th>
<th>Frequency stability $\sigma$ (fs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.290913MHz synthesized by HP8662</td>
<td>4290913±0.001</td>
<td>2.8×10^{-9}</td>
</tr>
<tr>
<td>4.290913MHz synthesized by HP8662</td>
<td>4290913±0.00001</td>
<td>6.4×10^{-12}</td>
</tr>
<tr>
<td>82.5001MHz synthesized by HP8662</td>
<td>8250010±0.0003</td>
<td>7.5×10^{-12}</td>
</tr>
</tbody>
</table>

It can be seen that if $f_x$ is measured by merely adopting phase coincidence detection circuits, in other words, discrete fuzzy areas are not processed, the frequency stability of $10^{-9}$/s is acquired, certainly not a satisfactory result. The measurement precision is determined by the resolution of nanosecond level of the phase coincidence detection circuit. Adoption of border effect yields results which are three orders of magnitude higher than the resolution of the phase coincidence detection circuit.

These experiments indicate that the method presented in this paper can acquire the border information of the coincidence detection resolution fuzzy area when fuzzy areas are complicated and discrete. This method eliminates the quantization error and significantly improves the measurement precision. It is the definition of the fuzzy area that makes it possible for us to employ the border information to improve measurement precision.

V. CONCLUSION

Based on the existing foundation of measurement precision, using border effect properly can lessen phase coincidence pulses in a discrete fuzzy area and reduce the randomness of gate movements, thus improving the measurement precision. This new measuring method not only is applied in the measurement of complicated frequencies, but also can greatly simplify the process of frequency link in the design of quantum frequency standard. Many designs relevant to frequency control will change.

ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China under Grant No. 61201288, the Open Fund of National Key Laboratory of Aerospace Dynamics under Grant No. 2013ADL-DW0402, Xi'an Science and Technology Plan Projects under Grant No. CXY1351(6), Shaanxi Natural Science Foundation Research Plan Projects under Grant No. 2014JM2—6128.

REFERENCES

Single-Bit-Output All-Digital Frequency Synthesis Using Multi-Step Look-Ahead Bandpass $\Sigma$-$\Delta$ Modulator-Like Quantization Processing

Charis Basetas and Paul P. Sotiriadis
Department of Electrical and Computer Engineering
National Technical University of Athens, Greece
Email: chbasetas@gmail.com, pps@ieee.org

Abstract—A configurable single-bit-output $\Sigma$-$\Delta$ modulator variant is presented, which improves upon stability and dynamic range. The derivation of the proposed modulation scheme from the $\Sigma$-$\Delta$ error-feedback structure and its mathematical description are analyzed, while its application in fractional-$N$ frequency dividers and all-digital frequency synthesizers is investigated. It is shown that stable high-order noise shaping is possible even when using noise transfer functions that are unstable in a conventional $\Sigma$-$\Delta$ modulator loop. Furthermore, multipliers can be avoided if the noise transfer function coefficients are chosen appropriately, leading to a low complexity hardware implementation. In special cases the architecture is equivalent to another $\Sigma$-$\Delta$ modulator with a different, stable, NTF than the initial one. The effectiveness of the proposed modulation scheme is demonstrated by specific examples for each application.

I. INTRODUCTION

The defining characteristic of $\Sigma$-$\Delta$ modulation and one of the reasons for its widespread usage is the presence of a 1-bit or few-bit quantizer, leading to highly efficient Digital-to-Analog or Analog-to-Digital Converters (DACs or ADCs), while oversampling is exploited in order to shape the elevated quantization noise out of the signal band. The simplicity of all-digital $\Sigma$-$\Delta$ modulator implementations has led to their widespread usage in several applications such as fractional-$N$ frequency synthesizers [1] and direct all-digital frequency synthesis architectures [2].

Since the introduction of $\Sigma$-$\Delta$ modulation, there have been many attempts to utilize higher-order loop filters, while maintaining stable operation for a wide range of input signals. Due to the fact that $\Sigma$-$\Delta$ modulation is a non-linear process, its analysis is particularly difficult and is usually based on linear or quasi-linear models. Many empirical stability criteria and optimization techniques can be found in [3], while a stability analysis of high-order $\Sigma$-$\Delta$ modulators is presented in [4]. Stability problems relate to the input amplitude, the loop filter zeros and poles and the number of quantization levels. An increase in the number of quantization levels improves stability [3], but a multi-bit output necessitates the use of a DAC in the case of a direct all-digital frequency synthesizer or a more complicated frequency divider in a fractional-$N$ PLL.

This paper proposes a low complexity extension of $\Sigma$-$\Delta$ modulation in order to improve stability and dynamic range. In [5] the basis of our methodology is set. There, the $\Sigma$-$\Delta$ error-feedback modulator is viewed as an optimization algorithm which minimizes a norm of the difference between the input and the output sequences in a certain frequency band, defined by a comparison filter. Viterbi decoding provides the optimal solution to the aforementioned problem, but its complexity renders it inapplicable for practical purposes. Reduced convolutional decoding algorithms, such as list decoding, may be used instead [6], [7]. In any case the complexity of these algorithms is prohibiting for a real-time implementation and their application is approximate for infinite impulse response (IIR) filters. Our approach mitigates these problems and offers a balance between complexity and stability.

The remaining of the paper is organized as follows. Section II formulates mathematically the optimization problem that our algorithm solves as an extension of $\Sigma$-$\Delta$ modulation. In section III the application of the Multi-Step Look-Ahead modulator in fractional-$N$ PLL dividers and in all-digital frequency synthesizers is investigated, accompanied by simulation results. Finally, section IV discusses conclusions based on the previous sections.

II. MULTI-STEP LOOK-AHEAD OPTIMIZATION ALGORITHM

A. $\Sigma$-$\Delta$ Modulation Basics

A general 1-bit $\Sigma$-$\Delta$ modulator block diagram is shown in Fig. 1a. The discrete-time description of this system using the z-transform is given by

$$U(z) = L_0(z)X(z) + L_1(z)Y(z)$$

where $y_n = Q(u_n)$ is the quantization operation of the $\Sigma$-$\Delta$ modulator. A linear approximation of the system behavior is obtained if we suppose that $Y(z) = U(z) + N(z)$, where $N(z)$ is the z-transform of the quantization noise $n_n = y_n - u_n$. In this case the quantization noise is usually treated as zero mean white noise [3]. This substitution yields the following linear equation

$$Y(z) = \frac{L_0(z)}{1-L_1(z)}X(z) + \frac{1}{1-L_1(z)}N(z).$$

It is convenient to define $STF = L_0(z)/(1 - L_1(z))$ and

$$NTF = 1/(1 - L_1(z)),$$

where STF and NTF stand for Signal Transfer Function and Noise Transfer Function respectively.
indicates that the STF modifies the signal as it passes through the modulator and the NTF shapes the quantization noise. The error-feedback topology depicted in Fig. 1b lets the signal pass through undistorted and is described by

\[ Y(z) = X(z) + \frac{1}{1+G(z)} N(z). \] (3)

It is a special case of the generic \( \Sigma \Delta \) modulator of Fig. 1a with \( L_0(z) = 1 + G(z) \) and \( L_1(z) = -G(z) \) and consequently \( STF = 1 \) and \( NTF = 1/(1+G(z)) \).

The selection of the NTF depends on the frequency band of the input signal. A high order NTF offers higher in-band noise attenuation, but instability issues usually arise. The stability of the NTF can be checked by certain empirical criteria, many of them presented in [3]. This important stability requirement greatly reduces the choices of the NTF. The modification of the \( \Sigma \Delta \) decision rule presented in the next section shows that an unstable NTF in a conventional \( \Sigma \Delta \) loop, offers stable operation when utilized in the Multi-Step Look-Ahead optimization scheme.

B. Modification of the \( \Sigma \Delta \) Error-Feedback Topology

The error-feedback \( \Sigma \Delta \) topology of Fig. 1b may be described by the following decision rule [5]

\[ y_n = \min_{y \in \{\pm 1\}} \left| x_n + c_n - y \right|. \] (4)

where \( k \) is the number of forward time-instances that we take into account. The arguments \( v_0, v_1, \ldots, v_k \) denote all possible future values of the output \( y_n, y_{n+1}, \ldots, y_{n+k} \). This way the effect of \( k \) future inputs in addition to the current input value is taken into account when considering the next output value. For \( k = 0 \), eq. (5) reduces to eq. (4), i.e. the error-feedback \( \Sigma \Delta \) modulator. It can be seen that, looking further ahead typically results in increased dynamic range and stability of the modulator for a specific NTF selection.

Further simplification is possible if only the last term or the last two terms of the sum in (5) are considered. Extensive simulations have shown that this modification does not incur any performance degradation, while it enhances the dynamic range and stability. Then, (5) reduces to

\[ y_n = \arg \min_{v_0 \in \{\pm 1\}} \left( \min_{v_1, v_2, \ldots, v_k \in \{\pm 1\}} \left| x_{n+i} + c_{n+i} - v_{i} \right| \right) \] (6)

where \( v_1, v_2, \ldots, v_{k-1} \) are imbedded in \( c_{n+k} \).

C. Analysis of the Multi-Step Look-Ahead 1-bit Modulator with IIR Loop Filters

Let us consider the general case of \( G(z) \), i.e., in the form

\[ G(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_l z^{-l}}{a_0 + a_1 z^{-1} + a_2 z^{-2} + \ldots + a_m z^{-m}} \] (7)

For a realizable modulator a delay element between the modulator output and the filter input is required, that is \( b_0 = 0 \). We can also set \( a_0 = 1 \), as it only scales the filter output. Additionally, for notation convenience we drop the dependence of \( v_0, v_1, \ldots, v_k \) on time index \( n \). With these assumptions and (7), \( e_n, e_{n+1}, \ldots, e_{n+k} \) are given by

\[ e_{n+j} = \sum_{i=1}^{l} b_i x_{n+j-i} - \sum_{i=1}^{l} b_i v_{j-i} - \sum_{i=1}^{m} a_i e_{n-i}, \quad j < l \] (8a)

\[ e_{n+j} = \sum_{i=1}^{l} b_i x_{n+j-i} - \sum_{i=1}^{l} b_i v_{j-i} - \sum_{i=1}^{m} a_i e_{n-i}, \quad j \geq l \] (8b)

where \( 0 < j \leq k \). In this way we can recursively express \( e_{n+j} \) in terms of \( x_i, n - l - i \leq i \leq n + k - 1 \) and \( y_i, n - l - i \leq i \leq n - 1 \), as well as \( e_i, n - m - i \leq i \leq n \) (which do not depend on \( v_0, v_1, \ldots, v_k \)) and a linear function of \( v_0, v_1, \ldots, v_{k-1} \). It turns out that this linear function is

\[ f(v_0, v_1, \ldots, v_{j-1}) = - \sum_{i=1}^{j} c_i v_{j-i} \] (9)

where the coefficients \( c_j \) can be recursively calculated using

\[ c_i = b_i - \sum_{j=1}^{i} a_j c_{i-j} \] (10)
with $c_1 = b_1$. So, using (8) and (9), (5) is rewritten as

$$y_n = \arg \min_{v_0 \in \{\pm 1\}} \left( v_0, v_1, \ldots, v_k \in \{\pm 1\} \sum_{j=0}^{k} |A_n^j| \right)$$

$$\text{(11)}$$

where $A_n^k$ is a quantity independent of $v_0, v_1, \ldots, v_k$ and $c_i$ are given by (10) with $c_0 = 1$ and $c_1 = b_1$. Equation (6) can be rewritten as

$$y_n = \arg \min_{v_0 \in \{\pm 1\}} \left( v_0, v_1, \ldots, v_k \in \{\pm 1\} \sum_{i=0}^{k} c_i v_{i-1} \right)$$

$$\text{(12)}$$

From this analysis we conclude that the extended modulator selects its next output $y_n$ based on its current state and input, imbedded in $A_n^0, A_n^1, \ldots, A_n^k$, as well as all the possible future outputs until time instant $n + k$, weighted by the comparison filter, namely the terms $c_k v_0, c_{k-1} v_1, \ldots, c_1 v_{k-1}, c_0 v_k$. The decision for the next output value depends on possible future outputs, instead of only the current one. This means that the Multi-Step Look-Ahead 1-bit modulator can find meaningful output sequences, when the $\Sigma$-$\Delta$ modulator suffers from instability, resulting in a highly deteriorated SNR performance. The number of look-ahead steps $k$ can be increased until stability is achieved.

### III. Applications

#### A. Fractional-N Frequency Division

Fractional-N PLLs are very popular in communication and timing applications because they offer high frequency resolution with relative architectural simplicity. The fractional divider of the PLL is typically programmed by the output sequence of a $\Sigma$-$\Delta$ modulator which is responsible for both achieving the desirable fractional division ratio on average, and shape the spectral power of the quantization error away of the divider’s output frequency. Perhaps the most common $\Sigma$-$\Delta$ modulator for PLLs is the third order 1-1-1 MASH [1].

The popularity of the 1-1-1 MASH lies in its 3rd order noise shaping capabilities and the simplicity of its design. But the drawback of using this modulator for the generation of the divider control signal is that the divider should have a multi-bit control input, which complicates the divider design. This work demonstrates a design with the same noise shaping performance and simplicity as the 1-1-1 MASH, while having a single-bit output.

In the following example we start with the 3rd order NTF $(1 - z^{-1})^3$. In order to stabilize the loop we choose $k = 2$ and make use of (12). In this case the implementation of (12) is of low complexity and no multipliers are needed due to the convenient coefficients of the NTF. Thus, the resulting digital circuit is fast, small and power efficient. The power spectrum of $2 \cdot 10^5$ output samples of the Multi-Step Look-Ahead modulator for a DC input with amplitude 0.51 is shown in Fig. 2.

#### B. All-Digital Frequency Synthesis

A new approach to all-digital frequency synthesis is proposed, where a sinusoidal signal $\{x_n\}$ of desirable frequency $\omega_0$ is generated by a LUT with high-resolution output of $N$ Bits, exactly as in standard DDS. Then $\{x_n\}$ is fed into a Band-Pass (BP) Multi-Step Look-Ahead modulator performing the 1-bit quantization and shaping the quantization noise away from the carrier frequency $\omega_0$. In principle the quantization step can be performed using a conventional BP 1-bit $\Sigma$-$\Delta$ modulator. However, the stability of conventional BP 1-bit $\Sigma$-$\Delta$ modulators of order $2M, M = 2, 3, \ldots$ requires sacrifices in noise suppression, i.e. the choice of the NTF is very limited. Instead, the Multi-Step Look-Ahead modulator used here offers superior stability and therefore a significantly larger design space for the NTF. This setup is depicted in Fig. 3.

In order to achieve stability and high-order noise shaping we choose $NTF = (1 - e^{-j\omega_0 z^{-1}})^2(1 - e^{j\omega_0 z^{-1}})^2$ and $\omega_0 = 2\pi \cdot 0.365$. So the quantization noise is shaped by a double pair of conjugate zeros placed on the unit circle at frequency $\omega_0$. In Fig. 4 the spectrum of $2 \cdot 10^5$ samples of the output $\{y_n\}$ is shown when the amplitude of the sinewave $\{x_n\}$ is 0.7. NTF zeros may be split to broaden the useful bandwidth near-in. The resulting output spectrum for a NTF with three pairs of conjugate zeros at frequencies $\omega_1 = 2\pi \cdot 0.36, \omega_2 = 2\pi \cdot 0.365$ and $\omega_3 = 2\pi \cdot 0.37$ ($6^{th}$ order NTF) is shown in Fig. 5. Stability is guaranteed when $k = 10$. The signal bandwidth...
is significantly increased at the cost of an elevated in-band noise level.

IV. CONCLUSIONS

It has been shown that a low complexity real-time alternative to 1-bit $\Sigma$-$\Delta$ modulation offering increased dynamic range and stability is possible. Furthermore, the number of look-ahead steps $k$ is a design parameter capable of providing the perfect balance between performance and circuit complexity, depending on the application. Specific application examples have been given for the cases of a fractional-N frequency divider and a direct all-digital synthesizer architecture with simulation results demonstrating the noise shaping capabilities of the proposed Multi-Step Look-Ahead modulator.

ACKNOWLEDGMENT

Described work was partially supported by Broadcom Foundation USA.

REFERENCES


Fig. 4. Power spectrum of $2 \cdot 10^5$ output samples of the band-pass Multi-Step Look-Ahead modulator for $k = 2$ and a sinusoidal input with amplitude 0.7.

Fig. 5. Power spectrum of $2 \cdot 10^5$ output samples of the band-pass Multi-Step Look-Ahead modulator with split zeros for $k = 10$ and a sinusoidal input with amplitude 0.4.
Abstract—This work discusses hardware implementation considerations for a novel Multi-Step Look-Ahead modulation architecture which improves on the stability and dynamic range of conventional Σ-Δ modulators for all-digital frequency synthesis applications. The basic theoretical concepts of the architecture are analyzed and an appropriate general hardware implementation of the required mathematical operations is presented. It is shown that hardware complexity reduction is possible when noise-shaping filters with convenient coefficients are utilized. Moreover, FPGA and IC implementation examples for a specific noise-shaping filter are given, accompanied by power, area and delay estimations.

I. INTRODUCTION

The rapid advances in deep-submicron CMOS technology combined with the multi-faceted and stringent requirements of modern wireless communications protocols, have resulted in the intensive research of highly integrated all-digital frequency synthesizers. Digital circuits in sub-100nm CMOS technologies exhibit superior characteristics, namely noise and PVT robustness, faster switching times, smaller die footprint etc. compared to the degrading analog devices characteristics such as increased leakage current and lesser degree of scalability.

All-digital frequency synthesis is an ongoing research subject and many of its challenges have not yet been completely resolved [1]. E.g., dithering techniques have been used as a remedy for spurious tones [2], while single-bit output has been proposed for DAC-less architectures [3]–[5]. Dithering however leads to an elevated noise floor, which makes all-digital frequency synthesis unsuitable for many noise-sensitive applications. Noise shaping using Σ-Δ modulation type techniques may be used to lower the noise floor within a specific frequency band. Yet, noise shaping filters of order higher than two and high input amplitude may result in stability problems of the loop [6]. Increased dynamic range and stability may be achieved, simultaneously, if future inputs and states of the modulator are taken into account when calculating the next output sample.

This work deals with hardware implementation aspects of a new Multi-Step Look-Ahead noise-shaping architecture with single-bit output. Its reconfigurable noise-shaping properties and single-bit output make it an ideal building block for DAC-less low-noise all-digital frequency synthesizers. The theoretical background is analyzed in [7], while a brief overview is presented herein.

The remainder of the paper is organized as follows. Section II presents an overview of the Multi-Step Look-Ahead modulation scheme, setting the path for section III, where a general hardware implementation architecture is introduced. FPGA and IC implementation of a specific MSLA modulator are also discussed. Finally, section IV summarizes the results and investigates further research requirements.

II. MULTI-STEP LOOK-AHEAD MODULATION OVERVIEW

The conventional Σ-Δ error-feedback structure (Fig. 1) constitutes the basis of the Multi-Step Look-Ahead (MSLA) modulator. The output is determined by the comparison of the filtered difference between the high resolution input and the single-bit output against the current input. This decision rule is formulated as [8]

\[ y_n = \arg \min_{y \in \{ \pm 1 \}} |x_n + e_n - y| \tag{1} \]

where \( y_n \) is the output, \( x_n \) is the input and \( e_n = \sum_{i=0}^{n-1} (x_i - y_i) g_{n-i} \) is the filtered error sequence \( \{x_i - y_i\} \). Sequence \( \{g_i\} \) is the impulse response of the noise-shaping filter \( G \) and \( g_0 = 0 \). The frequency response of the comparison filter \( G \) allows the signals in the desired frequency range to pass, while it greatly attenuates out-of-band signals. Thus, the single-bit output tracks the input in the pass-band, pushing the quantization noise out of the signal band. The relation between \( G \) and the noise transfer function (NTF) of the modulator is \( G(z) = (1 - NTF)/NTF \). A higher order filter allows for better noise attenuation, but given the stability requirements for filters of order \( N > 2 \), the available design space for the NTF is greatly reduced.

Fig. 1. The error-feedback Σ-Δ modulator topology
This limitation may be overcome when \( k \) future inputs of the modulator are taken into account. Then, the modified decision rule is

\[
y_n = \arg \min_{v_0 \in \{\pm 1\}} \left( \min_{v_1, v_2, \ldots, v_k \in \{\pm 1\}} \left| y_{n+k} + e_{n+k} - v_k \right| \right)
\]  

(2)

where the arguments \( v_0, v_1, \ldots, v_k \) denote all possible future values of the output \( y_n, y_{n+1}, \ldots, y_{n+k} \) and \( k \) is the number of look-ahead time-instances. If the comparison filter is

\[
G(z) = \frac{\sum_{i=1}^{l} b_i z^{-i}}{1 + \sum_{m=1}^{m} a_i z^{-i}}
\]

(3)

then

\[
e_{n+j} = \sum_{i=1}^{l} b_i x_{n+j-i} - \sum_{i=1}^{j} b_i v_{j-i} - \sum_{i=j+1}^{l} b_i y_{n+j-i} - \sum_{i=1}^{m} a_i e_{n-i}, \ j < l 
\]

(4a)

\[
e_{n+j} = \sum_{i=1}^{l} b_i x_{n+j-i} - \sum_{i=1}^{l} b_i v_{j-i} - \sum_{i=1}^{m} a_i e_{n-i}, \ j \geq l 
\]

(4b)

where \( 0 < j \leq k \).

Through extensive simulations it has been concluded that superior stability and dynamic range is achieved even if only the last term or the two last terms of the sum in (2) are considered, leading to the simplified expression

\[
y_n = \arg \min_{v_0 \in \{\pm 1\}} \left( \min_{v_1, v_2, \ldots, v_k \in \{\pm 1\}} \left| y_{n+k} + e_{n+k} - v_k \right| \right)
\]  

(5)

where \( v_1, v_2, \ldots, v_{k-1} \) are embedded in \( e_{n+k} \).

In [7] it is shown that an equivalent expression for (2) is

\[
y_n = \arg \min_{v_0 \in \{\pm 1\}} \left( \min_{v_1, v_2, \ldots, v_k \in \{\pm 1\}} \sum_{j=0}^{k} |A_j - \sum_{i=0}^{j} c_i v_{j-i}| \right)
\]  

(6)

where \( A_j \) is a quantity independent of \( v_0, v_1, \ldots, v_k \) and \( c_i \) are given by

\[
c_i = b_i - \sum_{j=1}^{i-1} a_j c_{i-j}
\]  

(7)

with \( c_0 = 1 \) and \( c_1 = b_1 \). Using the same reasoning (5) may be rewritten as

\[
y_n = \arg \min_{v_0 \in \{\pm 1\}} \left( \min_{v_1, v_2, \ldots, v_k \in \{\pm 1\}} \sum_{j=0}^{k} |A_j - \sum_{i=0}^{j} c_i v_{j-i}| \right)
\]  

(8)

In order to evaluate (6) or (8) a proper expression for \( A_j \) is needed. It turns out that

\[
A_j = \sum_{i=0}^{j} c_i x_{n+j-i} - \sum_{i=j+1}^{l} c_i (x_{n+j-i} - y_{j-i}) + \sum_{i=0}^{m-1} d_i e_{n-i}
\]  

(9)

where \( d_i \) are coefficients depending on the filter coefficients \( a_i, l \) is the order of the numerator of the comparison filter \( G \) and \( m \) is the order of the denominator. Finally, \( e_n \) is calculated by the difference equation implied by the comparison filter

\[
e_n = \sum_{i=1}^{l} b_i \left( x_{n-i} - y_{n-i} \right) - \sum_{i=1}^{m} a_i e_{n-i}
\]  

(10)

III. HARDWARE IMPLEMENTATION OF THE MULTI-STEP LOOK-AHEAD MODULATOR

A. Reconfigurable Hardware Implementation Architecture

From the analysis of the previous section it should be clear that a hardware implementation of the MSLA modulator should evaluate (6) or (8). The output of the modulator \( y_n \) is a function of \( A_n^0, A_n^1, \ldots, A_n^k \). It can be shown that the optimization problem (6) can be transformed into the problem of simultaneously satisfying a number of linear inequalities. Based on that, a decomposition of the hardware required for deriving the solution, whilst maintaining low complexity, is achievable. The solution may be as simple as the sign of \( A_n^k \) for special cases of (8), effectively transforming the MSLA into an equivalent conventional \( \Sigma \Delta \) modulator with a different, stable, NTF than the one that we started with. This is the case for the NTF \( (1 - z^{-1})^3 \) and \( k = 2 \). Another example of the mapping between \( A_n^0, A_n^1, \ldots, A_n^k \) and \( y_n \) is depicted in Fig. 2 for NTF \( (1 - z^{-1})^2 \) and \( k = 1 \). Similar mappings in higher dimensions are obtained for more sum terms in (6), while different filters change the mapping region boundaries.

The calculation of \( A_n^j \) is based on (9). The coefficients \( c_j \) and \( d_i \) are precalculated based on the comparison filter coefficients, while \( k + l - 1 \) previous values of the high resolution input, \( l - 1 \) previous values of the single-bit output and \( m - 1 \) previous values of \( e_n \) should be stored in memory elements. The proposed architecture for the computation of \( A_n^j \) is shown in Fig. 3. It consists of an IIR filter for the calculation of \( e_n \), three multiply and accumulate (MAC) units for the computation of the three sums in (9) and three adders. The memory units and the clock signals have been omitted in the figure for clarity. This is the building block of the MSLA modulator. For the evaluation of (8), one \( A_n^k \) computation
The complexity of each A\textsuperscript{0} \text{comp.} unit is depicted in Fig. 4. The more than one or two units in parallel for the computation of (6), but thorough simulations have shown that no reduction in hardware complexity and increase in maximum possible even if it is implemented in a low-end FPGA. Further reduction in hardware complexity and increase in maximum clock frequency is possible if the filter coefficients are simple enough, e.g. powers of two or a sum of powers of two, so that hardware multiplication units are replaced by an adder in combination with a shift operation.

B. FPGA Implementation Results

As a test-case, the MSLA modulator was implemented in a low-end Xilinx Spartan 3E Starter Board. A multiplier-less low-pass modulator design was chosen with NTF = (1 − z\textsuperscript{-1})\textsuperscript{3}. The FPGA implements (8) with k = 2. With this value of k, stability is maintained for a maximum sinusoidal input amplitude A = 0.61. The design was optimized for speed and operates with a maximum clock frequency of 142 MHz. The high-level synthesis tool reports 16 16-bit adders and 242 flip-flops for the entire design. This accounts for just 1% utilization of the total FPGA resources.

This test-case implementation demonstrates the MSLA modulator’s suitability for FPGA device implementations. Its hardware requirements are minimal even for a low-end FPGA, meaning that implementations for higher values of k and for higher order comparison filters are well within the capabilities of low-end FPGA devices.

C. Standard 65nm CMOS Physical Design

The MSLA modulator with the characteristics of the previous subsection has also been implemented in a standard 65nm CMOS technology with 1.0V supply voltage. The EDA tools utilized in the design flow are the industry standard Cadence® Encounter® RTL Compiler (Synthesis) and the Cadence® Encounter® Digital Implementation (EDI) System (Place & Route). The structural netlist of the modulator consists of 540 standard cells, taking up a total area of 4,695 μm\textsuperscript{2} with an estimated power consumption of 1.724 mW at clock frequency 400 MHz. The layout of the IC physical design is shown in Fig. 5. Aiming at the highest possible operating clock frequency using standard cells, a maximum clock frequency of 850 MHz was achieved, while the synthesized circuit occupied an area of 4,697 μm\textsuperscript{2} with an estimated power consumption of 3.394 mW.

The hardware implementation in a common IC technology exhibits the low-power and small area characteristics of the MSLA modulator. This means that an MSLA module may easily be integrated as part of a larger design with minimal area and power overhead. Of course, as it is an all-digital design, it will benefit from any reduction in the scale of integration.

IV. CONCLUSIONS AND FUTURE INVESTIGATIONS

In the previous sections a hardware implementation architecture for the MSLA was proposed, along with various proposals for the reduction of the required hardware, such as the careful selection of the comparison filter in order to avoid multipliers. The MSLA modulator exhibits high order noise-shaping properties with low hardware complexity and high operating frequency, rendering it ideal for all-digital frequency synthesis applications.
Many design aspects require further investigation and we are currently working on several techniques to make the design faster and more energy efficient. Significant improvements are expected from the bit-width optimization of each basic component of the MSLA hardware architecture, as well as from more advanced pipelining and parallelization of the required arithmetic operations. Finally, more comparison filters with a wide range of different orders and frequency characteristics need to be tested and implemented.

ACKNOWLEDGMENT

Described work was partially supported by Broadcom Foundation USA.

REFERENCES

Compact Clocks for Industrial Applications: the EMRP Project IND 55 MClocks

S. Micalizio, F. Levi, A. Godone, C. E. Calosso, B. François
Istituto Nazionale di Ricerca Metrologica, INRIM
Torino, Italy
s.micalizio@inrim.it

S. Guérandel, D. Holleville, E. De Clercq, L. De Sarlo, P. Yun, J. M. Danet, M. Langlois
LNE-SYRTE, Observatoire de Paris, Paris, France
stephanie.guerandel@obspm.fr
david.holleville@obspm.fr

R. Boudot, M A. Hafiz, Femto-St, Université Franche Comté, Besançon, France, rodolphe.boudot@femto-st.fr

E. Sahin
National Metrological institute of Turkey, Tubitak, Kocaeli, Turkey, ersoy.sahin@tubitak.gov.tr

C. Afföldterbach, S. Kang, F. Gruet, M. Gharavipour, G. Mileti
Laboratoire Temps Fréquence (LTF)
Université de Neuchâtel (UniNe)
Neuchâtel, Switzerland
gaetano.mileti@unine.ch

B. Desruelle
Muquans, Talence, France
bruno.desruelle@muquans.com

Abstract—Vapor cell atomic clocks are an interesting technology because they combine compactness, low power consumption and excellent relative frequency stability. Recently, due to better performing laser sources and innovative techniques to prepare and detect the atoms, several cell-based prototypes exhibiting unprecedented frequency stability have been developed. These techniques allow a reduction in the transfer of laser noise to the atoms, improvement of the signal-to-noise ratio and subsequently the clock’s frequency stability. The project IND55 Mclocks funded by the European Metrological Research Programme (EMRP) proposes to develop high performances vapor cell clocks for industrial applications. Three technologies are investigated: 1) the pulsed optical pumping (POP) scheme; 2) the cold atoms approach, and 3) the Coherent Population Trapping (CPT). The results related to the first period of activity are presented.

Keywords—vapor cell clock; frequency stability; laser pumping; cold atoms

I. INTRODUCTION

Atomic frequency standards provide the ultimate source of accuracy and stability for all modern communication, navigation and timekeeping systems, and nowadays commercially available devices are deployed in many strategic industrial fields. As these industrial activities are essentially exploiting microwave frequencies, it turns out that vapor-cell clocks are particularly suited to fulfill industrial requirements. In fact, besides working in the microwave regime, vapor-cell clocks are compact, portable, reliable, with low power consumption and exhibit a good short-term frequency stability.

However, the development of compact and high performing frequency standards is of interest in several scientific and technological applications, as well as in industry, such as, e.g., improved local oscillators for future primary frequency standards or improved clocks for the Galileo space segment. In industry, forthcoming Galileo services, such as autonomous landing of airplanes, mooring of ships or vehicle autopilots will need cm positioning accuracy. This implies the need for new clocks with very good frequency stability and more than 10 times better accuracy than current commercial products.

Recently, due to better performing laser sources and to innovative techniques to prepare and detect the atoms, several cell-based prototypes exhibiting unprecedented frequency stability have been developed.

Examples of these techniques are pulsed laser pumping [1], coherent population trapping (CPT) [2], isotropic laser cooling [3] and light shift suppression methods [4]. For certain laboratory prototypes of such clocks frequency stabilities in the order of $10^{-13}$ at 1 s and in the range of $10^{-14}$ or better for the medium-long term were measured, a result even better than passive H-masers. In addition, the clocks based on cold atoms are on the way to providing an accuracy rivaling commercial Cs atomic clocks, drastically reducing the need for periodic frequency calibration.

In this context, the project IND55-MClocks [5] funded by EMRP addresses the following objectives: 1) to develop a vapor-cell clock based on the pulsed optical pumping (POP) principle with a fractional frequency stability of $10^{-15}$ at 1 s and in the $10^{-14}$ range at $10^5$ s; the clock will be targeted on industrial applications in terms of size, power consumption and reliability; 2) to develop a vapor-cell clock based on cold Rb atoms with performance comparable to that of POP in the short term but with better long term performance including an accuracy within an order of magnitude of that of primary standards; the project will identify the compromises required in order to obtain the expected performance while still targeting industrial applications; 3) to investigate alternatives principles such as CPT, to study the possibility of realizing a clock optimized in terms of compactness.

In this paper, we present the status of the project; specifically, we will describe the realization of a compact laser
system and the design of the magnetron cavity for the POP clock. We will report on the development of Rb clock based on cold atoms (Rubiclock) and, finally, preliminary results of a Cs vapor cell CPT clock based on push-pull optical pumping will be presented.

II. POP Rb CLOCK

Within the frame of the JRP IND55, we will develop a Rb vapour-cell clock based on the pulsed optical pumping (POP) technique and approaching the requirements of industry and advanced technology. The laboratory prototype of the POP Rb clock already implemented at INRIM represents the state of the art for vapor cell frequency standards, exhibiting a frequency stability of $10^{-13}$ at 1 s and reaching the $10^{-15}$ range for measurement times up to one day [6].

In the project we expect to get the same results but adopting innovative solutions allowing to reduce the overall size of the clock. In this regard, we are going to realize a small-volume microwave cavity and to develop a compact laser source module.

A. The POP clock physics package

Several parts of the clock physics package for the POP clock are newly developed by UniNe-LTF, such as the Rb vapor cells, the microwave cavity, and a frequency-stabilized laser system with pulsed optical output. The Rb cells are custom-made for this project and have been designed specifically to reach a stem temperature coefficient 100 times lower than for the cells used in previous clock studies [7]. The Rb cell will be placed in a highly compact magnetron-type microwave cavity (see Fig. 1), similar to the one described in [8]. This cavity has a very small overall volume of 45 cm$^3$, more than 2-times smaller than for a fundamental-mode cylindrical cavity, but still maintaining a highly homogenous microwave magnetic field for achieving high-contrast Ramsey signals. Preliminary POP clock experiments using a previous version of the magnetron-type microwave cavity have been conducted at UniNe-LTF. In these experiments clear Ramsey fringes on the clock signal were obtained, showing a fringe contrast of 35% at a central fringe width of 160 Hz, and resulting in a short-term clock stability of $2.1 \times 10^{-13}$ τ$^{-1/2}$ [9]. The frequency-stabilized laser system including an acousto-optical modulator (AOM) for switching the optical output has also been developed and realized (optics assembly volume of 1 liter only) and was used successfully to operate a POP clock [10]. These new building blocks will contribute to achieve a very compact clock physics package, in view of a clock realization for industrial applications.

III. A COLD ATOMS Rb FREQUENCY STANDARD: RUBICLOCK

One of the goal of the JRP IND55 is to develop a prototype vapour-cell clock based on cold atoms which can deliver a short term stability comparable to that of active hydrogen maser (relative stability below 5×10$^{-13}$ at 1 s), while at the same time having an accuracy at the level of a few 10$^{-15}$, which is inherently superior to the GPS+H-maser systems currently in use. A key feature of this activity is to demonstrate this development in a form which is geared towards industrial applications in terms of cost, size and projected reliability. In this respect, the presence within the JRP-Consortium of an industrial JRP-Partner will be instrumental in guiding the development. The prototype will close the existing gap between currently available commercial products and atomic fountains both in terms of stability and accuracy. This requires sacrificing in part the simplicity and miniaturisation potential of the POP clock to allow the use of an ultracold atomic sample.

A. Rubiclock physics package

The physics package is composed of several sub-systems. Main of them are the atomic resonator with its 3 magnetic shields, the laser system and the microwave synthesis. Resonator and synthesis have been made at SYRTE, and the laser system has been developed by Muquans. Inside the atomic resonator, we find the microwave cavity and the vacuum chamber, where the atoms are cooled, interrogated and detected. Vacuum chamber is mainly in titanium, allowing use of commercial components as dispenser and pumps, and the upper part is surrounded by the microwave cavity. To perform the isotropic light cooling, the upper part of the camber is a fused silica bulb allowing optical access in all directions. The inner walls of the copper cavity act as an integrating sphere and 6 free-diverging fibers feed the cooling zone with cooling and re-pumping beams. To avoid loss of optical power along the vertical direction we add a vertical retro-reflected beam, which is also used as a detection beam. Microwave cavity is tunable thanks to a copper ring at its top, allowing a tuning range of the resonance frequency of about 20 MHz.
Fig. 2: CAD of Rubiclock atomic resonator with its 3 magnetic shields. At the center, the fused silica bulb and the microwave cavity.

B. Laser system

The laser system has been developed by Muquans Company in collaboration with the LP2N. This system uses some reliable fibered telecom components at 1.5 µm, followed by a frequency doubling stage to reach the 780 nm rubidium D2 line. It generates more than 200 mW of optical power at 780nm, providing all the beams and frequencies needed for atom cooling, preparation and detection phases. Thanks to the use of fibered components, the whole setup including command electronics is very robust and compact (rack 19” x 6U). It has been tested and validated during 2 parabolic flights campaigns.

C. Synthesis chain

Microwave synthesis chain has been designed by SYRTE’s Electronic Service. The local oscillator is a 10 MHz AR quartz. Specifications are compliant with a Dick effect less than $2 \times 10^{-13}$. The Synthesis uses a PLL-DRO to generate a 7 GHz signal. A DDS signal is mixed to reach the 6.8 GHz clock frequency. A second DDS is used to control the frequency of a second output needed to generate the repumping beam of the laser system. This synthesis has been characterized and it does not degrade the noise of local oscillator. This setup has been integrated in a 3U rack.

D. Results

The Rubiclock prototype has been assembled at the beginning of 2014. Since it participated in 6 parabolic flights, validating all the sub-systems and the technological choices. Thanks to microgravity, the clock can operate with longer interrogation times than in the lab, with a great win of the fringe contrast (see Fig. 3). In the same way, fringes with Ramsey interrogation time up to 400 ms has been demonstrated, but signal-to-noise ratio (SNR) was limited by the fluctuations of acceleration in the 0g plane.

Fig. 3: Ramsey fringe with a 80 ms interrogation time on Earth and in microgravity. Fringe contrast is greatly improved in microgravity.

Short term stability measurements have not be done yet, but we perform lot of work in the lab to reduce the noise sources and to increase the SNR. The quantum projection noise has been reached and shot-to shot Ramsey fringe SNR up to 1500 has been observed (see Fig. 4). These very good results let us expect a short term stability in the low $10^{-13}$ range, and the measurement will be made in the next weeks.

Fig. 4: (blue) measured shot-to-shot detected atom number fluctuations vs. atom number. (red) estimated ultimate quantum noise.

IV. CPT CLOCK USING THE PUSH-PULL TECHNIQUE

In the frame of the Mclocks project, we recently started in FEMTO-ST the development of a high-performance compact Cs CPT atomic clock based on the pioneering push-pull optical pumping technique, proposed by Jau et al. [11]. We reported the spectroscopy of high-contrast CPT resonances in Cs vapor cells [12] and demonstrated the possibility to detect high-contrast Raman-Ramsey fringes [13].
In this paper we report on first frequency-stability results of the CPT-PPOP based Cs vapor cell atomic clock. A promising short-term frequency stability of $3 \times 10^{-13}$ at 1 s is measured.

A. Experimental set-up

Figure 5 presents the Cs CPT clock experimental set-up.

The laser source is a distributed-feedback (DFB) diode laser tuned at 894 nm on the Cs D$_1$ line [14]. A MZ EOM (Photline NIR-MX800-LN-10) driven at 4.596 GHz with a low noise microwave frequency synthesizer by [15] is used to generate phase-coherent optical sidebands frequency-separated by 9.192 GHz. At the output of the EOM, the optical carrier rejection is actively stabilized thanks to an original microwave synchronous detector presented in [12]. The laser beam at the output of the EOM is sent into an annex reference Cs cell. Saturated absorption spectroscopy is used to stabilize the laser frequency. At the output of the EOM, an acousto-optic modulator (AOM) is used for two main functions. Its first role is to compensate for the optical frequency shift due to the presence of buffer gas in the CPT cell. This optical shift was measured to be -122 MHz. The second function is stabilize the laser power. For this purpose, a fraction of the laser power is extracted, converted into a voltage with a photodetector, compared to a stable voltage reference. This allows the generation of an error signal used to correct permanently the RF power that drives the AOM. A Michelson delay-line and polarization orthogonalizer system is used to produce the PPOP scheme [11]. At the output of the Michelson system, the diameter beam is expanded to 2 cm thanks to a pair of convergent optical lenses to cover the whole diameter of the vapor cell atomic resonator. The atomic resonator is a 2-cm diameter and 5-cm long cylindrical Cs vapor cell filled with a $N_2$-Ar buffer gas mixture of total pressure 15.3 Torr and partial pressure ratio $r = P_{Ar}/P_{N2} = 0.4$ to operate in the Dicke regime. The cell temperature $T_{cell}$ is stabilized at 29 °C where the CPT signal height is maximized. The presence of buffer gas causes a collisional frequency shift of the Cs clock transition of about 9.3 kHz. A static magnetic field of 89.16 mG parallel to the laser beam propagation direction is applied to split the hyperfine ground-state Zeeman transitions. The ensemble is surrounded by a double-layer mu-metal magnetic shield. The CPT resonance is monitored by detecting the laser power transmitted through the cell. The resulting signal is analyzed by a computer that drives the microwave synthesizer output frequency. The clock output signal is compared to the signal of a reference hydrogen maser for frequency stability measurements.

B. Experimental results

The short-term relative frequency stability $\sigma_y(\tau)$ of a passive atomic clock is well-approximated by [16]:

$$\sigma_y(\tau) \sim \frac{\Delta \nu}{\nu_c \text{SNR}} \tau^{-1/2}$$

where $\Delta \nu$ is the clock resonance full-width at half maximum (FWHM), $\nu_c$ is the clock transition frequency, SNR is the signal-to-noise ratio in a 1 Hz bandwidth of the detected signal and $\tau$ is the integration time of the measurement.

Fig. 6 reports the typical clock signal for an incident laser power $P_L$ on the cell of 700 μW.

The clock resonance is well-fitted by a lorentzian function. The CPT linewidth $\Delta \nu$ is 564 Hz. The clock signal $S$, defined on Fig. 6 as $S = H - y_0$, is 0.114 V. The resonance contrast $C$, defined as the ratio between the CPT signal $S$ and the dc background $y_0$, is 22%. The discriminator slope $D = S/\Delta \nu$ is measured to be 0.2 mV/Hz.
Figure 7 plots the Allan deviation of the clock frequency. The latter is measured to be $3 \times 10^{-13}$ at 1 s, going down to about $3.2 \times 10^{-14}$ at 100 s. A large bump, from a few seconds to hundreds seconds averaging time, presently prevents the Allan deviation to decrease with a perfect white frequency noise slope. This is currently under investigation but could be explained by thermal effects (the laboratory room is not temperature stabilized), a non-optimized gain of the EOM bias voltage servo loop and feedback from the EOM fiber input face that slightly affects the laser frequency. An additional optical isolation stage should be added in a near future. Nevertheless, these short-term frequency stability performances, comparable to those of best vapor cell atomic clocks, are very encouraging for the development of a high-performance Cs vapor cell CPT atomic clock.

Table shows main contributions to the clock short term frequency stability at $\tau = 1$ s. The noise budget is in excellent agreement with the measured clock fractional frequency stability. The clock frequency stability is currently mainly limited by the laser AM noise and the laser FM-AM noise process. The following contribution is the Dick effect. The latter will be reduced later thanks to a novel ultra-low noise frequency synthesis chain [17] that should reject the Dick effect contribution to a level lower than $7 \times 10^{-14}$ at 1 s.

<table>
<thead>
<tr>
<th>Noise source</th>
<th>$\sigma_f(1s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shot noise</td>
<td>$8.5 \times 10^{-15}$</td>
</tr>
<tr>
<td>Detector noise</td>
<td>$3 \times 10^{-14}$</td>
</tr>
<tr>
<td>LO phase noise</td>
<td>$1.1 \times 10^{-14}$</td>
</tr>
<tr>
<td>Laser AM noise</td>
<td>$2.8 \times 10^{-15}$</td>
</tr>
<tr>
<td>Laser FM-AM noise</td>
<td>$1.7 \times 10^{-13}$</td>
</tr>
<tr>
<td>Total $\sigma_f(1s)$</td>
<td>$3.02 \times 10^{-15}$</td>
</tr>
</tbody>
</table>

V. CONCLUSIONS

We presented in this work the preliminary results related to the project IND 55 Mclocks funded by EURAMET. Three different technologies addressing different applications are investigated.

For the POP clock, we implemented a compact laser-source module with an overall volume of 1 liter. A design of a magnetron microwave cavity has been also realized.

For the Rubiclock clock, we develop a prototype of the compact cold atom clock in collaboration with Muquans Company, integrating industrial compatible technologies. An important work to reduce noise sources has been done and the metrological characterization of the prototype will begin in the next weeks. First results let expect a short term stability in the low $10^{-13} \tau^{1/2}$ range. In the same time, we start a theoretical work on systematic effects, as the cavity pulling effect, to prepare the evaluation work on long term stability and accuracy.

We implemented a CPT-based Cs vapor cell atomic clock using the push-pull optical pumping technique. The optics part of the clock, compatible with further integration, mainly combines a single diode laser, a Mach-Zehnder electro-optic modulator, an acousto-optic modulator for laser power stabilization and a Michelson delay-line system. An encouraging short-term frequency stability of $3 \times 10^{-13}$ at 1 s was demonstrated, in good agreement with the signal-to-noise ratio limit. Laser intensity effects were found to be the main limitation to the clock short-term frequency stability performances.

REFERENCES


Advances of Chip-Scale Atomic Clock in Peking University

Zhao Jianye1*, Zhang Yaolin1, Lu Hao Yuan1,2, Hou Dong1, Zhang Shuangyou1, Wang Zhong1

1Department of Electronics, School of Electronics Engineering and Computer Sciences, Peking University, Beijing, China
2Department of Electronic Engineering, Fudan University, Shanghai, China
zhaojianye@pku.edu.cn

Abstract—The authors are developing chip-scale atomic clocks (CSACs) based on the 85Rb coherent population trap (CPT) transition. As an intermediate milestone, we have developed a miniature atomic clock prototype. In this paper, we report on the design combining the miniature physics part, low-power digital control circuits and low-power microwave system, and the process that enable an atomic clock to be made with an overall size of 20 cm³ volume, power consumption about 400 mWatts, and an Allen Deviation at 100 s of 2.9E-11.

Keywords—CSAC; CPT; Atomic clocks; MEMS

I. INTRODUCTION

Atomic clocks are applied to keeping time very accurately by locking the oscillators to stable and precise atomic hyperfine transitions, and play an essential role in the modern communications and navigations. Atomic clock technology continues to evolve in different directions. One of them goes after the excellent stability performance [1]; the other is pursuing the low power consumption and small volume [2]. The relatively large size and power consumption of atomic clocks with excellent stability have prevented the wide applications when the scenarios require the clocks portable and battery-powered. Therefore, the demand for CSACs is steadily increasing with the emergence of broadband and secure communications and precise location and navigation systems [3].

Since the idea to combine the CPT spectroscopy with the micro-electromechanical systems (MEMS) for the fabrication of CSACs is proposed in [4], CSACs based on the CPT [5] have been an active research area for a few years. Many great research institutions are devoting to the realization of small-size, highly stable and low-power atomic clocks. The MEMS vapor cells and Vertical Cavity Surface Emitting Lasers (VCSELs) are the key part of the miniature atomic clocks, can make the atomic clock (CSAC) physics packages with size 1cm³ and frequency stabilities below 2×10⁻¹⁰/τ¹/₂[3].

In this paper, we report our advances in the miniature atomic clock prototype based on the 85Rb coherent population trap transition.

II. ATOMIC CLOCK PHYSICS

Our miniature atomic clock is based on the 85Rb CPT transition, and uses the 5S½ mF=0 hyperfine transition between the F=3 and F=2 ground states. The CPT excitation scheme is illustrated in Fig.1. We use the MEMS and anodic bonding technologies to fabricate the chip-sized 85Rb cells [6, 7]. The 85Rb atoms are generated based on the chemical reaction between rubidium chloride (RbCl) and barium azide (BaN₆). The nitrogen (N₂) gas from the decomposition of BaN₆, remaining in the cells acts as the buffer gas. The MEMS Rb cells are fabricated with inner dimensions of 3 mm length and 3 mm radius. A 10% CPT resonance contrast (CPT signal/Optical absorption) is observed with the temperature at ~ 50. The linewidth of the CPT signal is less than 5 kHz.

Several key components must be assembled to make physics parts of our miniature atomic clock. The architecture is shown in Fig.2. A vertical cavity surface-emitting laser (VCSEL) modulated by 3GHz radio frequency (RF) is used to generated linearly polarized dichromatic 794.9 nm light. A quarter waveplate then converts the linearly polarized light to circularly polarized light. The Rb cell and VCSEL are heated by a high-temperature ceramic material, and the temperature is stabilized within 0.01K. The ensemble is magnetically shielded with μ-metal shield (not show in Fig.2) in order to prevent spurious environmental magnetic fields from affecting the resonant frequency of the clock. A photodetector (PD) collects the transmitted light, converts optical signal to electrical signal used to lock the 794.9 nm light to the absorption transition and the 3GHz RF frequency to the CPT resonant frequency of the 85Rb, respectively. The whole volume of the physics parts is about 2cm³.
III. ELECTRONICS DESIGN

A block diagram of our miniature atomic clock is shown in Fig.3. The main controller is a MSP430 microcontroller. This microcontroller generates two audio frequency signals to modulate the RF signal and direct current source for the two frequency servo loop. The optical frequency of the VCSEL driven by the current source is stabilized to the $^{85}$Rb D1 absorption line. The 3-GHz signal is synthesized by multiplying around 300 times from the 10-MHz temperature compensated crystal oscillator (TCXO). In the modulation of RF signal, a new microwave frequency modulation scheme is achieved by directly injecting triangular current signal into the charge-pump of voltage-control oscillator (VCO), which is one part of RF synthesis [8]. The signal from the PD is amplified by an operational amplifier and contains two frequency modulation signals for the locking of the laser and RF. This signal is rapidly digitized by an analog-to-digital converter (ADC) and the error signals are digitally filtered and demodulated by a new servo algorithm based on discrete Fourier transformation (DFT) algorithm. The laser and RF error signals are applied to two digital-to-analog converters (DACs), which control the laser DC bias and TCXO tuning respectively.

IV. EXPERIMENTAL RESULTS

We have demonstrated a complete 20 cm$^3$ miniature atomic clock with power consumption about 400mW. Fig.4 shows the frequency drift of the stabilized 10MHz from the TCXO, which demonstrates a short-term stability of $2.9 \times 10^{-10}/\tau^{1/2}$ for $\tau$=1-100 seconds, as shown in Fig.5. The current power consumption is limited by the physics part. The physics package, thermal management need to be more carefully designed. In the future, instead of heating the laser and cell together, we will control the temperatures of the laser and cell respectively, to further reduce the power consumption. In our prototype, the new servo algorithm based on DFT algorithm can optimize the long-term stability of the clock, and reduce the complexity of electric circuits.

ACKNOWLEDGMENT

The authors wish to acknowledge the following team members for valuable technical, theoretical, and moral support: Wu Chenwei, Zhang Yu, Yang Zhibin, Tian lu, Chen Zhen, Gao Shan from Department of Electronics, School of Electronics Engineering and Computer Sciences, Peking University.

REFERENCES


An Atomic Frequency Micrometer Based on the Coherent Population Beating Phenomenon

WANG Zhong, ZHAO Jianye, ZHAO Xiaona, LIU Li, ZHUANG Yuxin and LI Dawei
School of Electronics Engineering & Computer Science
Peking University
Beijing, P. R. China
zw@pku.edu.cn

Abstract—We have demonstrated an atomic frequency micrometer based on the coherent population beating phenomenon, which enables us to obtain the beat frequency between the measured signal and the atomic transition frequency. The beat frequency and its fluctuations are detected and accurately measured through digital signal processing, which is capable of up to mHz or higher frequency resolutions (for GHz signal). The frequency discrimination via our method is comparable to that of the Ramsey fringes method, and the working range is no longer limited by the width of line shape. This enables us to achieve an atomic clock by actively compensating the frequency shift, which eliminates the need for a phase locking loop, and broadens the working range with increasing reliability. This novel scheme could be extended to the optical atomic clocks, optical frequency comb, atomic spectroscopy and other related researches.

Keywords—micrometer; CPB; CPT

Atomic clocks have become increasingly important to scientific research as well as our daily lives, and research studies relating to frequency control and measurement have attracted attention from around the world. Typically, a precision frequency measurement will require a frequency standard as a reference, which usually is an atomic clock, as it has very good frequency precision and stability. The standard frequency of an atomic clock is locked onto an atomic transition frequency, which is very stable, and acts as the ultimate reference frequency. For better accuracy during the measuring process, the measured high frequency signals are beat with reference signal of the atomic clock, which converted the high frequency measurement to low frequency measurement. It is conceivable notion that if we directly beat the measured frequency with the atomic transition frequency, we could omit the intermediate process with the atomic clock, and achieve the frequency precision measurement with the atomic transition itself being the frequency standard, which would be much more accurate and reliable. In this paper we will demonstrate how the coherent population beating (CPB) phenomenon and the rapid development of digital signal (DSP) processing technology helped us to achieve this goal.

Based on the CPB effect, we can directly obtain the frequency difference between the measured signal and the atomic transition—the beat frequency. The main portion of the measured high frequency signal is equal to the transition frequency. We then obtain the small frequency difference between the measured signal and the atomic transition frequency via the CPB effect, and measure this “magnified” beat frequency in a manner similar to the fine adjustments of a micrometer. Hence, we have named this frequency measurement system the “atomic frequency micrometer”.

With the application of the CPB phenomenon in this atomic frequency micrometer, we can directly use the atomic transition as the reference signal to measure an radio frequency (RF) signal frequency. This precludes the need for an additional precise frequency standard. Since the DCPB effect has converted the high frequency (GHz or Higher) signal to a much lower frequency (kHz) measurement, it enables us to accurately read and measure the beat frequency via a low stability (better than $10^{-5}$) crystal oscillator, achieving accuracy within a few millihertz or higher. And then we can obtain the frequency of the measured gigahertz order RF source with millihertz accuracy, which equals the atomic transition frequency plus/minus the CPB frequency. We now have the potential to improve the performance of atomic clocks with the achievement of a CPB atomic clock, which uses this accurate measurement results to actively compensate for fluctuations of the RF source. This novel scheme may be widely used in atomic clocks, magnetometers [3], chip scale atomic clocks (CSAC) [4,5], and optical clocks [6,7], etc.

The CPB phenomenon has been observed and studied by various groups (under different titles) and has been theoretically and experimentally approved [1,2,8] CPB occurs in a typical three level $\Lambda$ system (Fig.1). When the frequency difference of the two pump laser fields $\omega_1=\omega_2=\omega$, equals the ground state hyperfine splitting frequency $\Delta_n$, all the electrons will be trapped within the two ground states. This is the well-known coherent population trapping (CPT) phenomenon. However, when the $\omega_i$ experiences a small detuning from the hyperfine splitting ($\omega_i\neq\Delta_n$), a portion of the electrons will be pumped up to the excited state. Due to the frequency difference of the $\omega_i$ and $\Delta_n$, the excited state population $\rho_s$ will coherently oscillate with their beat frequency, hence the term: detuned coherent population beating.
From the density matrix, the analytical solution of $\rho_{33}$ can be obtained \(\text{[11]}\) as

$$
\rho_{m} = \text{Re} \left[ A \exp \left( - \left( \gamma_{1} + \frac{\Omega^{2} - \Delta_{1}^{2} \lambda_{1}}{\gamma} \right) t + i\Delta_{1} t \right) + \frac{2\kappa_{3} A}{\gamma} \right] \tag{1}
$$

The formula conditions and details of the derivation have been described in Ref.1 and 11. Where \(A\) is constant, \(\gamma_{1}\) is the coherence decay rate of the two ground states. \(\Delta = |\omega_{21} - \Delta_{21}|\) is the detuning frequency, \(\Omega\) is the Rabi frequency (we assumed \(\Omega = \Omega_{1} = \Omega_{2}\)), in the second term

$$
\kappa_{3} = \frac{\omega^{2}}{4(\omega^{2} - (\omega_{1} - \Delta_{21}^{2}) \lambda_{2}^{2})} \tag{2}
$$

$$
\lambda_{3} = -\gamma_{2} + i\Delta \tag{3}
$$

They are constant and not related to time. According the equation we will see a damping oscillation of \(\rho_{33}\). The oscillation spectrum will exhibit a Lorentzian lineshape with linewidth:

$$
\gamma_{2} + \frac{\omega_{1}^{2} - \Delta_{1}^{2}}{\gamma} \tag{4}
$$

Usually, we have \(\Delta \ll \Omega\), to the linewidth can be simplified to \(\gamma_{2} + \frac{\omega_{1}^{2}}{\gamma}\), which is the same as CPT resonance line width\(^{9}\). In this manner CPB oscillation can be seen as the time domain transform of the CPT resonance, but its center frequency is converted to a very low frequency Equation (1) shows the important features of CPB oscillation, 1) its frequency is equal to \(\Delta = |\omega_{21} - \Delta_{21}|\), 2) it is a damping oscillation. 3) Its line width will be less than or equal to CPT resonance.
Fig. 3. (a) The theoretical result of $\rho_{33}$, the horizontal axis is relative time and the vertical axis is the relative population of the excite state.

Fig. 3. (b) The experimental result of the laser intensity, the horizontal axis is time $t$, the vertical axis is the relative laser intensity $I$.

Since $\Delta$ is the difference frequency between the RF and the atom’s (which could be $^{85}\text{Rb}$ $^{87}\text{Rb}$ or $^{133}\text{Cs}$ etc.) ground state hyperfine splitting $\Delta_{21}$, we can easily obtain the RF frequency $\omega_{21} = \Delta_{21} \pm \Delta$. The CPB effect enables us to take the hyperfine splitting frequency as reference and convert the measurement to a low frequency region, which is crucial to data sampling and processing. This low frequency measurement allows us to use the measured signal itself as the reference (in our experimental system the measured OCXO signal is used as the reference frequency of FPGA) to accurately measure the beating frequency and fluctuations related to the hyperfine splitting. In this scheme, the frequency measuring precision for the RF signal is related to the beating frequency measurement accuracy and the system’s physical stability.

Fig. 4. Microwave frequency $f_0$ is increasing by 1Hz/step (marked by dots) and the corresponding oscillating frequency $f_c$ is represented by a series of triangles. $\tau = 20s$ is the time sampling interval. In our experiment $\Delta_{21} > \omega_{21}$, so a negative sign is added to the $f_c$. The frequency resolution is 0.01Hz and the processing time is about 100 seconds for each data point.

We have tested the measurement accuracy of this atomic frequency measurement system. In the experiment, an analog signal generator is employed to generate the RF signal, of which, the frequency shift is less than $1 \times 10^{-10}/200s$. Since $^{85}\text{Rb}$ is used, at the first data point we set the $f_c = \omega_{21}/2\pi = 3035731310.00\text{Hz}$, and measured the CPB frequency $f_c = \Delta_{21}/2\pi = 1281.34\text{Hz}$. We then increased the $f_0$ by 1 Hz per step, and for each step, take the reading of the corresponding oscillation signal within 20 seconds. A total of 10 steps were in completed 200 seconds to ensure the RF signal itself total shift is smaller than 0.5Hz. Fig.4 illustrates our experimental result. For every step of the $f_0$ increase, the measured $f_c$ has a corresponding decrease, and the probable deviation of $f_c$ is less than 0.5 Hz. In correspondence to the 10 Hz increase of $f_0$, the $f_c$ decreased 9.90Hz, and the entire deviation is about 0.1Hz. In this experiment, the software processing time is about 100 seconds for each data point, and the frequency resolution is 0.01Hz, which is one order better than the limit of the equipment and the experimental system instability. If it is needed, we can also improve the resolution to mHz or higher through more accurate calculations. Although the improvement of calculation accuracy will require more processing time and using more powerful software and hardware, if we are able to limit the DCPB oscillation to within the order of kHz, the development of digital processing technology will enable us to reach the physical system limit, which is in the order of mHz for high stability atomic clock systems.

Fig. 5. Sketch map of the CPB frequency equal precision measurement: the square wave in the middle is the signal converted from the DCPB oscillation, using its rising edge to open and close the measuring gate. We then compare the $T_G$ with the standard signal $f_s$ to measure the DCPB oscillation frequency.

Based on this CPB scheme, there are other digital processing methods with which we can achieve this measurement. For example, the detected CPB signal can be shaped into a square wave with the same frequency, then measured with the typical equal precision measurement digital processing, as shown in Fig.5. The measurement error rate is

$$\delta \leq \frac{1}{n} = \frac{1}{\tau f_s} \quad (5)$$
Here, \( n \) is the standard signal pulse number in \( T_G \), and \( f_i \) is the standard signal frequency. As a result, the measurement accuracy is proportional to \( T_G \) and \( f_i \). However, for CPB frequency measurements, \( T_G \) will be limited by the damping time \( T_d \), which is about 10ms in our system. Since \( f_i \) is 3MHz, it gives a \( 1/(3\times10^4) \) measurement error rate; and for a 3kHz CPB signal, the frequency resolution is 0.1 Hz. To improve accuracy, we can periodically excite the oscillation and add a number \( N \) of \( T_G \)'s together, as long as the standard signal is continua, then the measurement error rate would become

\[
\delta \leq \frac{1}{n} = \frac{1}{NT_G} \tag{6}
\]

and the measurement accuracy is now proportional to \( NT_G \). In the real system, considering signal to noise ratio and picking up the stable part of the CPB signal, \( T_G \) will take about 1/3 to 1/2 of \( T_d \). Therefore, in one second \( NT_G \) will be about 0.2 second; the corresponding frequency resolution is \( 5\times10^{-3} \)Hz. By either increasing time to 10 seconds, or increasing \( f_i \) to \( 10^3 \)Hz (with the standard signal frequency instability less than \( 1\times10^{-5} \)), the measuring resolution could exceed mHz accuracy. Today, the OCXO frequency instability is usually better than \( 10^{-11} \), and the DSP technology supports \( f_i \) up to \( 10^9 \) Hz. Theoretically speaking, by increasing the frequency of standard signal to gigahertz, the measurement resolution can be up to \( 10^{-5} \) to \( 10^{-6} \) Hz in 10 seconds of processing time. The measured RF frequency can be obtained by adding/subtracting the DCPB frequency to/from the atomic ground state hyperfine splitting transition frequency. Hence, the measurement accuracy is also related to the stability of the atomic transition. Comparing with traditional methods, if the atomic transition frequency stability is the same, this CPB method of frequency precision measurement allows us to omit the intermediate process and directly take the atomic transition as the reference, which increases measurement stability, reliability, and accuracy. In addition, this scheme can also be employed in the atomic transition measurement for the research of atomic spectroscopy, when the RF signal is from a frequency standard.

This atomic frequency micrometer can also be directly used to achieve an atomic clock. As illustrated in Fig. 2, to generate the standard frequency of the atomic clock, the dashed (DDS) part is added into the system. The \( \omega_2 \), and its real time shift related to the hyperfine splitting will be obtained in the FPGA. The direct digital synthesizer (DDS) is used to generate a 10MHz signal, which takes the same OCXO frequency as its reference. Therefore, the output signal will shift with a certain rate following the shift of \( \omega_2 \). The FPGA will generate a delta frequency word (with resolution of up to \( 1/24^8 \) of the output frequency) for the DDS to compensate the output signal frequency shift. The DDS 10MHz output signal will be stabilized with the reference of the atomic transition.

Fig. 6 shows the Allan Deviation of the rubidium CPB atomic clock with FFT method, in which the digital processing period is 2 seconds and the frequency resolution is 1 Hz. In our preliminary experiment, a frequency instability of \( 3\times10^{-11} \) at 1000s of integration was observed (Fig. 6), which is comparable to the published data \(^4,5\) of CPT atomic clocks. We also performed the experiment on a DCPB atomic clock with equal precision measurement methods, and observed a frequency instability of \( 1\times10^{-11} \) at 1000s of integration\(^13\). In real CPB atomic clocks, the frequency resolution is limited by the processing period, and we can solve this problem through optimizing the software as well as using more powerful hardware.

Traditionally in high stability atomic clocks, the well-known Ramsey fringes method\(^14,15\) is used to improve frequency resolution of the atomic transition spectrum. With this method, two interrogating fields with time interval of T are applied to excite the Ramsey fringes, and amplitude of the fringes is proportional to the excited state probability\(^14\):

\[
P = 1/2[1 + \cos(\omega_0 - \omega)T]C \tag{7}
\]

Where \( C = \sin \omega T \) is a coefficient, \( \omega_0 \) is the atomic transition frequency, \( \omega \) is the frequency of the two interrogating fields. According to formula (7) the spectrum resolution can be improved by increase T. In 2005 Zanon et al proposed the CPT Ramsey method\(^16\), in which the second laser pulse is used to interrogate the excited state probability oscillation’s phase delay in T, and to obtain the fringes. Theoretically the spectrum resolution or the precision of the transition frequency measurement is proportional to the action time of the electromagnetic fields and the atoms, which is equivalent to T in Ramsey fringes scheme. In comparison, the frequency resolution of DCPB scheme described in formula (6) is proportional to \( NT_G \). An advantage of DCPB scheme is that we can extend the total effective action time by adding together N times the interrogation time \( T_G \). In the Ramsey method, the T is limited by the relaxation time of the two hyperfine levels. In DCPB method, \( NT_G \) are only limited by the time that can be spent on measuring one data point, which tends to more accurate measurement and higher resolution. This corresponds to directly obtaining the \( (\omega_0 - \omega) \) of formula...
(7), which what the Ramsey fringes method is going to detect. The excited state probability oscillates with the beat frequency between the interaction field and the hyperfine splitting. This phenomenon also occurs in two level systems. In fact, it is the same basic principle of the Ramsey fringe formation. However, under the continual wave interrogation in a two level system, this oscillation will be mixed with the Rabi oscillation, and cannot be directly observed.

Actually the CPB method is based on the same physical effect as the Ramsey method; the difference being that Ramsey fringe is in the frequency domain, while DCPB oscillation is in the time domain. We can view the CPB method as another approach to improving the spectrum resolution of the atomic transition. For most atomic clocks (aside from beam and fountain clocks), this method is more concise, more convenient, and more direct.

With the support of today’s high speed DSP, the DCPB method’s ultra-high frequency discrimination is comparable to the Ramsey method (both up to mHz). Usually the width of Ramsey fringe can be narrowed down to 1Hz. For ground hyperfine states transition and CPT resonance, the linewidth will be between 10 to 10^3 kHz, and in general consideration, the frequency discrimination ability will be 1/1000 of the linewidth. Simultaneously, in the traditional approach, with the increasing ability for frequency discrimination, the narrowed line shape or fringe will decrease the frequency discrimination region. In contrast to the traditional phase locking loop scheme, in the DCPB scheme, the frequency discrimination region is no longer limited by the width of the atomic transition line shape or the Ramsey fringe. As a result, the CPB oscillation in our experiments has a much broader working range and can be observed from 10^2Hz to 10^4Hz. This broadened working range will benefit the system’s stability and reliability, and completely eliminate the problem of a frequency losing lock. In addition, the digitized processing enables us to actively reduce the noise disturbances via methods such as: choosing the stable part of the signal, eliminating the noise point, and using better processing programs, etc. All the above suggest that CPB scheme has the potential to improve the performance of atomic clocks.

With a conveniently digital control circuit and no need for a microwave cavity, the CPB atomic clock can be easily integrated and miniaturized, which holds promising implications for further CSAC development. As the measurement accuracy of the atomic frequency micrometer is unrelated to the measured frequency nor the atomic transition frequency, we can, in theory, extend this technique to the optical frequency region and utilize its benefits in optical clocks and optical frequency combs.

In conclusion, the CPB effect allows us to directly take the atomic transition frequency as reference to achieve a quantitative and accurate frequency measurement. The development of digital signal processing technology enables us to accurately analysis and use the CPB phenomenon. Based on this CPB scheme and DSP technique, we have demonstrated a digitized atomic frequency micrometer with excellent measurement resolution and accuracy. Its frequency discrimination is comparable to that of Ramsey fringes, and it has a much broader working range which increases system reliability. It provides a novel solution to achieve CSAC, microwave and optical atomic clocks which eliminate the traditional phase locking loop and have great potential to improve their performance. This atomic frequency micrometer also provides us a novel method for identifying atomic properties, which suggests possible applications for atomic spectroscopy, relevant atomic effects exploration, frequency control, magnetometer, and other related research areas.

**ACKNOWLEDGMENT**

This work was partially supported by the Major National Basic Research Program of China (2013CB922401) and the National Natural Science Foundation of China (No.11074012). The author would like to thank Dr. Wang Yiqiu, Chen Jingbiao, Zhu miao, and Deng Ke for their informative discussion, to thank Wang Xin for her constructive suggestions during the writing of this text.

**REFERENCES**


Digital servo system based on FPGA for optically pumped magnetometer

Zhou Sheng, Liu Chang, Wang Yanhui

Institute of Quantum Electronics, School of Electronics Engineering and Computer Science, Peking University, Beijing, China

Email: wangyanhui@pku.edu.cn

Abstract—This paper describes the design and implementation of a digital servo system based on FPGA for one optically pumped magnetometer. The digital servo system based on FPGA incorporates an analog to digital converter, a digital to analog converter and a direct digital synthesizer. In a simple test setup, the Zeeman transition frequency is detected and locked by this digital servo system. The measurement of the magnetic field is experimentally realized. This digital servo system is designed to detect and lock the Zeeman transition frequency in mind but should be adaptable to a variety of other similar systems. With a minor change, this digital servo system is already applied in our atomic clock.

Keywords—digital servo system; magnetometer

I. INTRODUCTION

The use of the servo system is widespread in frequency locking in atomic physics[1-5]. Alkali atomic magnetometer [6-8] is one of the best methods to measure the weak magnetic field by detecting the Zeeman transition frequency. The Zeeman transition spectrum could be detected and the Zeeman transition frequency could be locked by a servo system. One cesium optically pumped magnetometer is successfully realized as we report in [9]. For making the optically pumped magnetometer more practical and portable, a digital servo system is designed to replace the former servo system based on personal computer and NI acquisition card.

The digital servo system is designed and implemented with field programmable gate array (FPGA), analog to digital converter (ADC), digital to analog converter (DAC) and direct digital synthesizer (DDS). The Zeeman transition frequency is detected and locked by this digital servo system. Particularly, the intensity of light is detected by a photo detector and then converted to digital signals by ADC. The Zeeman transition spectrum is detected by sweeping the frequency of the radio frequency field which is generated by DDS controlled by FPGA and the Zeeman transition frequency is also locked by frequency switching method [10-11] based on FPGA.

The measurement of the magnetic field is realized in an experiment. With a minor change, this digital servo system is also applied in our atomic clock.

II. MAGNETIC MEASUREMENT SYSTEM

A. Description of the experimental apparatus

The schematic diagram of the experimental magnetic measurement system is shown in Fig. 1, which is the M_z configuration of the cesium optically pumped magnetometer.

The frequency of pumping light which is from a DFB laser is locked in $D_1$ line ($^6S_{1/2}, \text{F}=4 - ^2P_{1/2}, \text{F}'=3$) of cesium by

![Fig. 1: Schematic diagram of the probe for detecting Zeeman transition signal and the digital servo system for locking the Zeeman transition frequency.](image-url)
saturated absorption spectroscopy. The pump light is circularly polarized by the polarizer and quarter-wave plate. The power of the light is 0.7 mW and the collimated beam is 10 mm in diameter. The intensity of the light that has passed through the cylindrical cell, which is 25 mm in diameter and 25 mm long, is detected by a photodiode and the photocurrent is converted into voltage by analog circuits and then converted to digital signals by ADC. The static magnetic field to be measured is generated by an inner coil now. To reduce the influence of outer magnetic field, the cell is surrounded by magnetic shields. Beside the cell, there is a RF-coil to generate an RF field controlled by the digital servo system. And the whole system is put into a light-proof and temperature control box in order to avoid optical influence and maintain a temperature of 50 °C.

B. Description of the digital servo system

The servo system for locking the Zeeman transition frequency is a core component of the device. Fig. 2 shows the block diagram of the digital servo system. The system consists of the following hardware components: FPGA, ADC, DAC and DDS. These components constitute the digital servo system and require a layer of interface to convert signals to register values back and forth.

![Fig. 2: Schematic diagram of the digital servo system](image)

The Altera Cyclone IV (EP4CE6E22C8) is chosen to be the controller. The FPGA works with a low-noise 24 bits ADC (ADS1256, TI company) with 30K SPS to convert analog signals to digital signals. A 32 bits frequency tuning word DDS (AD9850, AD company) to generate radio frequency signal and a 16 bits DAC (DAC8534, TI company).

The control unit, including communications between FPGA and other chips and digital signal process, is programmable and compatible with hardware description language in a software (Quartus II). The communications between ADC, DAC and FPGA is handled over an SPI-compatible serial interface. Through the SPI-compatible serial interface, FPGA could configure the operation mode of ADC and DAC. And the parallel port model is applied between FPGA and DDS. By changing the 32 bits frequency tuning word, FPGA could control the output frequency of the DDS from 20kHz to 200kHz with the resolution of 0.0291Hz. The timing control and digital signal process, including calculating the error signal for locking the Zeeman transition frequency, is realized by FPGA.

C. Description of the locking process

The function of locking the Zeeman transition frequency $f_c$ is realized by using switching method. Before locking the Zeeman transition frequency $f_c$, the Zeeman transition spectrum needs to be detected. And the normal method is that FPGA controls the frequency of the output of DDS sweep from 20kHz to 100kHz and monitors the change of the power of the light. If there is a proper change in the power of the light, there is the Zeeman transition spectrum and the frequency corresponds to the lowest power of the light may close to the Zeeman transition frequency. In our program, the frequency corresponds to the lowest power of the light is called the center frequency $f_c$, which is close to the Zeeman transition frequency $f_c$. And the Zeeman transition frequency $f_c$ is detected by locking the $f_c$ to the center of the Zeeman transition spectrum. And the digital-locking process goes through the following steps in Fig. 3.

![Fig. 3: The block diagram of the locking process](image)

In locking process, the error signal is specially significant to quantize the difference between the target frequency and the frequency to be locked to the target. In Zeeman transition spectrum, the switching method is applied to generate the error signal. When $f_c > f_s$, the power of the light corresponds to the frequency $f_c - \Delta f$ is smaller than the power of the light corresponds to the frequency $f_c + \Delta f$. Thus, the voltage $V(-)$, which means the collected voltage from the photo detector when the RF is $f_c - \Delta f$, is smaller than the voltage $V(+)$, which represents the collected voltage from the photo detector when the RF is $f_c + \Delta f$. And the error signal $V_{err} = V(-) - V(+) < 0$. Similarly, when $f_c < f_s$, the error signal $V_{err} = V(-) - V(+) > 0$ and when $f_c = f_s$, $V_{err} = V(-) - V(+) = 0$. The switching method could be realized by digital circuits. In the program, $\Delta f$ is set to the line width[10], the error signal could well quantize the difference between $f_c$ and $f_s$ and then could be used to adjust $f_c$ to be closer to $f_c$. 472
The block diagram shows the states machines in digital circuits. The servo system repeats the six steps and eventually \( V_{err} \) appears to be very close to zero. At the same time, \( f_s \) is locked to \( f_L \).

### III. Measurement Results and Discussion

Fig. 4 shows the photo of the digital servo system. The digital servo system consists of AD circuit based on ADS1256, DDS circuit based on AD9850 and FPGA and DA circuit based on Cyclone IV and DAC8534. And in a simple test setup, the Zeeman transition frequency is detected and locked by this digital servo system. And a 250 seconds long recording of a constant magnetic field with the FPGA based digital-locking cesium magnetometer operated with the frequency switching method has been made. The upgrade rate of the field value is set at 1 Sample/s. And the peak-to-peak fluctuation of the measured magnetic field is about 300nT. Thus, the sensitivity of the FPAG based digital servo system digital-locking Cs magnetometer is about 300 nT/Hz^{1/2}.

### IV. Conclusion

This paper presents a digital servo system based on FPGA and this digital system is applied in an optically pumped magnetometer. With the frequency switching method, the digital servo system locked the Zeeman transition frequency.

### REFERENCES


Measuring Buffer-Gas Pressure in Sealed Glass Cells

T. U. Driskell, M. Huang, and J. C. Camparo
Physical Sciences Laboratories
The Aerospace Corporation
2310 E. El Segundo Blvd., El Segundo, CA 90245
james.c.camparo@aero.org

Abstract — In alkali rf-discharge lamps used for optical pumping in atomic clocks and magnetometers, a buffer-gas (Kr or Xe) allows electrons to extract energy from an rf-field, and these energized electrons eventually produce alkali resonant light. Contrary to naïve intuition, rf-discharge lamps can lose their noble-gas buffer over time. Recently, we began a long-term experimental program to better understand the mechanism of noble-gas loss in rf-discharge lamps, and needed a non-destructive means of measuring buffer-gas pressure in sealed glass cells. For this purpose, we employ the Kazantsev, Smirnova, and Khutorshchikov (KSK) technique, which is based on inferring buffer-gas pressure from the collision shift of an alkali ground-state hyperfine transition frequency \( \nu_{\text{hfs}} \). Here, we discuss the basic the KSK technique and two modifications that we have implemented for its improvement: use of a diode laser for optical pumping, and extrapolation of \( \nu_{\text{hfs}} \) to zero magnetic field. Testing our system’s long-term performance with a very low pressure reference cell (i.e., 3.3 torr Xe), we find a reproducibility of 0.2% and an absolute accuracy of 5%. Further, our systematic drift is less than one mtorr/month.

I. INTRODUCTION

The rf-discharge lamp of the Rb vapor-cell clock [1] and Hg ion-trap clock [2] are arguably the most important elements in these clocks’ physics packages: their efficiency for optical pumping determines the amplitude of the atomic clock signal, and variations in their brightness and spectrum can affect the clock frequency through the light-shift effect [3,4]. Nevertheless, though the technology of rf-discharge lamps traces back to the early decades of the 20th century [5], they remain one of the least well understood components of an atomic clock’s physics package.

Focusing on the alkali rf-discharge lamp illustrated in Fig. 1, optimum atomic signal generation is often achieved with the lamp operating near the “ring-mode” to “red-mode” transition [6], a region of operation where radiation trapping first begins to play a dominant role in the plasma physics [7]. Not only does the visual appearance of the discharge from the lamp change in this transition region, but the lamp can exhibit instability in the form of low-frequency pulsing [8]. Moreover, depending on the interaction of the discharge’s complex permeability and the electronic circuit powering the discharge [9], ion-acoustic waves can develop in the discharge giving rise to \( \sim 10 \text{ kHz} \) light intensity oscillations [10,11].

Conceptually, the alkali rf-discharge lamp is a relatively straight-forward device: a glass bulb containing several hundred micrograms of liquid Rb [12,13] and a few torr of noble gas (Xe or Kr) is placed within the inductor coils of an rf-oscillator. When power is supplied to the circuit a resonant field with a frequency of \( \sim 10^2 \text{ MHz} \) builds up, and it is this field that creates the discharge. Electrons oscillate in the field, and through elastic collisions with the noble gas atoms extract energy from the field. Rubidium atoms are ionized by these energetic electrons (\( \langle \text{KE} \rangle \sim 0.3 \text{ to } 0.4 \text{ eV} \)), and Rb\(^+\)/electron recombination at the bulb’s glass walls [7,14] produces the Rb resonance light that is used for optical pumping.

![Figure 1: Picture of a typical Rb rf-discharge lamp. A glass bulb constructed from an alkali-resistant glass (e.g., Corning 1720) houses a liquid pool of Rb and either a Kr or Xe noble gas at a pressure of several torr. The glass bulb sits in a metal base, which allow easy insertion into the inductor coils of an rf-oscillator circuit. The inductor coils surround the glass bulb, allowing rf-energy to excite electrons, which then ionize the Rb atoms. Recombination at the lamp’s glass walls produces the Rb resonance light that is used for optical pumping.](image)

This work was funded by U.S. Air Force Space and Missile Systems Center under Contract No. FA8802-14-0001.
electrons in the discharge ionize Kr or Xe (either directly or by stepwise excitation). The noble-gas ions are then accelerated by the plasma sheath, colliding with the lamp’s glass walls and irreversibly imbedding themselves there.

To better understand the mechanism of noble-gas loss in alkali rf-discharge lamps, we have begun a long term project to measure Xe loss in small alkali rf-discharge lamps. Briefly, similar to our procedure for measuring Rb loss [19], we monitor the Xe pressure in a sealed glass lamp as a function of lamp operating time. Clearly, we require a means of measuring buffer gas pressure in the lamp non-destructively, and we are employing the technique of Kazantsev, Smirnova, and Khutorshchikov (KSK) for that purpose [20]. In the following section, we will briefly outline the KSK technique. We will then discuss two modifications that we have made, and a one year test examining our Xe pressure measurement system’s precision and long-term stability.

II. THE KSK TECHNIQUE

With the KSK technique, the buffer gas pressure in a Rb rf-discharge lamp is estimated by measuring the Rb atoms’ ground-state hyperfine transition frequency, which for the unperturbed atom is 6834.682608 MHz [21]. This is accomplished by employing the lamp as the resonance cell in a separate Rb atomic clock, and comparing the clock’s output frequency to the unperturbed value. Since the “clock” frequency is typically dominated by the buffer-gas collision shift [22], the frequency of the sealed cell’s Rb atoms compared to the unperturbed value provides a good estimate of the buffer gas density (or pressure). In the case of Xe, the pressure shift, $\frac{\text{d} \nu}{\text{d} P_{\text{Xe}}}$, is $-1184 \text{ Hz/torr}$ [23]. Consequently, for (let’s say) a 4-torr Xe lamp the clock frequency of the lamp will be shifted from “truth,” $\Delta \nu_{\text{lamp}}$, by $-4.74 \text{ kHz}$. Further, if the Xe pressure changes in time, which is the object of our studies, then so too must $\Delta \nu_{\text{lamp}}$, providing a potentially sensitive means for monitoring a lamp’s noble-gas vapor pressure as a function of operating lifetime:

$$\delta P_{\text{Xe}}(t) = \left( \frac{\text{d} \nu}{\text{d} P_{\text{Xe}}} \right) \Delta \nu_{\text{lamp}}(t) = \left( \frac{\text{d} \nu}{\text{d} P_{\text{Xe}}} \right) (\Delta \nu_{\text{lamp}}(0) + \Delta \nu_{\text{lamp}}(t)).$$

III. THE MEASUREMENT SYSTEM AND PROCEDURE

Figure 2 shows a block diagram of our buffer-gas pressure measurement system, which is specialized for low pressure (~ 4 torr) Xe lamps. A VCSEL diode laser is locked to the Rb D₁ transition at 795 nm, and provides the optical pumping light that creates a hyperfine polarization in the “resonance-cell” lamp. The lamp is situated in an oven centered in a set of three mutually perpendicular Helmholtz coils, and is maintained at a temperature of $\sim 63 \text{ °C}$. In the original KSK technique a second lamp was used for optical pumping. However, not only does the VCSEL diode laser produce more efficient optical pumping, but tuning the laser provides some control over the light-shift effect, and hence an assessment of one possible source of systematic error. Two of the Helmholtz coils cancel out the Earth’s field in the laboratory, while the magnetic field produced by the third, $B_z$, provides a quantization axis.

Figure 2: Buffer-gas measurement system as described in the text.
The laser passes through a half-wave plate, allowing us to adjust the polarization axis of the laser, and keep this constant over multiple months. It then passes through a neutral density filter, which allows us to keep the light intensity at near optimum levels (i.e., large enough for good signal-to-noise, but not so large as to produce excessive light shifts). The laser beam is expanded so that from measurement-to-measurement, as the lamp is placed and re-placed in the measurement system, reflection and scattering differences of the laser light do not lead to systematic variability in pressure assessments. After passage through the lamp, the expanded laser beam passes through an aperture so that our signal comes from a reasonably homogeneous volume within the lamp.

The microwaves are generated by a modulated frequency synthesizer, which is referenced to a voltage-controlled crystal oscillator (VCXO); the optimum output signal strength from the synthesizer is 4 dBm. The microwaves propagate out of a 15 dB gain horn, with the lamp and oven in the near field. After passing through the lamp vapor, the transmitted light is collected by a lens and focused onto a photodiode, where the detected signal is employed in a standard feedback loop to lock the VCXO output frequency to the hyperfine resonance of the lamp’s Rb atoms.

For a resonant frequency of 3.3843.2 Hz, we infer an actual magnetic environment from our results. If, for example, the lamp is placed and re-placed in the measurement system, reflection and scattering differences of the laser light do not lead to systematic variability in pressure assessments. After passage through the lamp, the expanded laser beam passes through an aperture so that our signal comes from a reasonably homogeneous volume within the lamp.

The 10 MHz output frequency from the VCXO also proceeds to a mixer, where it is combined with the output from a low phase-noise (Fluke) synthesizer that is referenced to a Cs clock. Each day that we perform a pressure measurement, the Cs clock is referenced to UTC(GPS) using a GPS-disciplined Rb clock, and the Fluke synthesizer is adjusted to account for any difference between our Cs clock and UTC(GPS). The mixer’s output at ~ 1 kHz is sent to a frequency counter that is referenced to the GPS-disciplined Rb atomic clock, and from the absolute measurement of the mixer’s output frequency we can work backwards to determine the hyperfine resonance frequency for the Rb atoms in the lamp relative to UTC(GPS). From this measurement, Eq. (1) gives an absolute determination of the Xe pressure in the lamp.

In order to check for systematic variations over the course of our multi-month experiment, we have a set of five Xe “calibration” lamps containing 0.6, 1.5, 3.3, 10.1, and 14.4 torr of Xe. These lamps never experience the field of an rf-discharge, therefore we expect no change in their Xe pressures. Any apparent change is taken as a measure of systematic drift in our pressure measurement system.

The 0-0 hyperfine transition frequency [24], percent level changes in the magnetic field seen by the atoms will produce pressure errors on the order of millitorr. Achieving such long-term precision in a magnetic field environment can be a difficult proposition. To solve this problem, as illustrated in Fig. 3, we measure the resonant frequency in our lamps as a function of B, (i.e., the current through our Helmholtz coil). Since the 0-0 hyperfine resonance frequency is a quadratic function of B, even in the presence of a static laboratory field, we can easily extrapolate the resonant frequency results to zero (total) magnetic field. In this way we can remove long-term variations in the laboratory magnetic environment from our results.

IV. ACCURACY, SENSITIVITIES, AND STABILITY OF THE PRESSURE MEASUREMENT SYSTEM

A. Accuracy of the KSK Technique

To examine the accuracy of the KSK procedure, we used the technique to assess the pressure in a number of sealed cells from various manufacturers that were filled with different buffer gases and pressures. Several of these cells were constructed in a fashion similar to the lamps we intended to place under long term test (i.e., our calibration lamps); other cells were larger, and meant for a different series of experiments [25]. The complete list of cells is collected in Table I, where it may be noted that we limited the set to nominal fill pressures greater than ten torr. For fill pressures less than this, we had concerns regarding the manufacturers’ abilities to accurately fill the cells with buffer gas.

Figure 4 shows the KSK estimate of the pressure in the cell as a function of the nominal fill pressure quoted by the manufacturer. As anticipated, the relationship is linear with a slope of about unity and an intercept near zero. Obviously, some of the scatter about the regression line must be attributed to the manufacturers’ abilities to accurately fill the glass cells with buffer gas. Typically, these cells are placed on a glass manifold, then filled with alkali metal and a buffer gas of some pressure, and finally “pulled” from the manifold using a torch. During the pulling process, the glass cell is obviously heated, with the result that the density of buffer gas in the cell can change from its “cold” condition. Nevertheless, if, we assume that the manufacturer’s fill pressure is completely accurate, then we can upper bound the inaccuracy of the KSK technique. Those upper-bound errors suggest an absolute accuracy better than ~ 5%. Specifically, for our 3.3 torr calibration lamp, which we will discuss further below, the

![Figure 3: The 0-0 hyperfine resonance frequency in our 3.3 torr Xe calibration lamp as a function of B.](image-url)
The absolute accuracy of the KSK technique is better than ±0.17 torr.

**Table I**: List of sealed cells used to examine the accuracy of the KSK pressure measurement procedure.

<table>
<thead>
<tr>
<th>Buffer Gas</th>
<th>$P_{\text{nom}}$, torr</th>
<th>$P_{\text{KSK}}$, torr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ar</td>
<td>100</td>
<td>100.9</td>
</tr>
<tr>
<td>Kr</td>
<td>100</td>
<td>98.2</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>70.6</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>24.9</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>9.2</td>
</tr>
<tr>
<td>Xe</td>
<td>100</td>
<td>102.2</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>67.4</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>26.0</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10.1</td>
</tr>
<tr>
<td>N$_2$</td>
<td>100</td>
<td>97.6</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>47.5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>11.6</td>
</tr>
</tbody>
</table>

**Figure 4**: KSK estimate of pressure in sealed cells vs. manufacturer’s listed nominal fill pressure for the cells.

**Figure 5**: Sensitivity of our KSK pressure measurement system to relative laser-light intensity.

**Figure 6**: Sensitivity of our KSK pressure measurement system to microwave power.

**B. Sensitivities of the KSK Technique**

Figure 5 shows the change in estimated pressure of our 3.3 torr calibration lamp as a function of relative laser intensity. Overall, the estimated pressure is a nonlinear function of laser intensity, which is perhaps not too surprising given the fact that our measurements are likely influenced by the inhomogeneous light shift [26]. Nevertheless, under nominal conditions the estimated pressure only varies by ~0.2 mtorr/%.
temperature to ±0.5 °C, we should be able to make consistent pressure measurements at the level of several millitorr.

**Figure 7**: Sensitivity of our KSK pressure measurement system to resonance-lamp temperature.

C. **Long-Term Stability**

Figure 8 shows the estimated pressure in our 3.3 torr calibration lamp as a function of time. Measuring for over a year, we find no evidence of a systematic drift in our measurement system at the level of mtorr/month: \( \frac{dP_{Xe}}{dt} = (0.3 \pm 0.4) \text{ mtorr per month} \). The standard deviation of \( P_{Xe} \) over this period is 7 mtorr, implying a reproducibility of Xe pressure measurements over the long term of 0.2%.

**Figure 8**: Long-term stability of our pressure measurement system using a 3.3 torr Xe calibration lamp.

V. **SUMMARY**

In this work we have discussed our system for measuring buffer-gas pressure in sealed atomic resonance cells. In particular, our system has a reproducibility of 0.2%, and an absolute accuracy likely better than 5%. Moreover, any systematic drift in the pressure measurements appears to be less than one mtorr/month. Though our specific system was designed to investigate the physics/chemistry behind buffer-gas loss in alkali rf-discharge lamps, it is clear that systems similar to ours could be employed for measuring buffer-gas pressure in any atomic resonance cell.

**REFERENCES**


Majorana atomic transition research in H-maser’s magnetic state selection region

Aleynikov Mikhail

Time and Frequency Department, National Research Institute for Physical-Technical and RadioTechnical Measurements
VNIIFTRI, Moscow Region, Russia
e-mail: alejnikov@vniiftri.ru

Abstract — For efficient performance of a single-state selection system when using Majorana method, it’s extremely important to know how eventual angle of rotation for atom’s spin depends on a total magnetic field in a region between selection magnets. As it was shown in previous work, the angle of rotation for spinor F = 1 greatly depends on a transverse (in regard to the hydrogen beam axis z) displacement of the point where total field is zero. In this work, for the first time, the dependence of single-state selection system performance (or H-maser’s output power) on the currents of the transverse coils, that are placed at the region between magnets, is experimentally obtained. On the dependence, the angles of rotation for spinor that are equal to 0, π/2, π under corresponding quantity of the currents, i.e. transverse coordinates of the zero field, are explicitly determined. The optimal values of the currents, when the maximum of the single-state selection efficiency (the angle of rotation is equal π) is achieved, are defined. Moreover, the operation of the single-state selection system is confirmed by exploration of the H-maser’s power curve and also by double resonance method. The relative amount of the operating atoms in the beam, that defines single-state selection system efficiency, is approximately 70%.

Key words — H-maser; Majorana effect; state selection system; double resonance method.

I. INTRODUCTION

A quality of H-maser’s magnetic selection system performance manipulates its output power level and thus H-maser’s metrological properties [1-3]. A perfect operating state selection system increases H-maser’s output power and decreases its phase noise degree what may be used for synthesis of clock transition frequency, that lies in the microwave range [4].

In the previous work [4] the Majorana method provided a single-state selection operation was theoretically examined in details. In that paper the influence of the total field parameters on the state selection factor was explored. In particular the radial (transverse) displacement of the total field’s zero is one of significant parameters in that problem. This fact needs careful design of H-maser’s lower chamber in the region between selecting magnets. Advancing this question, in the present paper H-maser which has the lower chamber with four cylindrical layers of the magnetic shields in the area between magnets is experimentally investigating. For fine adjustment of the total magnetic field except Majorana coils there are two mutually orthogonal pairs of Helmholtz coils creating field in the radial plane and being inside shields.

Due to current tuning on these additional coils a shift of the total field’s zero occurs, this fact affects selective magnetic system factor and hence amplitude of H-maser’s exit signal. Thus, under measuring the out signal magnitude it’s possible to determine the atom spin’s rotation angle in the magnetic field created in the region between selective magnets when the all coil currents are fixed.

II. THEORY

The atomic beam intensity is proportional to H-maser’s output power delivered to the receiver or signal-noise ratio in 1 Hz bandwidth measured on the signal analyzer, and has the conventional expression [5]:

\[
\frac{P}{P_c} = -2q^2 \left( \frac{\Delta I}{I_{thr}} \right)^2 + (1 - 3q) \frac{\Delta I}{I_{thr}} - 1
\]

(1)

here \( I_{thr} \), \( P_c \) – threshold values of the atomic beam flux and power, \( \Delta I \) – efficient flux of the operating atoms entered in the storage bulb, equals \( I_{F=1,m=0} - I_{F=0,m=0} \), \( q \) – quality quantity described spin-exchange interaction in the storage bulb and defined as:

\[
q = C \frac{I_{tot}}{\Delta I}
\]

(2)

where \( C \) – constant depends on the set of the H-maser’s parameters, \( \alpha = I_{tot}/\Delta I \) – quality factor of the magnetic selection system, equals the flux ratio of the total atoms entered into the storage bulb to the efficient operating atoms [6].

Equations (1) – (2) describes the H-maser’s out power dependence on the total and efficient fluxes and also on the quality factor \( \alpha \). From (1) – (2) it can be shown out power increases and total flux decreases while parameter \( \alpha \) is decreasing and efficient flux being constant. As it was studied in [7] the analogous power dependence on the inverse quality factor of the line occurs.

The examination of the H-maser’s out power curve (1) – (2) is not only way to define \( \alpha \). Another method is associated with double resonance process which consists in using an additional transverse magnetic field in the storage bulb region \( H_T = H_z e^{i\omega t \cos(\omega z + \delta \Omega t)} \) near the Zeeman frequency. If the states of the hydrogen ground state hyperfine structure in magnetic field are denoted as \( |F = 1, m_F = 1\rangle \equiv |1\rangle, |F = 1, m_F = -1\rangle \equiv |2\rangle \),

magnetic transition.
m_F = 0 \equiv |2\>, |F = 1, m_F = -1\> \equiv |3\>, |F = 0, m_F = 0\> \equiv |4\> the interaction Hamiltonian takes the following form:

\[ H_{\text{int}} = H_{24} |2\>\<4| + H_{12} |1\>\<2| + H_{23} |2\>\<3| + h.c. \]  

(3)

The density matrix equation described the dynamic of the atomic ensemble in the storage bulb has the form:

\[ \dot{\rho} = i\hbar [\rho, H_0] + \frac{i}{\hbar} [\rho, H_{\text{int}}] + \dot{\rho}_{\text{flux}} + \dot{\rho}_{\text{relax}} \]  

(4)

here \( H_0 \) the Zeeman’s splitting, \( \rho_{\text{flux}}, \rho_{\text{relax}} \) the dynamic terms described atomic flux into the bulb and atomic relaxation in the bulb. Solving (3) – (4) assumed as a basis, Andresen [8, 9] found that to second order in the transverse field Rabi frequency \( X_{12} = \mu_{12} H_T / \hbar = X_{23} \) the small static field limit of the maser shift \( \Delta \) is given by:

\[ \Delta = - |X_{12}| \frac{\bar{\mu}}{r}(Y_2 Y_2 + |X_{24}|) \frac{\delta(p_{12} - p_{43})}{(Y_2^2 - \delta^2) + \left(\frac{1}{4} |X_{24}|^2\right) + (2\delta Y_2)^2} \]  

(5)

where \( X_{24} \) unperturbed maser Rabi frequency, \( Y_2 \) population decay rate, \( Y_2 \) hyperfine decoherence rate, \( Y_2 \) Zeeman decoherence rate, \( r \) atom flow rate, \( p_{12} - p_{43} = r / (2 Y_2) \) the steady state population difference between states \( |1\> \) and \( |3\> \).

Thus the maser shift induced by applied radiation near the Zeeman frequency has a discrimination shape with magnitude which is proportional to the flow difference between states \( |1\> \) and \( |3\> \). It should be noted that in the case of the ideal magnetic state selection performance this dependence is neglected.

### III. Experimental Results

All experiments have been fulfilling on the single H-maser setup. The main parameters of the higher vacuum chamber have the next values: \( Q_c = 35,5 \times 10^7 \) – loaded cavity quality factor, \( V_c = 17,5 l \) – cavity capacity, \( V_b = 2,95 l \) – bulb capacity; \( r \approx 1 s^{-1}, Y_2 \approx 1,44 s^{-1}, X_{24} \approx 1,1 s^{-1} \). In the lower chamber two selective 6-pole magnets 60 mm long, 7 kG and 10 kG magnetic strength near pole tips are installed; the distance between magnets equals 120 mm. In the region between the magnets two equal Majorana coils are about 100 mm apart from each other. These coils have the 35-mm diameter, the number of turns about 200. The coil axis agrees with atomic beam direction. Two pairs of additional transverse Helmholtz coils are equidistantly from the Majorana coils installed. The coils have 20-mm diameter and the number of turns about 100. The region of the all coils between the magnets is enclosed by the four layer cylindrical magnetic shields.

H-maser’s out power dependence on the transverse coil currents when Majorana coil currents being equal to \( I_M = 40 mA \) is shown in figure 1. The level of the output signal under the constant atomic flux from the hydrogen source. When the turning of the spin to the angle of \( \pi / 2 \) the number of the operating atoms decreases, but the number of the undesired atoms increases in comparison with the previous case. This fact leads to the additional spin-exchange interaction in the storage bulb, that corresponds to the minimum limit level of the H-maser’s power. When the turning is absent, what occurs in the region of the significant values of the transverse coil currents (the region of the zero field is greatly shifted from the beam axis), the number of the operating atoms remains equal to the case of the \( \pi \)-angle, but the number of the undesired atoms increases. The first case is the most arduous in practice. In the plane of the transverse currents (or transverse displacements of the zero total field) the case corresponds to the small circle with center in which zero of the field concurs with atomic beam axis.

To confirm the correct behavior of the single-state selection system the two independent experiments have been performed. In the first one the out H-maser’s power were measured versus the inverse line quality factor. The two curves are shown in figure 2: the first one corresponds to the turning of the spin to the angles of \( 0 \) and \( \pi / 2 \) (the single-state selection is switched off) since these cases concur, the second one corresponds to the \( \pi \)-angle turning (the single-state selection is switched on) (Fig. 1). When the spin-exchange interaction in the storage bulb is invariable and the atomic flux from the hydrogen source is great the output power level differs between the two single-state selection operating modes in 1,8 times. The fulfilled numerical calculation of the atom's trajectories [10] with mentioned low chamber’s parameters shows that \( \alpha_{eff} / \alpha_{an} \approx 1,4 \).

The maser shift \( \Delta \omega \) due to the detuning from the Zeeman frequency \( \Delta f_z \) of the transverse magnetic field \( H_T \) when the single-state selection is turned off and turned on is shown in figure 3. The measurements were performed at the great hydrogen flux when the coil’s currents implement the required turning (Fig. 1), the curves for the angles of \( 0 \) and \( \pi / 2 \) concurs since \( \alpha_{eff} = \alpha_0 \). The parameters of the discriminant curve approximating the experimental data have the values \( |X_{12}| = 0,72 s^{-1}, Y_2 = 1,15 s^{-1} \). As it is shown from the curves in figure 3 the steady state population difference between states \( |1\> \) and \( |3\> \) varies like \( (p_{12} - p_{43})_{eff} / (p_{12} - p_{43})_{an} \approx 5 \). The results of the numerical calculation when the single-state selection is turned...
off have the following quantities: $\rho_{11} = \rho_{22} = 0.43, \rho_{33} = \rho_{44} = 0.07$. The steady state population of the hyperfine structure states when single-state selection is turned on – $\rho_{11} = 0.19, \rho_{22} = 0.67, \rho_{33} = 0.12$.

IV. CONCLUSION

It is essential to carefully accomplish the area of the total magnetic field’s sign changing in the region between the selective magnets in order to perform the correct single-state selection operation based on the Majorana effect. In the present paper for this goal the four-layer cylindrical magnetic shields and two pairs of the transverse Helmholtz coils were used. The turning of the atom spin versus the transverse magnetic field was experimentally defined at the H-maser setup. By means of the current adjustment of the transverse and the Majorana coils the current configuration provided correct operation of the single-state selection system and replied to the spin turning of $\pi$ radian is found. Also to confirm the correct operation of the single-state selection the two independent experiments were performed. The experiments show the relative number of the desired atoms in the atomic beam is approximately 70% when the single-state selection is switched on. Moreover the H-maser’s output power increases approximately in 1.8 times, this fact leads to its improved metrological characteristics.

REFERENCES


Noise Investigation on Optical Detection in a Cesium Beam Clock with Magnetic State Selection

Liu Chang, Zhou Sheng, Wang Yanhui
Institute of Quantum electronics, School of Electronic Engineering and Computer Science, Peking University, Beijing, China
E-mail: wangyanhui@pku.edu.cn

Abstract—Noise sources in optical detection of a magnetic-state-selection cesium beam clock are analyzed in this paper. Atomic shot noise, photon shot noise, laser frequency noise and stray light noise are considered. Experimental measurements and estimations of the noise magnitude are made.

Keywords—frequency standard; cesium clock; laser; noise;

I. INTRODUCTION

We reported the use of optical detection in a cesium beam tube where state selection is accomplished by magnets [1]. Short-term frequency stability of $1.0 \times 10^{-11} \cdot \tau^{-1/2}$ is achieved. Theoretical and experimental results on short-term stability in frequency standards of different types are reported, such as in [2-5]. The short-term stability of an atomic frequency standard is dependent on the line width and the signal-to-noise ratio. The atomic-shot-noise limit is believed to be reached in a well-tuned magnetic-state-selection beam tube. When optical methods are applied to improve the performance, additional noise arises. Noise sources in an optically pumped cesium beam tube are analyzed [6]. In this paper, we give a description the main noise sources in the optical detection in our cesium beam clock. The noise power spectrum density measurement under different conditions helps with estimation of noise magnitude.

II. EXPERIMENTAL SET-UP

As shown in Fig. 1, cesium atoms are deflected by the inhomogeneous magnetic field caused by magnet A. Double beam scheme is used, i.e. two atomic beams are emitted and deflected symmetrically, which is not shown in the figure. The microwave cavity has two interaction lengths of 1 cm each and a microwave-free distance of 17 cm. An external cavity diode laser with line width of 100 kHz, which is frequency-locked to $(F = 4 \leftrightarrow F' = 5)$ transition of cesium D2 line with help of saturated absorption spectroscopy, serves as the light source. The laser-induced fluorescence of the atomic beam is collected by a pair of spherical mirrors and detected by a photodiode, which is different from the storage bulb in [1]. The photo current is converted to voltage signal with an equivalent resistance of $10^7$ ohm. The fluorescence signal is processed by a digital servo system similar to that in [7]. Power spectrum density of the fluorescence signal is measured with a digital data acquisition card from National Instruments, whose sample rate is set as 20 kHz. Square wave frequency modulation on the microwave input is adopted where the modulation frequency is 78 Hz. Allan deviations of the frequency output are measured against an H-maser, whose short-term stability is believed to be several-fold better than this clock.

III. NOISE IN THE DETECTION

A. Signal-to-noise ratio

The general equation for estimating short-term frequency stability in a frequency standard is

$$\sigma(1s) \propto \frac{1}{S/N \nu_0} \cdot \Delta \nu, \quad (1)$$

$\Delta \nu$ and $\nu_0$ are respectively the line width and central frequency. $S/N$ is the total signal-to-noise ratio which is defined as

$$S/N = \frac{U_o}{\sqrt{S_N}}. \quad (2)$$

$U_o$ is the amplitude of Ramsey spectral line. $S_N$ is the noise power spectrum density when there is only white noise, which is usually the case when we consider short-term frequency stability. The noise powers from different sources are
considered to be statistically independent. Hence their power spectrum density could be added to obtain the total noise power. The total signal-to-noise ratio can be written as

\[ S/N = \frac{1}{\sqrt{\sum_i (S/N_i)^2}} \]  

where \( S/N_i \) is the partial signal-to-noise ratio when only one noise source is taken into account.

### B. Atomic shot noise

In a thermal atomic beam device, the atoms arrive at the detection region with a particular rate, depending on oven temperature. The number of atoms arrived during certain period is modeled by Poisson process. The variance of atom number arriving during time \( \tau \) is

\[ \text{Var}(N_\alpha) = I_\alpha \tau . \]

\( I_\alpha \) is the rate of atom arrival, i.e. average number of atom arriving in one second. All \( F = 4 \) atoms at the detection region are detected, including both the effective atoms that undergoing microwave interrogation, and the residual \( F = 4 \) atoms due to the imperfection of state selection, the rate of which are defined as \( I_\alpha \) and \( I_r \) respectively. By the way, \( I_\alpha \) is proportional to the difference between the number of \((F = 3, m_F = 0)\) and \((F = 4, m_F = 0)\) atoms after magnet A, because they both interact with microwave and the actual Ramsey spectrum is a subtraction of them.

The partial signal to noise ratio of atomic shot noise is

\[ (S/N)_\alpha = \frac{I_\alpha}{\sqrt{I_\alpha + I_r}} . \]

Theoretically derivation of the ratio between \( I_\alpha \) and \( I_r \) is impractical because the lack of knowledge of exact atom trajectory near magnet A. The spectrum of atomic beam when sweeping the detection laser’s frequency is shown in Fig. 2. The spectrum is taken at laser power of 1.5 mW and oven temperature of 100°C. It is seen that the ratio is approximately \( I_\alpha / I_r = 4.5 \). This ratio varies little with the oven temperature which is experimentally verified, while \( I_\alpha \) and \( I_r \) can both be enhanced with high oven temperature.

### C. Photon shot noise

The fluorescence photons emitted by atoms are detected by some probability because the finite efficiency of the light collector and the photodiode. This can be equivalently treated as Poisson process as well, where the rate of the arrival of photons is

\[ \alpha + \beta = \lambda, \quad I_p = \beta \eta I_\alpha . \]

\( \beta \) is the average number of photons emitted by an atom. \( \eta \) is the detection efficiency of each photon. The partial signal-to-noise-ratio is

\[ (S/N)_p = \frac{\sqrt{\beta \eta I_\alpha}}{\sqrt{I_\alpha + I_r}} = \sqrt{\beta \eta (S/N)_\alpha} . \]

This noise is negligible compared with the atomic shot noise when \( \beta \eta \) is much greater than unity, which is easily satisfied when \((F = 4 \leftrightarrow F' = 5)\) transition, the cyclic transition, is used for detection.

### D. Laser frequency noise

For the atomic beam’s laser spectrum is not flat around \((F = 4 \leftrightarrow F' = 5)\) transition, fluctuations of the laser frequency will result in fluctuations of the fluorescence signal. This is an effect that correlates all \( F = 4 \) atoms in the detection region, making the noise power proportional to \( (I_\alpha + I_r)^2 \), which leads to a signal-to-noise ratio

\[ (S/N)_\ell \propto \frac{I_\alpha}{I_\alpha + I_r} . \]

independent of oven temperature [6, 8]. This indicates that a limit of short-term stability usually emerges when the oven temperature is high, because the laser frequency noise increased faster than atomic shot noise when the atomic beam flux intensity is large.

### E. Stray light noise

Fluctuations in laser intensity have two influences. One is on the average photon number emitted by an atom and the
other is on the stray light which is directly detected. The former is relatively small because the fluorescence is much weaker than scattered light, which is seen by the comparison of the peak and background in Fig. 2. Its property of independence of atoms makes the measurement of the stray light noise possible.

IV. MEASUREMENT AND ESTIMATION OF THE NOISE

A. Stray light noise measurement

In this experiment, the laser frequency is locked to the crossover line between \( (F = 4 \leftrightarrow F' = 4) \) and \( (F = 4 \leftrightarrow F' = 5) \) in a saturated absorption spectrum. Because the atomic beam spectrum is Doppler-free, laser locked to the crossover line hardly excites any atoms. The noise here is mainly the stray light noise. Fig. 3 shows the noise power spectrum density with respect to the background voltage, which is proportional to the power of the incident light. By the way, the noise power of the detection system is about one order of magnitude less than the measured value and has been subtracted from all the data in this paper.

It is experimentally verified that \( U_s \) increases linearly with laser power when the power is small and saturates when laser power is large. Hence this partial signal-to-noise will show a peak when the laser power is tuned. However, for its relatively small power density compared with the atomic shot noise and laser frequency noise, this noise is almost negligible with a large range of laser power when the oven temperature is not too low.

B. Laser frequency noise measurement

Measurement of the noise power spectrum density at various oven temperature is made when the laser is locked to \( (F = 4 \leftrightarrow F' = 5) \) transition. The measured stray light noise power has been removed from the data. The power densities are plotted with respect to \( U_s \), the amplitude of Ramsey spectral line in Fig. 4. In this case, the laser frequency noise and the atomic shot noise are coupled. However, they are proportional to different orders of atomic flux intensity as mentioned earlier. Fitted quadratic line of Fig. 4 indicates that the laser frequency noise is dominating within the temperature range in this experiment, from 95°C to 135°C. By direct measuring the frequency stability, we see that the short-term stability varies little within this temperature range, which is in accordance with the theory.

V. CONCLUSION

Noise sources in optical detection are analyzed in a cesium beam clock with magnetic state selection. The power spectrum densities of fluorescence signal are measured for estimation of the noise. This method is easy to realize and is possible for cases when the local oscillator cannot be locked. It is theoretically shown that the large ratio of \( I_s / I_r \) affect both atomic shot noise and laser frequency noise. The second-order relationship makes its influence on the laser frequency noise more serious, which leads to the limit of short-term stability at oven temperature above 95°C. Improvement on selection method that reduces this ratio will help to effectively enhance signal-to-noise ratio at high oven temperature.
REFERENCES


The Effect of Bend on the Ramsey Cavity

Fuyu Sun*, Xianhe Huang
School of Automation Engineering
University of Electronic Science and Technology of China
Chengdu, China

Abstract—Ramsey cavity is one of core components that compose the Cs beam tube in the Cs atomic beam clocks. In this work, the contribution from the waveguide bend on the field distribution of the cavity is carefully investigated by using combination method of Maxwell equations and Finite element simulation. We find that there exists TM11 mode inside cavity in addition to standing wave TE10p mode. Meanwhile, we also find that the cavity resonance frequency is closely related to the bend radius. These results demonstrate a better description of the microwave properties than previous work where the Ramsey cavity was usually studied as an ideal rectangular microwave cavity.

Keywords—atomic clock; ramsey cavity; waveguide bend

I. INTRODUCTION

The small cesium atomic clocks (SCACs) are the basis for many aspects of modern society, especially in measurement, navigation systems, and global communication. The related research work has been going on for several decades [1-6]. The Ramsey cavity is one of core components that compose the cesium beam tube in the two classes of SCACs [7]. Fig.1 is a structure of the Ramsey cavity typical of those used in traditional cesium beam tube, it is usually bent in the E-plane of the standard X-band waveguide and closed at the extremities. There are two holes which enable the atoms to cross the oscillating field region near the two cavity ends, and the two cut-off waveguides are attached to the cavity to prevent microwave leakage from beam holes.

The principal purpose of the Ramsey cavity is to provide separated oscillating microwave fields which extract the clock transition frequency [8-12]. Ideally, the microwave magnetic field should be a pure standing wave, its amplitude is sustained at every point of the cesium beam path in the cavity, and the cavity operating frequency depends only on the cavity longitudinal length for a given dimension of the transverse cross-section [12]. However, the aforementioned properties are based on assumption that the Ramsey cavity is an ideal rectangular cavity (ideal box). In this paper, it is pointed out that this is obviously not the case in an actual Ramsey cavity. In addition to standing wave TE10p mode, theoretical analysis shows that there exists non-standing wave mode in the Ramsey cavity. Comparisons of the microwave field between HFSS simulation and theoretical calculation are presented, it is pointed out that it is necessary to take into account mode in order to describe microwave field correctly. Finally, a simulation result of the cavity operating frequency is investigated as a function of the drift region length for different bend radius. The results in this paper provide new insights into the microwave properties of the Ramsey cavity.

Fig. 1. Structure of the Ramsey cavity

II. THEORETICAL ANALYSIS

The most important designing principle of the Ramsey cavity is to ensure that the cavity operating frequency is exactly the same as the clock transition frequency \( \nu_c = 9.19263177 \text{GHz} \), and the amplitude of the microwave field does not vary along the cesium beam path. Firstly, the reference coordinate system is established as shown in Fig. 1, the cesium atoms travel across the interaction region along y-direction, the dimension of the transverse cross-section of the cavity is \( a \times b = 22.86 \text{mm} \times 10.16 \text{mm} \) such that the TE10p mode is the only standing wave mode in the cavity according to the single-mode transmission condition, where \( p \) denotes the number of longitudinal modes at the clock transition frequency, mainly depending on the requirements of the length of the drift region of SCACs, is generally between 8 and 11 in a short Ramsey cavity. If the cavity operating mode is assumed to be pure standing wave TE10p mode, then the transverse magnetic field in right straight arm that excites the clock transition is \( H_x = \sin(\pi x/a)\cos(\pi y/b) \), where \( L \) represents the total longitudinal length of the cavity. Now, we let the transverse magnetic filed in the bend section is \( H_x = X(x)R(r)\Phi(\theta) \), since each field component is
independent of \( r \) (or \( y \)) on interface \( z = l \), where \( l \) is the length of the two straight arm, thus \( R(r) \) must be constant. Using variable separation method, we solve Helmholtz equation \( \nabla^2 H_{s-TM} + k^2 H_{s-TM} = 0 \) for bend section in cylindrical coordinate, then we have: 
\[
\rho^2 \left( \frac{d^2}{d\rho^2} + \frac{1}{\rho} \frac{d}{d\rho} \right) \gamma_{\text{bend}} - \nu_{\text{bend}} = 0 ,
\]
where \( d^2 X / dx^2 = -\gamma_{\text{bend}} X \), \( d^2 \Phi / d\theta^2 = -\nu_{\text{bend}} \Phi \), \( \gamma_{\text{bend}} \) and \( \nu_{\text{bend}} \) are the propagation constants with respect to the direction of \( x \) and \( \theta \) in the bend, and \( k \) is the wavenumber of the microwave in vacuum [13]. For a given size of the transverse cross-section of the Ramsey cavity, \( \gamma_{\text{bend}} \), \( \nu_{\text{bend}} \) and \( k \) must be constant, consequently, the solution of the equation above does not exist. As discussed above, because of the Ramsey cavity contains the bends, as a result, in order to guarantee continuity of the field between straight arm and bend section, non-standing wave mode in addition to standing wave \( TE_{10p} \) mode must be involved in the Ramsey cavity.

Since we are only concerned about the field distribution of interaction region where the atoms are submitted to the magnetic field and clock transition occurs, thus the right straight arm therefore has been chosen as studying object. Note that the standing wave mode is still the fundamental mode, and keep its period number in the cavity. In this case, the non-standing wave mode induces a perturbation of microwave properties, resulting in the change of the electromagnetic field distribution and the operating frequency of the Ramsey cavity. According to boundary conditions on the perfect conductor, electromagnetic field of any mode in the transverse cross-section shows the standing wave distribution, and then the net transverse magnetic field that excites the clock transition can be expressed as:

\[
H_x = H_0 \sin \left( \frac{\pi x}{a} \right) \cos \left( \frac{p\pi}{L_{\text{eff}}} z \right) + H_{s-TM} \quad (1)
\]

Where \( H_{s-TM} \) denotes the transverse magnetic field of non-standing \( TM_{mn} \) mode, and we define \( L_{\text{eff}} \) as the total longitudinal effective length of standing wave \( TE_{10p} \) mode. The corresponding longitudinal electric field component of \( TM_{mn} \) mode can be written as follows:

\[
E_z = \sum_{m,n} E_{mn} \sin \left( \frac{m\pi x}{a} \right) \sin \left( \frac{n\pi y}{b} \right) \left( e^{in\gamma_{\text{bend}}} + e^{in\gamma_{\text{bend}}} \right) \quad (2)
\]

Where \( \alpha = \pi \sqrt{(m/a)^2 + (n/b)^2 - (2f/c)^2} \) is the longitudinal cutoff constant of \( TM_{mn} \) mode, \( (m,n) \) are model indexes, note that, \( mn \neq 0 \) for \( TM_{mn} \) mode, \( f \) and \( c \) is the cavity operating frequency and the speed of light in vacuum, respectively. Obviously, the magnitude of \( H_{s-TM} \) changes exponentially with \( z \).

If the design is carried out by using an ideal box model, then \( H_{s-TM} = 0 \), \( L_{\text{eff}} = L = p\lambda_c/2 \), where \( \lambda_c = 46.5 \text{mm} \) is the waveguide wavelength of \( TE_{10p} \) mode in the straight arm. In order to obtain cavity operating frequency, the only variable we need to consider is the wavenumber \( p \). However, the method mentioned above often leads to a significant deviation between the experimental value and the theoretical expectation after the fabricating of the Ramsey cavity, as a result, one has to change the cavity length by adjusting the position of the two ends to meet the requirements of SCACs. In other words, \( L \) is no longer suitable as the total cavity length, because the waveguide wavelength in the straight arm and in the bend is not the same in an actual Ramsey cavity. As defined above, \( L_{\text{eff}} \) becomes the effective cavity length instead of \( L \). Since the cavity operating frequency for all modes is the same, and thus, for \( TE_{10p} \) mode, we have

\[
\frac{1}{a^2} + \frac{p^2}{L_{\text{eff}}^2} = \frac{4f^2}{c^2} \quad (3)
\]

In order to determine the operating frequency exactly the effective length \( L_{\text{eff}} \) of the Ramsey cavity must be known. However, for a given total length \( L \), we can not calculate the cavity operating frequency \( f \) directly like we did in an ideal rectangular cavity, since the effective length \( L_{\text{eff}} \) is not known yet. A more complete analysis of the microwave properties than ever presented here is quite helpful in understanding the Ramsey cavity.

### III. Simulation Results and Discussion

#### A. TM\(_{11}\) Mode Inside Cavity

In order to estimate the non-standing wave mode that may be exist in the Ramsey cavity we investigate the field distribution in detail by using the commercial finite element software Ansoft HFSS. The E-bent cavity which is most often implemented in cesium beam tubes of SCACs has been taken as an example (see Fig. 1). In this case, the phase difference between the oscillating fields at the two extremities of the cavity is equal to zero. Its fundamental mode is standing wave \( TE_{108} \) mode with \( p = 8 \) and the total longitudinal length \( l = 4\lambda_c = 186 \text{mm} \), the dimension of the cross-section of the beam hole is \( 3\text{mm} \times 6\text{mm} \), the length of the drift region is \( l_{\text{drift}} = 110\text{mm} \), and we define \( r_{\text{bend}} \) as the mean radius of the curvature of the bend. According to the field theory, to study the properties of non-standing wave \( TM_{mn} \) mode we only need to consider its longitudinal electric field component, \( E_z \), since \( TE_{108} \) mode has not contribution to \( E_z \).

For \( x = 7 \text{mm} \) and \( y = 3 \text{mm} \), the simulation results of the electric field on the \( z \) are shown in Fig. 2(a,b) for the \( r_{\text{bend}} = 6 \text{mm} \) example. In Fig. 2, approximately, the transverse electric field \( E_z = 0 \) and \( E_z \propto \sin(2\pi z/\lambda_c) \)
mainly represent the features of the standing wave $TE_{108}$ mode, it is obvious that the longitudinal electric field $E_z \neq 0$ shown in Fig. 2(c), as a result, the presence of non-standing wave $TM_{mn}$ mode is confirmed. According to Eq. (2), a significant contribution of $TM_{mn}$ mode to the net microwave field can be expected when $z$ is close to $l$, thus, as an additional proof of $TM_{mn}$ mode, the dependence of the longitudinal electric field $E_z$ on the $x$ and $y$ are plotted for $z = 26\,\text{mm}$, $y = 3\,\text{mm}$ and $z = 26\,\text{mm}$, $x = 7\,\text{mm}$ in Fig. 3, respectively. From Fig. 3 we have $m = n = 1$.

**B. Evaluation of Microwave Magnetic Field Inhomogeneity**

The microwave field experienced by the cesium atoms depends on the net transverse magnetic field, which includes the standing wave field and non-standing wave field. According to Eq. (4), the presence of the $TM_{11}$ mode introduces inhomogeneity of the microwave magnetic field along the cesium beam path. In order to evaluate quantitatively the field inhomogeneity, the longitudinal electric field $E_z$ as a function of $z$ for $x = a/2$ and $y = b/2$ is extracted from HFSS, as shown in Fig. 4(a). By using the magnitude of electric field at a certain point, e.g., $z = 29.5\,\text{mm}$, we have approximated $E_{11} = -0.39 \times 10^{-4} \,\text{V} / \text{m}$, and thus, the theoretical calculation results can be obtained from Eq. (2), as shown in Fig. 4(b). Fig. 4 shows that HFSS simulation results agree quantitatively with our theoretical calculation. So, again, the electromagnetic field distribution of non-standing wave $TM_{11}$ mode is confirmed in the cavity, and it works as the main weight of $TM_{mn}$ mode.

Fig. 5(a) shows the variation of $H_x$ versus $y$ (i.e., the coordinate along the cesium beam path) for $x = a/2$ and $z = 26\,\text{mm}$, it is obvious that the inhomogeneity is evoked by $TM_{11}$ mode. From Eq. (4), the $TM_{11}$ mode has no contribution to the net transverse magnetic field at $y = b/2$, so the transverse magnetic field of the standing wave $TE_{108}$ mode is about $-6.64 \times 10^{-4} \,\text{A} / \text{m}$. Let us now consider the influence of $TM_{11}$ mode on the uniformity. Using Eq. (1) and Eq. (4), the theoretical distribution of the net transverse magnetic field can be calculated, as shown in Fig. 5(b). Fig. 5 shows that HFSS simulation results are good agreement with results from theoretical model.

![Fig. 2. Electric field vs z.](image1)

![Fig. 3. Longitudinal electric field vs x, y.](image2)

As clearly shown in Fig. 2 and Fig. 3, the non-standing wave $TM_{11}$ mode must be involved in the Ramsey cavity, thus, the transverse magnetic field of $TM_{11}$ mode can be expressed as:

$$H_{x, TM} = \frac{2a^2bf}{a^2 + b^2} E_{11} \sin \left( \frac{\pi x}{a} \right) \cos \left( \frac{\pi y}{b} \right) (e^{-a x} + e^{a x}) \tag{4}$$

Where $\alpha = \pi \sqrt{\left(1/a\right)^2 + (1/b)^2 - (2f/c)^2}$, and $\varepsilon$ is vacuum permittivity. By inserting Eq. (4) into Eq. (1), the net transverse magnetic field $H_x$ experienced by the cesium atoms can be obtained immediately.

![Fig. 4. Longitudinal electric field vs z.](image3)
It clearly appears that the $TM_{11}$ mode in the Ramsey cavity is the main non-standing mode (see Fig. 4 and Fig. 5). It should be noted that there exists a slight difference between the two curves in Fig. 5, for this, a possible important reason is due to the existence of other non-standing wave mode in addition to $TM_{11}$ mode in the cavity. Despite the difference, it still presents a better approximation compared to the previous work where only the pure standing wave mode was considered. And thus, we can obtain quantitative evaluation for inhomogeneity of the microwave magnetic field at every point along the beam path for different bend radius. For example, in the center of the beam hole ($x = a/2, z = 3\text{mm}$), the variation of the microwave magnetic field as a function of $y$ can be calculated for $r_{\text{bend}} = 6\text{mm}$, we have: 

$$|H_x|/|H_x| = 2.25 \times 10^{-4} \cos \left(\pi y/b\right).$$

The stability of an atomic clock $\sigma \propto \Delta v/(S/N)$ is proportional to the clock signal halfwidth $\Delta v$, and $\Delta v \propto \alpha / \lambda_{\text{drift}}$ is inversely proportional to the drift region length $\lambda_{\text{drift}}$, where $S/N$ is the signal-to-noise ratio, and $\alpha$ is most probably velocity of the atoms in cesium beam tube. In general, $\lambda_{\text{drift}}$ is determined firstly under certain stability requirement of SCACs, so the inhomogeneity corresponding to different bend radius for a fixed drift region length is also investigated, we find that the field inhomogeneity is relatively small along the beam path when the bend radius is small, because the non-standing wave field can be sufficiently attenuated, thus minimizing the bend effects.

### C. Operating Frequency of the Ramsey Cavity

If the Ramsey cavity is considered as an ideal box, we have $L_{\text{eff}} = L$, Where $L_{\text{eff}}$ has the same meaning as in Eq. (1), according to Eq. (3). The cavity operating frequency only relates to the cavity longitudinal length $L$. However, as presented above, this is not the case in the Ramsey cavity. In addition to $L$, actually, the effective length $L_{\text{eff}}$ is a parameter which depends on both the drift region length $\lambda_{\text{drift}}$ and the bend radius $r_{\text{bend}}$. Unfortunately, there is no analytical solution available so far for the cavity operating frequency. Instead, using HFSS simulation, the bend radius contribution to the cavity operating frequency by varying the length of the drift region is investigated, as shown in Fig. 6. In a certain sense the operating frequency may can be regarded as the superposition of a constant frequency and an oscillation frequency, and the magnitude of the constant frequency and the amplitude of the oscillation frequency depend on significantly on the bend radius. We also find that the oscillation period of the oscillation frequency is approximately equal to $\lambda_{\gamma}/2$ according to the relation: $L = 2l + \lambda_{\text{drift}} + \pi r_{\text{bend}}$. Qualitatively we think that this is the result of longitudinal phase constant difference at different axial position.

In order to obtain the operating frequency, more attention is concentrated on the consideration of the total cavity length $L (p\lambda_{\gamma}/2)$ in the past, the selection of the cavity bend radius shows a great randomness. Nevertheless, as can be seen in Fig. 6, a frequency variation of more than $70\text{MHz}$ is observed for a given length of the drift region when the bend radius increases from $6\text{mm}$ to $15\text{mm}$. Furthermore, the frequency sensitivity is particularly significant in the case of a small bend radius, and the constant frequency and the amplitude of the oscillation frequency are shown to further increase as the bend radius decreases. In Fig. 6, the frequency change is relatively gentle when the bend radius is large enough. The results presented here provide a new understanding of the cavity operating frequency of different bend radius involved in designing Ramsey cavity.

### IV. CONCLUSION

This work shows that the microwave properties of the Ramsey cavity are critically linked to its bend radius. The bends will cause both the standing wave mode and the non-standing wave mode to determine microwave properties.
of the Ramsey cavity. Approximately, the key modes may be expressed as a sum of $TE_{10}$ mode and $TM_{11}$ mode. The microwave field distribution from the theoretical model agrees with the HFSS simulation, it is pointed out that it is necessary to take into account $TM_{11}$ mode in order to describe the field distribution correctly. Lastly, a simulation result of the cavity operating frequency change related to the drift region length under different bend radius is also investigated, intuitively, the operating frequency may be regarded as the superposition of a constant frequency and an oscillation frequency for a given drift region length. The obtained conclusions in this paper demonstrate a better description of the microwave properties of the Ramsey cavity, and the similar conclusions hold for the Ramsey cavity with a phase difference of $\pi$ between the two oscillating magnetic fields [14]. Our work may help in designing Ramsey cavity and realizing high performance cesium beam frequency standards.

REFERENCES

Design of the new NIM6 fountain with collecting atoms from a 3D MOT loading optical molasses

Fang Fang, Weiliang Chen, Kun Liu, Nianfeng Liu, Rui Suo and Tianchu Li
National Institute of Metrology (NIM), Beijing, China
E-mail: fangf@nim.ac.cn

Abstract—We report the design of a new cesium fountain clock NIM6, which is under construction in NIM. Besides some improvements on the vacuum system, Ramsey cavity and microwave synthesizer to reduce the Type B uncertainty. Another major improvement on NIM6 is to collect more atoms from a MOT loading optical molasses and optical pumping to get a better signal to noise ratio at the detection. The atom distribution will be more uniform compared with a 2D MOT loading optical molasses, and the diameter of the cloud can be adjusted by the intensity and detuning of lights during the post cooling to keep the collisional-induced frequency shift low. The atom numbers can be further increased by a new de-pumping - optical pumping procedure to pump atoms to the \(|F=3, m_F=0\rangle\) clock state directly. With a new cryogenic sapphire oscillator (CSO) based frequency synthesizer, NIM6 is aiming to reach the quantum projection noise, thus leading to a reduced Type A uncertainty compared with NIM5.

Keywords—Atomic fountain clock; frequency measurement; optical pumping

I. INTRODUCTION

Since the first laser-cooled cesium fountain clock developed in 1995 [1], many national metrology institutes worldwide started to build their own cesium fountain clocks. The recent reported Type B uncertainties of fountain clocks have reached down to a few parts in \(10^{16}\) [2-6]. In order to contribute to the maintenance of the global timescale UTC/TAI, all type of the uncertainties need to be included. Some works have been done to develop ultra-stable microwave synthesizers to reduce the Type A uncertainty to the quantum projection limitations. And a typical fractional frequency instability of a few parts in \(10^{16}\) \((\tau/s)^{1/2}\) is obtained. In such a case, it is very important to collect more number of atoms to push the uncertainty further down.

A new cesium fountain clock NIM6 is under construction in the National Institute of Metrology China. Besides some improvements on the design of the magnetic shielding and the vacuum system, a new shape of Ramsey cavity will be used to reduce the microwave leakage and distributed cavity phase induced frequency shift. Another major difference from NIM5 is that NIM6 collects atoms from a MOT loading optical molasses (OM), and an optical pumping will also be applied to increase the number of atoms, and lead to a better signal to noise ratio at the detection. A new method of optical pumping with a de-pump light is proposed to prepare atoms on the \(|F=3, m_F=0\rangle\) clock state directly without a state selection cavity. A new cryogenic sapphire oscillator (CSO) based frequency synthesizer is also under developing to reduce the phase noise of the microwave source, thus to improve the instability and reach the quantum projection noise.

II. THE DESIGN OF PHYSICAL PACKAGE

A cutaway of the new fountain physical package is shown in the figure 1. The entire physics package is enclosed in a layer of soft iron and the flight tube is surrounded with another three layers of \(\mu\) metal shielding with a shield factor of about \(10^5\). Atoms are collected in the lower MOT chamber and then launched to the upper optical molasses chamber with a small angle (10°) to reduce the background Cs atoms flying into the detection chamber directly. The lower MOT chamber is pumped by a 20 l/s ion pump and the upper OM chamber is pumped by a 40 l/s ion pump. A getter pump is on the top of the flight tube to keep the ultra-low pressure in the atom interrogation region.

Fig. 1: The vacuum system design of NIM6 with a MOT loading optical molasses.

The new fountain will be operated in a lab with the temperature fluctuation less than 0.3 K, and no active temperature control system will be added outside the flight tube. Instead, an isothermal liner will be surrounded the flight tube. It is a special type of heat pipe, made of a thin glass layer

---

This work is supported by the Ministry of Science and Technology of China (2013YQ09094301) and Chinese NSF (11174260)
of vacuum tube filled with pure water. The thermal conductivity of an isothermal liner is about 10 times higher than that of coppers. It also works like a low pass filter reducing the temperature fluctuations of the system. The temperature of the atom interrogation region will be more uniform and stable. With precision standard Pt thermometers measuring the temperatures of a few locations on the flight tube, the average temperature uncertainty is estimated to be less than 50 mK.

The new Ramey cavity has a loaded quality factor $Q_{\text{loaded}}$ approximately 8,000. Microwaves are coupled in through four 5 mm diameter holes evenly distributed in the horizontal directions from four rectangular waveguides. The amplitudes and phases on each side can be adjusted individually to reduce the residual travelling microwaves inside the cavity, thus to reduce the distributed cavity phase induced frequency shift. The connections between different parts of the cavity are sealed with indium to reduce the microwave leakage.

Cs atoms will be collected in the lower MOT chamber and launched to the center of the OM chamber. The separation between these two centers is 280 mm, the flying time is about 50 ms with a launching velocity of 5.5 m/s. The final temperature of the cloud after launching can be adjusted by the intensity and detuning of the post cooling beams to make sure the cloud expended enough when reaching the OM center in order to reduce the collisional shift. The atom velocity and temperature will be re-adjusted in the optical molasses region and then launched vertically to the flight tube. The advantage of this design is not only being able to collect more atoms compared to a direct optical molasses loading like NIM5, the background Cs gas in the detection chamber is also reduced due to a differential pumping. Furthermore, the atom density distribution is more uniform than loading OM from a 2D-MOT. Another feature here is that the cooling beams for MOT and OM are applied at different times. The lights for the MOT cooling, OM cooling can be provided by only one tapered amplifier with a total output power of about 800 mW. The output laser beam from a TA is split into three parts: two of them double pass two acousto-optical modulators (AOM1 and AOM2) respectively and are used for the MOT cooling beams. The third zero-order beams from these two AOMs double pass another two AOMs respectively and are used for the OM cooling beams. The third beam TA is used for the detection beams. The system setup is shown in figure 2.

### III. SELECTING CLOCK STATE WITH OPTICAL PUMPING

An expression for the instability of an atom clock is written as [7,8]:

$$\sigma_\tau = \frac{1}{\pi Q_{\text{loaded}}} \sqrt{T_c \left( \frac{1}{N_{\text{at}}} + \frac{1}{N_{\text{at}} N_{\text{ph}}} \cdot 2\sigma_\phi^2 + \gamma \right)}$$  \hspace{1cm} (1)

Where $T_c$ is the duration of the fountain cycle, $\gamma$ is the microwave oscillator noise, and the atomic quality factor $Q_{\text{at}}$ is about 9.2x10$^8$ (clock frequency of 9.2 GHz divided by the Ramsey linewidth of 1 Hz). The $N_{\text{at}}$ is the atom number at the detection. The first term in the brackets is the atomic projection noise. NIM5 collected atoms from an optical molasses directly, and about 10$^9$ atoms were collected in 600 ms (1.5x10$^6$ atoms at the detection [9]). In NIM6, atom numbers can be increased about 7 times by collecting atoms from MOT, with an additional optical pumping [10], 2x10$^7$ atoms can be easily obtained at the detection, which corresponding to a quantum projection noise limited instability of $1 \times 10^{-14}$ ($t/s^{1/2}$).

In a Cs fountain clock, atoms are approximately evenly distributed in the 9 Zeeman sublevels of the $|F=4>$ hyperfine state, and only atoms on the $|F=4, m_F=0>$ sublevel are selected to the $|F=3, m_F=0>$ clock state in a selection cavity to measure a clock transition. Normally, about 90% of the atoms which have been cooled and launched are removed by a radiation pressure pulse and lost from the fountain signal. The optical pumping has been used to accumulate atoms on the $|F=4, m_F=0>$ state in NPL’s fountain [11], and the atom numbers increased 4 times. Here, we propose a new method to do optical pumping. Right after launching, atoms are first de-pumped to the $|F=3, m_F=0>$ state by turning off the repumping light earlier than the cooling light during the post-cooling stage. If not enough number of atoms being de-pumped, a light on resonance with $|F=4> \rightarrow |F=3>$ transition can be applied. In this stage, the polarization of the light is not critical, and 75% of atoms will be de-pumped to the $|F=3>$ state by only scattering one photon. Then, atoms are optical pumped to the $|F=3, m_F=0>$ state with a linear polarized light to make a $|F=3> \rightarrow |F=3> \pi$ transition. In this situation, $|F=3, m_F=0>$ state is the “dark state” since the transition between $|F=3, m_F=0>$ and $|F=3, m_F=0>$ is forbidden by the angular momentum selection rules. The atoms excited to the $|F=3, m_F>$ states will decay spontaneously to $|4, m_F, |3, m_F>$ and adjacent sublevels. Those ending up in the clock state $|3, 0>$ will remain “dark” in this state and stop scattering photons. Eventually, atoms will be populated either on the $|F=4$ different Zeeman sublevels or $|F=3, m_F=0>$ clock state. More number of atoms on the clock state can be obtained if the de-pumping light is presented shortly during the optical pumping. In this circumstance, the average photon numbers scattered per atoms will increase, leading to a higher temperature of the atomic cloud. And some of will be lost before detection due to a q10 mm diameter aperture on the Ramsey cavity. The procedure needs to be optimized according to the detection signal. The advantage of the proposed method is not only increasing the atom number on the clock state, but also selecting the clock state by an optical pumping, and the state selection microwave cavity is not necessary anymore.

![Fig. 2 Schematics of the optical system setup for the MOT and OM cooling beams.](image-url)
Theoretical calculation from Clebsch–Gordan coefficients indicates that the atom population on the |F=3, m_F=0> sublevel reaches 92% with only scattering 2 photons per atom in an ideal case, as shown in figure 3(a). The atom distribution on each Zeeman sublevels of F=4 state is assumed to be 1. The theoretical results of atom distribution on the F=3 sublevel after optical pumping is shown in figure 3(b). The atom number on the clock state |F=3, m_F=0> is increased 2.3 times compared to the number with a routine method only selecting |F=4, m_F=0> state. The total population on the other m_F states is less than 3% of the atom number on the clock state, while the averaged scattering photon number per atom is less than 4. If more atoms are wanted, an additional de-pumper pulse can be applied shortly to excite atoms from the |F=4> back to |F=3> state, and then optically pumped to the |3, 0> clock state.

In NIM6, an external cavity diode laser will be locked to the crossover of F=3 $\rightarrow$ F'=3 & 4 transitions. The rest of the light is split into two beams: one is pass through an AOM to tune on resonance with the |F=3> $\rightarrow$ |F'=4> transition for the repumping of the MOT, OM and detection, the other beam pass another AOM to tune on resonance with the |F=3> $\rightarrow$ |F'=3> transition for the optical pumping. The de-pumping light will be provided by directing a small amount of light from the detection to pass through an AOM and tune on resonance with the |F=4> $\rightarrow$ |F'=3> transition.

IV. Summary

A new cesium fountain clock NIM6 is under construction in the National Institute of Metrology China. Besides some improvements on the design of the Ramsey cavity to reduce the distributed cavity phase shift and microwave leakage, NIM6 is also aiming to collect more atoms from a MOT loading optical molasses (OM) and optical pumping, leading to a better signal to noise ratio at the detection. A new method of optical pumping is proposed. The atoms are first de-pumped from F=4 to F=3 state, and then optically pumped and accumulated on the |F=3, m_F=0> clock state. The atom number on the clock state is increased and the state selection microwave cavity is not necessary anymore with this de-pumping - optical pumping procedure. A new cryogenic sapphire oscillator (CSO) based frequency synthesizer is also under developing to reduce the microwave phase noise in order to reach the quantum projection noise, thus leading to a lower Type A uncertainty.

Acknowledgment

The authors would like to thank Dr. X. Yan, Mr. P. Lin and Dr. K. Szymaniec for the helpful discussions. The work is supported by the Ministry of Science and Technology of China (2013YQ09094301) and Chinese NSF (11174260).

References

Advances in the atomic fountain clock at SIOM

Yuanbo Du*, Rong Wei†, Richang Dong*,†, Fan Zou*† and Yuzhu Wang*
*Key Laboratory of Quantum Optics, Center for Cold Atom Physics
Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Science
Shanghai, 201800, China
†University of Chinese Academy of Sciences, Beijing, 100049, China
Email: weirong@siom.ac.cn

Abstract—This work presents several advances in our $^{87}\text{Rb}$ atomic fountain clock (AFC) at Shanghai Institute of Optics and Fine Mechanics (SIOM). We directly lock the local oscillator (LO) of AFC, and make some improvements on the physical system, including optimizing the coupling of microwave and adjusting down the working temperature of interrogation region, etc. A short-term fractional frequency stability of $2.7 \times 10^{-13} \tau^{-1/2}$ is obtained by comparing the AFC with an H-maser. In accuracy evaluation of AFC, a self-comparison method is introduced, and an evaluation precision of $6 \times 10^{-16}$ is obtained at the average time of 300,000 s. In addition, a GPS common view instrument is set up to compare AFC with frequency standards at other time-keeping laboratories.

I. INTRODUCTION

The atomic fountain clock (AFC) is the most accurate primary frequency standard running in the International Atomic Time (TAI) system, and is utilized to characterize the accuracy of TAI [1] [2]. Records of the type-B uncertainty and the long-term fractional stability of AFC are as low as $1.1 \times 10^{-16}$ [3], and less than $6 \times 10^{-17}$ [4], respectively. Construction of a $^{87}\text{Rb}$ AFC started at Shanghai Institute of Optics and Fine Mechanics (SIOM) in 2005, and since then physical and microwave experiments have been performed, closed-loop operation of AFC has been realized, with its fractional stability and the total uncertainty evaluated of $5 \times 10^{-13} \tau^{-1/2}$ and $2.4 \times 10^{-15}$, respectively [5]. Lately, some improvements have been made on AFC. The LO is directly locked by the fountain, microwave coupling of the Ramsey cavity is optimized, and the working temperature of Ramsey interrogation region is adjusted down. A self-comparison method is introduced to evaluate some systematic biases of AFC. In addition, a GPS common view instrument is set up to compare AFC with frequency standards located at other time-keeping laboratories.

II. EXPERIMENTAL IMPROVEMENTS AND RESULTS

To build an independently-running AFC with standard frequency signal output, we directly lock LO to the resonance spectra of the atomic fountain. The LO locking scheme of our AFC is illustrated in Fig. 1. Acting as the LO of AFC, an OCXO delivers a 5MHz signal to the microwave synthesizer, and the synthesizer generates interrogation signal and feed it into the Ramsey cavity. After the process of interrogating the cold atoms, the frequency-sensitive transition probability is obtained by detecting the time of flight (TOF) signals of the $^{87}\text{Rb}$ in the two hyperfine levels, with frequency error derived [6], and a feedback is given to the OCXO. Thus, in the LO locking scheme, AFC runs independently and outputs a standard frequency signal, which lays the foundation for many high-precision measurement fields where a clock group is unavailable [7].

We make some improvements on the physical system to develop the performance of AFC. The first improvement is to adjust the Ramsey cavity and microwave chain in order to decrease the uncertainty due to power-related shift, which is about $2.0 \times 10^{-15}$. The apparatus figure of the Ramsey cavity is as illustrated in Fig. 2(a). Measured with a network analyzer, the return loss difference between the two ports decreases from the previous result of 5dB in the peak and 2.5dB on the sideband to 0.5dB in the peak and 1.5dB on the sideband, and the quality factor of the Ramsey cavity after the improvement is 9600, as shown in Fig. 2(b). The evaluation of the power-related shift is ongoing, and it is expected that the uncertainty due to the power-shift reduced to less than $1.0 \times 10^{-15}$ [8] [9]. The second improvement is to adjust down the working temperature of the Ramsey interaction region. The temperature was previously $53.2^\circ C$. The higher temperature will bring in a larger blackbody radiation (BBR) shift and uncertainty [10]. To reduce the BBR shift, we rub the end cap surface of Ramsey cavity, and adjust down the resonant temperature of the Ramsey cavity according to a calculated function of resonant temperature in terms of cavity length. Finally, the working temperature is

978-1-4799-8866-2/15/$31.00 ©2015 IEEE 495
adjusted down to 29.0°C, with the temperature gradient of the Ramsey interaction region decreasing correspondingly. The measured temperature stability of the Ramsey cavity after the improvements is illustrated in Fig. 3. Ultimately the evaluation value of the BBR shift is improved from $17.3(2) \times 10^{-15}$ to $13.0(1) \times 10^{-15}$. Other improvements include replacement of the heating wire with non-magnetic carbon fiber, which is helpful in reduction in the residual magnetic field sourced from heating wiring, optimizing the magnetic environment with 60% reduction in residual magnetic field of the Ramsey interrogation region, and improving the vacuum degree of the vacuum chamber from $2 \times 10^{-7}$Pa to $6 \times 10^{-8}$Pa.

In the LO locking scheme, noise of the voltage-control-circuit of the LO will degenerate the stability of AFC. Thus, a new electronic based on a low-noise digital-to-analog convertor (DAC) with a fractional frequency resolution of $1.8 \times 10^{-13}$ is made, whose noise contribution to AFC is negligible. After improvements of the physical system and adoption of the new control electronic, the short-term fractional frequency stability of AFC compared with the H-maser is $2.7 \times 10^{-13} \tau^{-1/2}$, which arrives at $1.8 \times 10^{-15}$ at the average time of 40 000 s, as illustrated in Fig. 4. The long-term stability of the comparison data deviates the fitting line at the average time of longer than 10 000 s, during which the stability index of the H-maser becomes significant.

In the widely adopted accuracy evaluation process, a certain systematic bias is evaluated as following step: 1) to find out the physical parameter $x$ with respect to frequency bias $y(x)$ to measure a series of frequency differences $y(x_1)$ and $y(x_2)$ in $x = x_1$ and $x = x_2$, respectively, 3) to compute the systematic bias and its uncertainty utilizing the function as

$$y_{bias} = \frac{y(x_1) - y(x_2)}{x_1 - x_2}; \quad \sigma_{bias} = \frac{y(x_1) - y(x_2)}{x_1 - x_2} \sigma_x.$$  \hspace{1cm} (1)

Usually frequency differences $y(x_1)$ and $y(x_2)$ is obtained by comparing the AFC with other frequency standards, mainly active H-masers, and the evaluation precision is limited by the long-term drift of the H-maser. We introduce a self-comparison method to evaluate systematic biases of AFC and in this method external reference frequency standards become unnecessary any more. The self-comparison scheme is as shown in Fig. 5(a). In the $nth$ self-comparison period, AFC alternately works in the two states with $x = x_1$ and $x = x_2$, error signal $y_n(x_1)$ and $y_n(x_2)$ are obtained in each state, and AFC is locked by referring $y_n(x_1)$. After a long-time experiment, a series of $y_n(x_1)$ and $y_n(x_2)$ are recorded, and the systematic bias could be evaluated, if they are introduced into function (1). In the self-comparison scheme, either an external frequency reference or a clock group is unnecessary, thus, the evaluation process is simplified, and the drift of the external reference is avoided. In addition, noise due to fluctuation of other systematic bias is partly canceled, and the evaluation precision of comparing experiments is better than the total uncertainty of AFC. However, it prolongs the comparison time necessary for arriving at a certain precision in the evaluation, as the stability degenerates due to Dick effect. We conduct an experiment of verifying the validity of the self-comparison scheme. Provided $x_1 = x_2$, the Allan deviation of fractional frequency difference between a series of $y_n(x_1)$ and $y_n(x_2)$ is as illustrated in Fig. 5(b), with the fitting line of $5.5 \times 10^{-13} \tau^{-1/2}$, and it reaches $6.0 \times 10^{-16}$ at the average time of 300 000 s. The self-comparison method is being utilized to evaluate the collisional shift, DCP shift, light shift, and so on.

To source the frequency of our AFC to TAI, a GPS common view instrument has been set up in our laboratory [11]. The
Fig. 5. (a) The self-comparison scheme, in which $x_1$ and $x_2$ are values of the physical parameter of a specific systematic bias, $y_1(n)$ and $y_2(n)$ are fractional frequency error obtained correspondingly. (b) Allan deviation of difference between the frequency error signals $y_n(x_1)$ and $y_n(x_2)$, with the fitting line $5.5 \times 10^{-13} \tau^{-1/2}$.

preliminary comparison result of AFC with IGST shows that the fractional frequency stability of $3.5 \times 10^{-15}$ at the average time of 120 000 s, limited by noise of the transfer technology.

III. Conclusion

In conclusion, we make several advances in our $^{87}$Rb AFC. Several adjustments of the physical system have been made to decrease the effects of microwave power shift, BBR shift, and so on. We realize LO locking, and a short-term fractional frequency stability of $2.7 \times 10^{-13} \tau^{-1/2}$ is obtained. A self-comparison scheme is introduced to evaluate the accuracy and uncertainty of AFC, and an Allan deviation of $6 \times 10^{-16}$ is achieved at the average time of 300 000 s. In addition, a GPS common view instrument has been set up to compare the standard frequency output of the AFC with remote frequency standards located at other laboratories.

ACKNOWLEDGMENT

We thank Wei Guan and Haibo Yuan in National Time Service Center (NTSC) for their assistant in data processing of remote time comparison results based on transfer technologies GPS common view and GPS PPP. This work is supported by National Natural Science Foundation of China (Grant Nos. 61275204 and 91336105).

REFERENCES

Comparison of Frequency Estimators for Interrogation of Wireless Resonant SAW Sensors

Victor Kalinin
Transense Technologies plc
Bicester, Oxfordshire OX25 3SX, UK
E-mail: victor.kalinin@transense.co.uk

Abstract— Statistical simulation is used to evaluate performance of four different frequency estimators for interrogation of resonant wireless SAW sensors. The first one is based on DFT and quadratic interpolation, the second one employs a weighted least-squares estimate of the phase difference between signal samples. The third and the fourth methods use singular value decomposition and apply a weighted linear predictor in the case of constant sine wave amplitude and an iterative least squares method in the case of decaying sine wave. Numerical receiver model includes additive and phase noises, SAW response limiting, non-linear phase distortions and parasitic SAW resonances. Experimental results are also obtained for all the four frequency estimators.

Keywords—wireless resonant sensor; SAW resonator; frequency estimator; wireless reader; systematic error; random error

I. INTRODUCTION

Wireless resonant SAW sensors are often interrogated by means of a pulsed RF excitation of the SAW resonators and a measurement of the frequency of their free oscillations picked up by a receiver of the reader [1-3]. Estimation of the frequency of the SAW response is usually performed by means of DFT and quadratic interpolation between the frequency points to find the maximum of the energy spectrum. This method is known to be very close to an unbiased maximum likelihood estimator (MLE) for a complex sine wave in the presence of additive white Gaussian noise (AWGN) and it attains Cramer-Rao lower bound (CRLB) within a wide range of SNR values [4]. However, there are a number of other frequency estimators that have been recently developed by a signal processing community for the complex sine wave in the presence of AWGN. Some of them, for instance, a phase-based frequency estimator (PBFE) employing weighted least-squares estimate of the phase difference between neighboring filtered signal samples [5] might be less computationally complex than MLE and thus characterized by lower latency. Another modern model-based parametric method uses singular value decomposition (SVD) to separate signal and noise subspaces and then applies a generalized weighted linear predictor (GWLP) to the left and right principal singular vectors to find the phase difference between their components [6]. It was developed for a constant amplitude sine wave but there is a version applicable to a damped complex sine wave [7], which could be better suited to the SAW resonator response. It also employs SVD and finds the phase difference between the components of the left and right principal singular vectors by means of an iterative weighted least squares (WLS) method. It has been recently shown in [8] that a similar SVD based method applied to a real signal can, under certain conditions, noticeably reduce amount of random frequency errors caused by AWGN compared to the DFT based method.

However, the problem is that wireless SAW readers do not just add a Gaussian noise to the sensor response. They also introduce a phase noise as a result of the frequency down-conversion, limit the SAW response amplitude in the IF channel and distort its phase. Besides, the SAW response may or may not contain a parasitic mode responses in addition to the main one so the analysis of [5-8] taking only AWGN into account and relying on CRLB as a benchmark is not sufficient. A more realistic model of the reader taking into account the phase noise of the local oscillator and the receiver non-linearity was applied to the analysis of the errors of the DFT based estimator in [9]. The aim of this paper is to use a similar model to investigate theoretically performance of the three parametric estimators mentioned above and compare them with the DFT based estimator. Results of this analysis are presented in sections II-III. Section IV contains experimental results obtained for each of the four estimators and some recommendations on their selection.

II. DESCRIPTION OF THE FREQUENCY ESTIMATORS

A. Model of the Reader and the SAW Sensor Response

The wireless reader used in research is a compact reader built on the RF ASIC transceiver chip MLX11006 and TI TMS320F2808 DSP chip [9, 10]. Its transmitter generates 10 dBm RF pulses lasting 2-10 us within the frequency range from 420 MHz to 440 MHz. A double super heterodyne receiver has the sensitivity of at least -83 dBm for the bandwidth of up to 3.5 MHz and the SNR = 17 dB. It has a limiting amplifier in the 1st IF channel to improve the reader’s dynamic range. The receiver has quadrature outputs producing signals I(t) and Q(t) at the nominal 2nd IF 1 MHz (when the interrogation frequency coincides with the SAW resonant frequency).

The model of the non-linear characteristics of the receiver is shown in Fig. 2 in [9]. Soft limiting starts at the input power of -40 dBm and reaches saturation at -28 dBm. Non-linear phase distortions start at -50 dBm and reach the peak value of...
2° at about -10 dBm. The phase noise spectrum of the Rx local oscillator is shown in Fig. 4 in [9]. Its floor is -120 dBc/Hz and a local maximum of -74 dBc/Hz is reached at 90 kHz offset.

The SAW response at the receiver output is modeled as

\[ R(t) = F'(A_{out}(A_w \exp(-t/\tau)) \cos(\Theta(t)) + n_t(i)), \]

\[ Q(t) = F'(A_{out}(A_w \exp(-t/\tau)) \sin(\Theta(t)) + n_t(i)) \]

where \( \Theta(t) = \omega t + \Phi_{out}(A_w \exp(-t/\tau)) + \Psi(t) \), \( \omega \) is the frequency of SAW natural oscillations at the 2nd IF output, \( \tau \) is the time constant of the loaded resonator, \( A_{out} \) and \( \Phi_{out} \) are output amplitude and phase as functions of the input amplitude, \( A_w \) is the maximum amplitude of the SAW response at the Rx input, \( \Psi(t) \) is the random phase corresponding to the phase noise spectrum, \( n_t(i) \) is AWGN and \( F' \) means filtering in the output bandwidth filters. Depending on the input peak power of the SAW response, its envelope at the Rx output can have a shape consisting of a constant and an exponentially decaying parts as shown in Fig. 3 in [9].

Signals \( I(t) \) and \( Q(t) \) are sampled in the ADCs built into the DSP chip with their maximum possible rate of \( F_t = 4.166 \) Msps. A 16.8 us sampling window length is selected to give minimum random frequency errors caused by AWGN for the loaded Q factor of the SAW resonator around 5000-7000 in the absence of limiting.

B. DFT Based Frequency Estimator

This estimator is applicable to any envelope of the SAW response. It calculates the power spectral density \( S(\omega) \) of the complex SAW response \( s(t) = I(t) + jQ(t) \) in order to eliminate a bias from the frequency estimate:

\[ S_n = \left[ \sum_{k=0}^{N-1} (I_k \cos(\phi_{kn}) + Q_k \sin(\phi_{kn})) \right]^2 + \left[ \sum_{k=0}^{N-1} (Q_k \cos(\phi_{kn}) - I_k \sin(\phi_{kn})) \right]^2, \quad \phi_{kn} = 2\pi kn/N \]

(2)

where \( S_n \) is the \( n \)th line of the energy spectrum and \( I_k \) and \( Q_k \) are the \( k \)th samples of the I and Q signals, \( N = 512 \) giving the line spacing \( \Delta f = 8.138 \) kHz. The search of the maximum of \( S(\omega) \) is done in two steps. First, a certain number of spectral lines of \( I(t) \), six for instance, are calculated with a coarse step \( 6\Delta f \) in order to provide sufficient tracking bandwidth of the reader. Second, five spectral lines of \( s(t) \) are calculated with the fine step \( \Delta f' \) around the maximum coarse line. The resonant frequency estimate is found as a position of the maximum of \( S(\omega) \) by means of parabolic interpolation.

The algorithm requires \((64N_s + 28) \) flops \( (N_s = 70) \) is the number of samples in the SAW response).

C. Phase-based Frequency Estimator

This is the estimator [5] developed for the constant amplitude complex sine wave. It may still give acceptable results for a slowly varying SAW response envelope. First, PBFE uses a digital heterodyning and low-pass filtering of the input samples to move their frequency \( 0.25 < \omega / (\pi F_s) < 0.75 \) to the baseband range. Then the resulting \( N_s \) samples are decimated giving real \( a_{\text{re}} \) and imaginary \( a_{\text{im}} \) parts

\[ a_{\text{re}} = (-1)^n(I_{2n} - Q_{2n-1}), \quad a_{\text{im}} = (-1)^n(I_{2n+1} + Q_{2n}), \quad n = 1 \ldots N_s/2. \]

The next stage is to calculate phase angles between the neighboring complex samples

\[ \phi_n = \angle(b_{\text{im}} + j b_{\text{re}}), \]

\[ b_{\text{re}} = a_{(n+1)\text{re}} + a_{(n+1)\text{im}} m_n, \quad b_{\text{im}} = a_{(n+1)\text{im}} m_n - a_{(n+1)\text{im}} m_{n+1}. \]

(4)

The weighted least squares frequency estimate is found as

\[ \omega / F_s = \sum_{n=1}^{N_s-1} \theta_n \frac{\sin(\frac{\pi}{2} T_s)}{\sin(\frac{\pi}{2} T_s)}, \quad \theta_n = \frac{\pi}{2} \]

(5)

The number of operations required by PBFE is small, \((11N_s - 6) \) flops assuming the use of an arctangent lookup table with linear interpolation. This is a clear advantage of the estimator.

D. Estimator based on SVD + GWLP

The signal in this estimator is also modelled by the constant amplitude complex sine wave [6]. The frequency estimate is based on the phase difference between neighboring elements of the principal singular vectors of a relatively compact signal matrix \( S \):

\[ \begin{bmatrix} S_1 & S_{M+1} & \cdots & S_{(K-1)M+1} \\ \vdots & \vdots & \ddots & \vdots \\ S_M & S_{2M} & \cdots & S_{KM} \end{bmatrix} = [U_s, U_n] \begin{bmatrix} \Lambda_n \end{bmatrix} [V_s, V_n]^T. \]

(6)

Here, \( N_s = KM \) with \( K = M \), and the SVD elements for \( S \) are as follows: \( \Lambda_n \) is the principal singular value corresponding to the signal subspace, \( U_s \) is the diagonal matrix of singular values corresponding to the noise subspace, \( U_n \) and \( V_s \) are the principal left and right singular column vectors with the length \( M \) and \( K \) respectively, \( U_n \) and \( V_s \) are the matrices containing singular left and right column vectors corresponding to the noise subspace, and \( U_n \) stands for conjugate transpose.

A rough estimate of the SAW response frequency is found using GWLP approach:

\[ \omega_{\text{re}} / F_s = \angle(x_1^T W x_2) \]

(7)

where \( x_1 = [U_1 \ldots U_{(M+1)}]^T, x_2 = [U_{M+2} \ldots U_M]^T \), and the elements of the weighting matrix are:

\[ W_{\text{val}} = [M \min(m, k) - m^k] \exp([m-k]\omega / F_s) / M. \]

(8)

Since \( \omega / F_s \) is unknown, the value of \( \omega_{\text{re}} / F_s \) is found from (7) and (8) by means of three iterations.

A more accurate estimate of the frequency is obtained as the value closest to \( \omega_{\text{re}} / F_s \) from a set of values

\[ \omega_{\text{re}} / F_s = \angle(y_1^T W y_2) + 2 \text{mi} / M, \quad \text{mi} = \left[ -\frac{M}{2} \ldots \frac{M}{2} \right]. \]

(9)

where \( y_1 = [V_1 \ldots V_{M+1}]^T, y_2 = [V_{M+2} \ldots V_M]^T \).

The number of operations required to find the frequency from (7)-(9) and known vectors \( U_i \) and \( V_i \) is approximately \((24N_s + 72 N_s^{1/2} - 5) \). It is more difficult to determine the number of operations required for SVD. Complexity of Golub-Reinsch algorithm for a real matrix is \( O(21 N_s^{3/2}) \) [11]. SVD
of a complex matrix takes many more operations. Finding only principal singular vectors may require less operations but most likely the SVD+GWLP estimator is still more computationally involved than the DFT-based estimator.

E. Estimator based on SVD + WLS

The SAW response at the Rx output does not have a constant envelope unless it is strongly limited. The SVD-based method described in [7] can be used for estimation of frequencies and to damping factors α of multiple decaying complex sine waves. If the reader excites just a single resonator, it makes sense to model the SAW response by a single damped sine wave and regard parasitic modes that may be present in the signal as noise. In this case, the SVD+WLS method [7] uses the same decomposition of the signal matrix (6). The difference is in the way how the optimal (M-1)-(M-1) weighting matrix is found from the tridiagonal matrix $A$:

$$W_1 = A^{-1}, \quad A_{m,m-1} = 1 + \alpha^2, A_{m,(m+1)} = A_{(m+1),m} = -\alpha e^{-j\omega T/\Delta}.$$  \hspace{1cm} (10)

The WLS procedure gives the following rough estimate of the frequency based on the left singular vector,

$$\omega_L/F_S = \angle (g), \quad g = (x_d^H W_1 x_1)^{-1} (x_d^H W_1 x_2),$$  \hspace{1cm} (11)

that can be calculated again by means of three iterations.

A more accurate estimate is found as a the value closest to $\omega_0$ from a set of values

$$\omega_{R,i}/F_S = \left[ \angle (h_i) + 2\pi i \right] / M, \quad i = \left[ \frac{M}{2} \right], \ldots, \left[ \frac{M}{2} \right]$$  \hspace{1cm} (12)

where $h_i$ is obtained as a result of an iterative procedure

$$h_i = [(T_1^H h)^H W_2 T_1 h]^{-1} (T_1^H h)^H W_2 T_2 h,$$  \hspace{1cm} (13)

$$W_2 = (BB^H)^{-1},$$  \hspace{1cm} (14)

where $B$ is the bidiagonal $(K-1)\times(K-1)$ matrix with $B_{k,k} = -h_1$ and $B_{k,(k-1)} = 1$. In (13), we have

$$T_1 = [I_{K-1} \ 0_{(K-1)\times1}], \quad T_2 = [0_{(K-1)\times1} \ I_{K-1}],$$

and the column vector

$$h = S^T (G^T)^T, \quad G = [g \ g^2 \ \ldots \ g^M]^T$$  \hspace{1cm} (15)

where $^T$ stands for pseudo-inverse.

The complexity of this estimator is at least not lower than that of the previous one.

III. STATISTICAL SIMULATION OF THE ESTIMATORS

To compare the performance of the four frequency estimators, they are implemented as a Matlab code. The input signal for the estimators $s(t) = l(t) + jQ(t)$ with a known frequency is also generated by a Matlab code simulating the reader according to (1). The mean error $\Delta f$ of the measured frequency and its standard deviation $\sigma_f$ are obtained for each estimator on the basis of 1000 trials. Influence of different reader and SAW response properties is investigated separately in a number of numerical experiments assuming $Q = 5000$ at 430 MHz.

A. Additive Noise and Receiver Non-linearity

First, the phase noise and parasitic modes are excluded from the model. The graphs of $\Delta f$ and $\sigma_f$ plotted against the maximum power of the SAW response at the Rx input $P_{in}$ are shown in Fig. 1 for $f = \omega 2\pi/1 MHz$. As one can see, the SVD-based estimators ($M = 10, K = 7$) do give approximately 1.35 times lower random errors than the DFT-based estimator for $P_{in} < -20$ dBm. This improvement is smaller than the one reported in [8] (the factor of 2.43 in the experiment). One explanation is that a much larger number of samples ($N_s = 504$) was used in [8], and the sampling window of 56 us exceeded the optimum signal length for the DFT-based estimator. It is also clear that the PBFE gives the largest random errors at all input powers and especially large errors at $P_{in} < -60$ dBm due to a high SNR threshold inherent to this method.

Fig. 1b shows that the signal phase distortions taking place in the limiting region at $P_{in} > -30$ dBm cause significant systematic errors especially large for the SVD-based estimators. They are surprisingly small for PBFE, probably, due to its averaging property. On the other hand, the systematic errors of this estimator grow very rapidly as soon as the SNR goes below its threshold.

B. Phase Noise

Adding the phase noise to the reader model changes the picture quite noticeably. As one can see in Fig. 2a, the random errors do not keep decreasing with the increase of SNR (or input power). This effect was shown in [9] for the DFT estimator and a similar behavior is inherent to other estimators as well. The simulations show that the influence of the phase noise (as it is defined in the model) on the SVD-based estimators is almost two times stronger than on the DFT-based

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{Standard deviation (a) and the mean value (b) of the measured frequency against the input power of the SAW response in the absence of phase noise.}
\end{figure}
estimator. The phase noise dominated region for SVD starts at $P_{in} > -65$ dBm while for DFT it starts at -55 dBm. Interestingly, the PBFE is the list sensitive estimator to the phase noise. PBFE random errors are smaller than those for other estimators at $P_{in} > -50$ dBm so this estimator seems to be quite attractive for the applications with a strong SAW response although its performance for weak responses (in the AWGN dominated region) is much worse.

C. Influence of the SAW response frequency

So far, the simulations have been performed only at $\omega / \pi F_s = 0.48$ and it is a known fact [5] that the performance of an estimator may depend on the frequency. Simulations also performed at $0.19 < \omega / \pi F_s < 0.77$ show (see Fig. 3) that the mean error oscillates within quite a narrow range of ±80 Hz for all estimators in the phase noise dominated region ($P_{in} = -40$ dBm) and only for PBFE it changed dramatically from -110 kHz to 32 kHz in the additive noise dominated region ($P_{in} = -70$ dBm) because the SNR for this estimator is below the threshold.

Variation of $\sigma_f$ with the frequency is quite small in the phase noise dominated region (see Fig. 4) and a little more noticeable in the additive noise dominated region, especially for SVD+WLS estimator. From a practical point of view, it should not cause problems for the SAW reader.

![Fig. 2. Standard deviation (a) and the mean value (b) of the measured frequency against the input power of the SAW response in the presence of phase noise.](image)

![Fig. 3. Variation of the mean error $\Delta f$ with the measured frequency at $P_{in} = -40$ dBm and $P_{in} = -70$ dBm.](image)

![Fig. 4. Variation of the standard deviation $\sigma_f$ with the measured frequency at $P_{in} = -40$ dBm and $P_{in} = -70$ dBm.](image)
D. Influence of a parasitic mode

Parasitic modes in the SAW response may originate from (a) the main resonant peaks of other resonators present in the sensor that are slightly excited by the interrogation pulse, (b) parasitic longitudinal modes of the other resonators getting close to the main resonant peak and (c) parasitic transversal modes. The most dangerous case is (b) when the detuning between the parasitic mode and the main response is around 70 kHz. This causes the largest pulling effect on the maximum of $S(\omega)$.

Simulations show that it is mainly the mean error $\Delta f$ that is affected by the interference in the form of the parasitic resonant response. Moreover, $\Delta f$ depends on the phase of the interference relative to the main response. Fig. 5 shows the variation of the maximum mean error with the signal-to-interference ratio (SIR) for all four estimators at $P_{in} = -40$ dBm. The SVD-based methods are the most sensitive to the parasitic response. DFT-based estimator is twice less sensitive compared to them and the PBFE is twice less sensitive than the DFT-based estimator.

IV. EXPERIMENT AND DISCUSSION

Experimental testing of the four frequency estimators is performed by means of generation of SAW responses by the actual SAW reader [10] connected to a SAW resonator working at 435 MHz and having $Q \approx 10000$ through a directional planar microstrip coupler. Variation of the input power of the SAW response is achieved by changing the gap between the two microstrips coupled back to back from 1.6 mm to 16.1 mm, which causes the change of $P_{in}$ approximately by 40 dB, roughly from -35 dBm to -75 dBm. Each SAW response sampled by the ADC is logged on a PC disk. Instead of processing the sampled response in the DSP of the reader, 1000 responses are saved and processed in the PC off line using the same Matlab routines as those used for theoretical simulations of the frequency estimators.

Initially, the interrogation frequency is selected to be close to the SAW resonant frequency so the sampled response has the frequency approximately equal to 1 MHz. Variation of the standard deviation and the mean value of the measured frequency against the gap width is presented in Fig. 6 for this case. The character of the curves is similar to the one of the theoretical curves shown in Fig. 2. The phase noise dominates at the gap width below 6 mm and the additive noise dominates at the gaps wider than 9 mm. However, the advantage of PBFE for strong SAW responses is not as evident as it is in theory: $\sigma_f$ is just 2-2.5% lower than that for the DFT-based estimator. As expected, $\sigma_f$ rises with the gap width faster for PBFE than for any other estimator when the SAW response is weak.

The SVD+GWLP estimator performance is significantly worse for the real signal over a wide range of its magnitudes than that for other estimators. On the other hand, SVD+WLS estimator shows the standard deviation only 1-3% higher than that for the DFT-based estimator when the SAW response is strong. This difference is smaller than the one seen in Fig. 2. For the weak SAW response, this estimator also does not show any advantage over the DFT-based estimator, probably, because the SAW response magnitude does not reach sufficiently low value.
The fact that the difference in the values of distortions of the strong SAW response near 1 MHz IF. variation of 0.6-2 kHz because of limiting and phase properties of the estimators; it happens because the SAW response magnitude drops by 20 dB at 150 kHz detuning due to a limited spectral width of the 10 us interrogation pulse. The mean value of the measured frequency experiences a peak of 0.6-2 kHz because of limiting and phase distortions of the strong SAW response near 1 MHz IF.

The fact that the difference in the values of $\sigma$ for PBFE and SVD+WLS estimator in the phase noise dominated region is not as big as predicted by simulations indicates that the phase noise model used in the simulations does not describe the real phase noise accurately enough. Nevertheless it does give a correct qualitative picture. Both the simulations and the experiment show that selection of the best frequency estimator for SAW sensor interrogation depends on the application. If it mostly deals with weak SAW responses, allows slow interrogation and processing of large number of samples in the reader then the SVD+WLS method can provide smaller random errors or a longer read range. In the case of very strong SAW responses that need to be processed as quickly as possible, PBFE can be an attractive option although its implementation using fixed point arithmetic should be done with care to ensure small errors in calculation of the arctangent function. If the SAW response magnitude can change within a wide range of values and processing should be relatively fast then the best option can be the DFT-based estimator. In any case, one should take into consideration particular non-linear and noise properties of the reader’s receiver as well as a presence of parasitic SAW responses in the input signal.

V. CONCLUSIONS

Four different frequency estimators, one DFT-based, one phase based, and two SVD-based, were implemented as a Matlab code for measurement of the frequency of natural oscillations of a passive wireless SAW resonant sensor. Their performance was compared using a mathematical model of a SAW reader performing time domain pulsed interrogation of the sensor as a source of test signals. The model included non-linearity of the reader’s receiver, its additive noise and the phase noise of its local oscillator as well as a parasitic SAW response. Random and systematic errors for all four estimators were calculated using statistical simulation. They were also found experimentally by using a real reader and a SAW sensor. Recommendation on selection of the frequency estimator were proposed based on the theoretical and experimental results.

REFERENCES


Acoustic Power Gain Induced by 2D Electron Drifting

Lei Shao *
Department of Mechanical Engineering
University of Michigan
Ann Arbor, MI, USA
shaolei@umich.edu

Kevin P. Pipe
Department of Mechanical Engineering
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI, USA
pipe@umich.edu

Abstract—In this work, amplification of surface acoustic waves (SAWs) by electron drift in a nanometer-scale two-dimensional electron gas (2DEG) is analyzed analytically. We compare the amount of acoustic power gain per SAW radian produced by electron drift in a bulk GaN thin film layer and in a GaN-based 2DEG layer. Calculations suggest that acoustic amplification in a 2DEG is independent on the SAW frequency while only a very narrow bandwidth of SAWs could be amplified in bulk. Furthermore, the peak power gain per SAW radian occurs at a more practical carrier density for a 2DEG than for a bulk material.

Keywords—acoustoelectric coupling; two-dimensional electron gas; surface acoustic wave; III-V nitride semiconductors

I. INTRODUCTION

Interactions between acoustic waves and electrons have been extensively studied for more than half a century. In 1953, the acoustoelectric effect was first theoretically described by Parmenter, who predicted that a travelling longitudinal acoustic wave could drive electrons to generate an electric current [1]. This effect was then thoroughly studied by many investigations in subsequent years [2-4]. In 1961, Hutson, McFee, and White demonstrated that bulk ultrasonic waves could be substantially amplified by drifting electrons if the electron velocity exceeds the speed of ultrasonic waves in CdS which is a piezoelectric semiconductor [5]. Later studies examined coupling between surface acoustic waves (SAWs) and electrons in a bulk or thin-film semiconductor [6-8]. The experimental approaches for such studies involved depositing a semiconducting layer with high electron mobility on a piezoelectric substrate, or simply bringing a semiconducting slab close enough to a piezoelectric substrate that the air gap is less than one wavelength of the SAWs propagating in the piezoelectric substrate [7, 8]. The oscillating electric field excited by SAWs travelling in a piezoelectric substrate interacts with drifting electrons in the semiconductor layer, which could amplify (or attenuate) the SAWs if the electron drift velocity is higher (or lower) than the SAW velocity. This effect was extensively studied during the 1970s and 1980s with demonstrations of many varieties of acoustoelectric devices, such as convolvers, correlators, etc. [9-11]. During the early 1990s, attention shifted to coupling between SAWs and two-dimensional electron gases (2DEGs), in particular the attenuation of SAWs by a 2DEG [12, 13], the use of SAWs to probe quantum effects in 2DEGs [14], the generation of SAWs in a dynamically screened 2DEG [15, 16], and the transport of multiple [17, 18] or single [19, 20] electrons by SAWs.

However, very few studies are available in the literature about the amplification of acoustic waves in a 2DEG; there is a single investigation of a 2D-space-charge-wave-based analytic model [21] which does not revert to the well-established acoustic attenuation model by a 2DEG [13] when the electron drift velocity is set to zero. In this work, we analytically study the amplification of SAWs by electron drift in a 2DEG, and compare the amount of SAW power gain induced by a 2DEG with that induced by a bulk material, in terms of SAW frequency and material constants such as carrier densities and carrier mobility.

II. ANALYTIC FORMULATION

A. SAW – electron interaction in bulk materials

To model the acoustoelectric amplification of SAWs by electron drift in a 2DEG, we first revisit the amplification of SAWs in a bulk piezoelectric semiconductor [5, 7]. It has been well-understood that as a SAW propagates through a piezoelectric semiconductor, it induces an electric field that causes bunching of carriers and hence a periodic variation in electrical conductivity. In the presence of a DC current, this periodic conductivity leads to an additional periodic potential that causes bunching of carriers and hence a periodic variation in electrical conductivity. The oscillating electric field excited by SAWs travelling in a piezoelectric substrate interacts with drifting electrons in the semiconductor layer, which could amplify (or attenuate) the SAWs if the electron drift velocity is higher (or lower) than the SAW velocity. This effect was extensively studied during the 1970s and 1980s with demonstrations of many varieties of acoustoelectric devices, such as convolvers, correlators, etc. [9-11].

The equations that describe piezoelectric coupling in a material can be written as:

\[ T = eS - \varepsilon_\epsilon E, \]  
\[ D = \varepsilon_\epsilon S + \varepsilon E, \]
where \( c \) is the elastic constant at constant electric field, \( \varepsilon \) is the dielectric permittivity at constant strain, \( E \) is the electric field, \( T \) is the stress, \( e_p \) is the piezoelectric constant, and \( D \) is the electric displacement. To derive the ultrasonic amplification constant [22], one first obtains an expression for the electric displacement, \( D \), in terms of the \( E \) field from Poisson’s equation and the current-density-space-charge continuity equation. This relation between \( D \) and \( E \) is then used to eliminate \( D \) from the Eq. (2), so that an expression for \( E \) in terms of the strain, \( S \), is obtained, which is then substituted in the Eq. (1). The result is a complex elastic stiffness constant, \( c' \), which contains all of the electrical effects. The acoustic amplification constant, \( \alpha \), is then obtained from the imaginary part of \( (c')^{1/2} \), yielding [5, 7]:

\[
\alpha = k \frac{K^2}{2} \frac{\omega / (\omega \gamma)}{1 + (\omega / (\omega \gamma))^2 (1 + k^2 L^2)}/
\]

where \( k \) is the acoustic wavenumber, \( K^2 \) is the effective electromechanical coupling coefficient which accounts for the surface boundary condition and is slightly different than the coefficient \( K^2 = e^2 / \varepsilon \) used for bulk waves [13], \( \omega \) is the acoustic frequency, \( \gamma = v_d / v_s - 1 \) is the dimensionless electron drift parameter (for which \( v_d \) is the electron drift velocity and \( v_s \) is the acoustic velocity), \( \omega = \sigma / \varepsilon \) is the conductivity, \( L^2 \) is the 3D Debye screening length (for which \( k_b \) is the Boltzmann constant, \( T \) is the absolute temperature, \( e_0 \) is the electron charge, and \( N_d \) is the bulk carrier density). \( L^2 \) can also be written as \( D / \omega \) [22], where \( D = \mu k_b T / e_0 \) is the diffusion constant and \( \mu \) is the electron mobility.

**B. SAW – electron interaction in a 2DEG**

In this work, rather than a bulk semiconductor, we consider a nanometer-scale conductive 2DEG layer of thickness \( d << 1/k \) on an insulating piezoelectric substrate, where \( k \) is the SAW wavevector and \( 1/k \) is the typical SAW penetration depth into the substrate. To adapt Eq. (3) to describe the SAW-2DEG interaction in such a structure, the following two modifications are required:

1. For this geometry, the bulk conductivity used in \( \omega \) needs to be replaced by \( \sigma_s \), [13, 23], where \( \sigma_s = \sigma_d \) is the 2DEG sheet conductivity;

2. An effective dielectric constant \( \varepsilon' = \varepsilon + e_0 \) need to be used which accounts for the half-space with dielectric constant \( e_0 \) above the conductive layer [23].

Substituting the above into Eq. (3) yields

\[
\alpha = k \frac{K^2}{2} \frac{\sigma_s / (\sigma_u \gamma)}{1 + (\sigma_s / (\sigma_u \gamma))^2 (1 + k \Lambda)^2}/
\]

where \( \sigma_u = v_d \varepsilon' \), and \( \Lambda = k_b T / e_0 n_s \) is the 2DEG screening length (in which \( n_s \) is the 2DEG sheet carrier density). This is the governing equation for the interaction of SAWs with drifting electrons in a 2DEG. We note here that this equation is only valid in the case that the surrounding materials of the 2DEG are highly resistive, contributing much lower currents than the 2DEG. Under conditions of no electron drift \( (\gamma = -1) \) and a SAW wavelength much greater than the 2D electron screening length \( (kA \) small), Eq. (4) reverts to the well-known equation describing attenuation of SAWs by a 2DEG without magnetic fields [13]:

\[
\alpha = -k \frac{K^2}{2} \frac{(\sigma_s / \sigma_u)}{1 + (\sigma_s / \sigma_u)^2}.
\]

Fig. 1 shows a layer structure supporting this interaction. When a SAW propagates through a 2DEG, the intensity of the SAW is increased (reduced) if the electron drift velocity is larger (smaller) than the SAW velocity. Based on Eq. (4), Fig. 2 shows \( \alpha \) as a function of electron drift velocity in a GaN-based 2DEG for four different values of carrier density. Here we choose GaN-based materials for demonstration purposes due to their large acoustic velocities, high electron mobility, and strong piezoelectric coupling [24]. The calculated interaction strength assumes an AlGaN/GaN 2DEG heterostructure grown on a sapphire substrate with an electron mobility of 1500 cm²/Vs at 300 K. We approximate \( K^2 \approx K^2 \) and use values for the epitaxial structure of \( e_p = 1.14 \) C/m², \( c = 397 \) GPa [25], and \( \varepsilon = 5.16 e_0 \) [26] at high frequencies. The
assumed frequency of the SAW used for demonstration here is 0.8 GHz and its velocity is 4.6×10^5 cm/s, consistent with the most typical 2-μm-GaN-on-sapphire structure [27]. We note that because the current flow primarily occurs in the 2DEG rather than in the GaN thin film layer underneath it, the interaction between SAWs and the GaN layer can be neglected. Positive gain indicates that the SAW is amplified, while negative gain indicates attenuation. A positive electron velocity indicates that the SAW is travelling in the same direction as the drifting electrons, while negative indicates the opposite direction. The SAW-2DEG interaction is nonreciprocal about zero drift velocity because zero gain occurs when the electron velocity equals the SAW velocity. Fig. 3 shows α as a function of 2DEG density for four different values of electron velocity, based on the material constants of the same layer structure used in Fig. 2. The gain decreases after reaching its maximum as carrier density keeps increasing, because the travelling piezoelectric field accompanying the SAW has a reduced bunching effect on local carrier density when this carrier density is high [22].

III. COMPARISON BETWEEN BULK AND 2D ELECTRON DRIFT

In order to gain a better insight into the properties of acoustoelectric amplification of SAWs induced by a 2DEG versus bulk, we then study the acoustic power gain per radian (α/k). The solid lines in Fig. 4 show the SAW power gain per radian as it travels through a DC-biased 2DEG, for the four different values of 2DEG density used in Fig. 2, while the electron drift velocity is held constant at 3×10^6 cm/s. As the SAW wavelength decreases (SAW frequency increases) to the sub-micron range and becomes comparable to or less than the 2D electron screening length Λ, the electrons cannot move rapidly enough to respond to the oscillating field generated by the travelling SAWs, and the gain per radian drops quickly. However, for typical SAW wavelength on the order of several micrometers, kΛ is often small. Thus, the SAW power gain per radian becomes independent with SAW frequency (as shown by the dashed lines) and can be derived from Eq. (4) as

\[ \alpha_{2DEG}/k = \frac{K_{SM}^2}{2} \left( 1 - \frac{1}{\frac{1}{\sigma_s/(\sigma_{SM}\gamma')} + \sigma_s/(\sigma_{SM}\gamma')}^2 \right). \]  

(6)

We note here that only a very narrow bandwidth of SAWs could be substantially amplified in a bulk or thin film semiconductor. This is easily understood by written Eq. (3) in the similar form as Eq. (6) while assuming \( k \cdot L_0 = 0 \) for a typical SAW wavelength, yielding

\[ \alpha_{bulk}/k = \frac{K_{SM}^2}{2} \left( 1 - \frac{1}{\frac{1}{\omega_s/(\omega_{SM}\gamma')} + \omega_s/(\omega_{SM}\gamma')}^2 \right). \]  

(7)

Clearly, peak SAW power amplification occurs near the frequency of \( \omega_s/(\omega_{SM}\gamma') \). High-frequency SAWs are not amplified in the bulk case due to reduced conductivity modulation, while for low-frequency SAWs, carriers are able to quickly redistribute to screen the SAW-induced piezoelectric field.

The distinctive difference for amplification of SAWs in a 2DEG versus bulk is that the physical overlap between the
SAW penetration depth (which extends a distance of approximately 1/k into the piezoelectric substrate) and the 2DEG conducting layer is very small as shown in Fig. 5. In a 2DEG-on-piezoelectric-insulator structure, the field modulation associated with the incident SAW relaxes by producing displacement currents within the conductive 2DEG channel [23]. For higher-frequency SAWs, the reduced penetration (due to smaller 1/k) leads to a larger overlap-to-penetration ratio and consequently an increase in displacement currents and an increase in \( \omega_0 \). This compensation effect only occurs in the SAW-2DEG interaction among all types of structures used for acoustoelectric amplification, contributing to frequency-independent gain per radian (or gain per wavelength) for a 2DEG. Therefore, it could enable a much wider bandwidth for acoustic amplification in the 2DEG case versus the bulk.

Finally, we study the optimal carrier densities for acoustic amplification. According to Eq. (6), the acoustic power gain per radian is maximized \( (k^2/4 \sigma_s/|\sigma_u|) = 1 \) when \( \sigma_s/(\sigma_u) = 1 \) and therefore a sheet carrier density

\[
\sigma_s = \frac{\nu_s e^* \gamma}{\epsilon_0 \mu}.
\]

(8)

However, based on Eq. (7), for a bulk material the peak power gain per radian occurs at \( \omega/(\omega_0) = 1 \) and therefore a carrier density

\[
N_s = \frac{\omega^* \gamma}{\epsilon_0 \mu}.
\]

(9)

To understand the relationship between Eqs. (8) and (9), we rewrite the 2D sheet carrier density as \( n_s = d \cdot N_s \) in Eq. (8) in which \( d \) is the thickness of the conducting layer, yielding \( N_s = \frac{\nu_s^* e^* \gamma}{d} \). The conducting layer could be considered as bulk when its thickness \( (d) \) is at least or comparable to the SAW penetration depth \( (1/k) \); substituting \( d = 1/k \) into the previous equation renders the same result as shown in Eq. (9).

Based on Eq. (8), for a 2DEG formed at an AlGaN/GaN interface with an electron velocity of \( 3 \times 10^6 \) cm/s and mobility of 1500 cm²/(V·s), the sheet carrier density at the peak acoustic power gain is approximately \( 1 \times 10^{14} \) cm⁻². On the other hand, for bulk GaN with a SAW frequency of 0.8 GHz, an electron velocity of \( 3 \times 10^6 \) cm/s, and a mobility of 600 cm²/(V·s) [28], the carrier density for peak gain is approximately \( 1 \times 10^{14} \) cm⁻². Since a carrier density as low as \( \sim 10^{14} \) cm⁻² is difficult to obtain in many bulk piezoelectric semiconductors such as GaN due to unintentional dopants, while it is typical to obtain \( \sim 10^{10} \) cm⁻² in a GaN-based non-degenerate 2DEG, a 2DEG could yield a higher maximum gain for practical carrier densities.

IV. Conclusion

In summary, we quantitatively studied the amplification of SAWs in a DC biased 2DEG in which the electron drift velocity exceeds the SAW velocity. Our calculations suggest that for acoustoelectric amplifications of SAWs, 2DEG is a better candidate than bulk or thin film semiconductors because a much wider bandwidth of SAWs can be amplified in a 2DEG at a more practical carrier density. The analytic model presented in this work could support future studies of acoustoelectricity, directional acoustic waves generation [29], and solid-state acoustics in general.

REFERENCES


Figure 5. Illustration of SAW penetration depth and its overlap with the conducting layer. (a) Low-frequency SAW interacts with a bulk piezoelectric semiconductor. (b) Low-frequency SAW interacts with a 2DEG located on top of a piezoelectric insulator. (c) High-frequency SAW interacts with a 2DEG located on top of a piezoelectric insulator.
Modelling and control of a travelling wave in a finite beam, using multi-modal approach and vector control method

Sofiane GHENNA\textsuperscript{1,2}, Frédéric GIRAUD\textsuperscript{1,2}, Christophe GIRAUD-AUDINE\textsuperscript{1,3}, Michel AMBERG\textsuperscript{1,2}, and Betty LEMAIRE-SEMAIL\textsuperscript{1,2}

\textsuperscript{1}Laboratoire d’Electrotechnique et d’Electronique de Puissance IRCICA, 50 avenue Halley, 59650 Villeneuve d’Ascq, France
\textsuperscript{2}Université Lille1, 59650 Villeneuve d’Ascq, France
\textsuperscript{3}Arts et Métiers Paris-Tech Centre de Lille, 8 bd Louis XIV, 59046 Lille

Abstract—This paper presents a new method to produce and control the vibration amplitude and direction of a travelling wave in a finite beam, using multi-modal approach. A closed loop control of the transducer vibration is applied using vector control method. The modelling in rotating frame and the decoupling according to two-axis allows to obtain a double independent closed loop control. This allows to regulate the vibration amplitude of the travelling wave directly. An analytical modelling is presented, with experimental validation, showing good performances even in the presence of perturbations.

Keywords: Traveling wave, Linear motor, Langevin transducer, Multimodal approach, Vector control

I. INTRODUCTION

Several studies are focused on the generation of travelling waves on a beam, to realize linear motor for instance. In [1], a travelling wave on a beam is generated by forces produced by two transducers. These forces are shifted by 90\textdegree{} and generated at the center frequency between the two modes. The advantage is that the impedance matching is no more required and so changing direction can be achieved by changing the phase difference from 90\textdegree{} to −90\textdegree{}. However, these methods are in open loop control and cannot control the vibration amplitude in the presence of disturbance. In [8], the authors proposed a closed-loop control of transducers to produce a travelling wave using the second method, even changing frequency, the travelling wave can be obtained by this control. However the vibration amplitude is low because of the used piezoelectric actuator, this travelling wave is controlled in direction but not in amplitude.

In this work, this latest approach is improved by controlling the vibration of each transducer in a rotating frame. It is then possible to control the amplitude of each actuator and their relative phase shift. This allows the control of both direction and vibration amplitude of the produced travelling wave, and the effect of the beam is rejected as a perturbation. Hence, it is necessary to model the transducers. Experiments have shown that it was possible to control phase and amplitude even in transient and obtained a large travelling wave with standing wave ratio nearly equal to one. The principal applications domain of travelling wave are acoustic levitation [9]–[11], piezoelectric miniature robot [12], [13], Transportation of objects using linear ultrasonic motor [14]–[18]. In the first part, the principle of a new excitation method of vibration modes, using multi-modal approach is presented, with analytical modelling, showing the possibility to excite the beam with vibration amplitude. In the second part, the design of the beam and horns are addressed, then the control method is proposed, allowing to control the travelling wave in both direction and vibration amplitude, using vector control method. Finally experimental results are provided in the last part.

II. MULTI-MODAL APPROACH

A. Forced vibration in finite beam

We consider a thin beam, with rectangular cross section denoted by A, length by L, \( \rho \), \( E \) are respectively density and modulus of elasticity, \( I \) is the quadratic momentum of the beam. The transverse vibration of a uniform elastic homogeneous isotropic Euler-Bernoulli beam [19] can be written in Cartesian coordinates as

\[
EI \frac{\delta^4 w}{\delta x^4} + \rho A \frac{\delta^2 w}{\delta t^2} + r_a \frac{\delta w}{\delta t} = p(x,t) \tag{1}
\]

with: \( w(x,y,z,t) = w(x,t) \) the deflection of the beam at point \( x \) and time \( t \), and \( p(x,t) \) denotes the load per unit length of the beam at point \( x \) and time \( t \), while \( r_a \) represents the coefficient of external damping of the beam.

It is possible to determine the deformation mode shapes, and the frequency spectrum of the beam, by using the analytical
model of Euler-Bernoulli eq.1 developed in [20]. In this paper, harmonic vibrations are only considered enabling the use of the complex notation so that: \( w \) with \( w \) is the real part of \( w \). If we take into account the contribution of every vibration mode excited in the beam, then the vibration can be written by:

\[
\mathbf{w}(x,t) = \sum_{n=1}^{\infty} W_n \phi_n(x)e^{j\omega t}
\]  

(2)

Where \( \mathbf{w} \) includes real and imaginary part of the vibration, at location \( x \), and time \( t \), \( \omega \) is the angular frequency of the flexural wave, and \( j = \sqrt{-1} \). \( \phi_n(x) \) is the deformation mode shapes of the \( n^{th} \) mode, and is fixed for the beam.

\( W_n \) represents the modal amplitude of the \( n^{th} \) mode. We suppose that we can control each mode independently. Hence, the deformation \( \mathbf{w}(x,t) \) depends theoretically on an infinite degrees of freedom which are \( W_n \). This property is used to generate and to control the travelling wave detailed hereafter. we restrict the study to two vibration modes denoted A and B. In this paper, only harmonic vibrations are considered enabling the use of the complex notation with \( \mathbf{X} = \mathbf{X}e^{j\omega t} \), where \( \mathbf{X} \) is the deformation mode shapes, and their modal amplitude. If we consider two given points of the beam \( x_1 \) and \( x_2 \) (position of the actuators), the harmonic vibration at these positions is given by:

\[
\mathbf{W}(x_1) = \mathbf{W}_A \phi_A(x_1) + \mathbf{W}_B \phi_B(x_1)
\]

(4)

\[
\mathbf{W}(x_2) = \mathbf{W}_A \phi_A(x_2) + \mathbf{W}_B \phi_B(x_2)
\]

(5)

where \( \mathbf{W}(x_1) \) and \( \mathbf{W}(x_2) \) are the vibration amplitude of the first and the second actuator, denoted respectively by \( \mathbf{W}_A \) and \( \mathbf{W}_B \). Introducing the matrix notation, eq.4 and eq.5 become:

\[
\begin{pmatrix}
\mathbf{W}_1 \\
\mathbf{W}_2
\end{pmatrix} =
\begin{pmatrix}
\phi_A(x_1) & \phi_B(x_1) \\
\phi_A(x_2) & \phi_B(x_2)
\end{pmatrix}
\begin{pmatrix}
\mathbf{W}_A \\
\mathbf{W}_B
\end{pmatrix}
\]

(6)

giving rise to:

\[
\frac{\mathbf{W}_1}{\mathbf{W}_2} = \phi_{x_1,x_2} \frac{\mathbf{W}_A}{\mathbf{W}_B}
\]

(7)

Inversely if the position of transducers does not represent nodes

\[
\frac{\mathbf{W}_1}{\mathbf{W}_2} = \phi_{x_1,x_2}^{-1} \frac{\mathbf{W}_A}{\mathbf{W}_B}
\]

(8)

The vibrations of transducers impose the dynamic of the modal amplitudes. Then \( \mathbf{W}_A \) and \( \mathbf{W}_B \) can be controlled directly by the vibration amplitude of the transducers. In this paper this principle is used to produce control a travelling wave, using excitation with vibration amplitude.

**C. Condition for travelling wave**

It has been shown in [3] that at fixed frequency a travelling wave can be obtained with the same modal amplitudes \( \mathbf{W}_A = \mathbf{W}_B \) shifted by 90°. In rotating frame this is modelled by:

\[
\mathbf{W}_A = \mathbf{W}_A e^{j\psi_A}
\]

(9)

\[
\mathbf{W}_B = \mathbf{W}_A e^{j\psi_B}
\]

(10)

with:

\[
\psi = \psi_A - \psi_B = \pm \frac{\pi}{2}
\]

(11)

then the necessary condition is

\[
\mathbf{W}_A = \pm j\mathbf{W}_B
\]

(12)
Where $W_A$ and $W_B$ are rotating in the fixed frame, and fixed in the $(d,q)$ frame as described in Fig.3(a). To achieve this, these modal amplitudes must be controlled in order to generate a travelling wave.

Fig. 3. Argument and magnitude in rotating frame

### III. EXPERIMENT VALIDATION

#### A. Design of the beam

The analytical model of Euler-Bernoulli eq.1 was simulated to determine the frequency spectrum of the beam given in Fig.5, the deformed mode shapes presented in Fig 5. A beam made of aluminium was chosen, because of the excellent acoustical characteristics of this material, which parameters are given in Tab.I. The dimensions of the beam has been chosen in such a way that the resonance frequency of transducers 28 kHz is in the range between the two resonance modes.

![Normalized mode shape A and B](image)

**Fig. 5. Simulated normalized mode shape A and B**

**Table I**

<table>
<thead>
<tr>
<th>Beam’s Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young’s modulus $E$</td>
</tr>
<tr>
<td>Poisson’s ratio $\gamma$</td>
</tr>
<tr>
<td>Density $\rho$</td>
</tr>
<tr>
<td>Length $L$</td>
</tr>
<tr>
<td>height $h$</td>
</tr>
<tr>
<td>width $b$</td>
</tr>
</tbody>
</table>

#### B. Design of the horns

In order to design the horns, the contact nature between the actuator and the beam must be ensure:

- Punctual Contact: for not to change the boundary conditions (free-free)
- Contact ensuring the stability of the beam: to prevent cause side movements (rotation, translation ...).
- Continuous and reversible contact for transmitting the vibrations. In this case: a screwed contact on a smallest possible diameter of the cone, to limit the surface of contact as shown in Fig.6. The horn is designed with $\lambda/2$ wavelength, to avoid the interaction forces between the different parts of the transducer. The design of the horns can be found in [22].

#### C. Position of transducers

The two piezoelectric actuators are positioned near the anti-nodes of each mode, in order to optimizing the electromechanical couplings and to obtain a purely transverse displacement for a non a non zero matrix $\phi_{x_1,x_2}$. The modal analysis of Fig.5 can be used to determine the position of the transducers on the anti-nodes, or the formula described in [21] can be used.

$$x_1 = x_2 = n \frac{\lambda}{2} + \frac{7}{8}$$

where $x_1$, $x_2$ are the position of the transducers from both ends of the beam, $n$ is an integer, $\lambda$ is the wavelength. The two transducers are placed at $x_1 = x_2 = 60$ mm, at the ends of the beam, with wavelength of 41.5 mm.
D. Experimental test bench

An aluminium beam is actuated by two Langevin transducers, which were associated with horns. The whole system is fixed on a solid support, which allows to move the beam for measuring the vibration velocity at every point of the beam, using an interferometer (OFV-525/-5000-S) as shown in Fig.7. A graphical user interface is used to control the vibrations amplitude of the two actuators independently, through a DSP (TI 2812) and two amplifiers (HSA 4051).

E. Identification of vibration modes

A cartography was performed to measure the vibration amplitude of each point in the beam, for the mode shape A and B as shown in Fig.8. This is used to determine the deformation mode shapes \( \phi_{A,x_1} \) and \( \phi_{B,x_2} \) we obtain:

\[
\begin{pmatrix}
\phi_{A}(x_1) \\
\phi_{B}(x_1) \\
\phi_{A}(x_2) \\
\phi_{B}(x_2)
\end{pmatrix} = \begin{pmatrix}
-0.9680 \\
-0.6949 \\
1 \\
-0.7447
\end{pmatrix}
\] (14)

F. Control of excitation’s vibration amplitude

The modal amplitude and its relative phase can be controlled directly by controlling the vibration amplitude of transducers accordingly to eq.8. By substituting eq.12 in eq.6 we obtain:

\[
\begin{pmatrix}
W_1 \\
W_2
\end{pmatrix} = W_B \begin{pmatrix}
\phi_{A}(x_1) & \phi_{B}(x_1) \\
\phi_{A}(x_2) & \phi_{B}(x_2)
\end{pmatrix} \begin{pmatrix}
1 \\
j
\end{pmatrix}
\] (15)

In this paper \( \psi_B \) is fixed to zero \( (W_B = W_b) \) giving rise to:

\[
\begin{pmatrix}
W_1 \\
W_2
\end{pmatrix} = W_B \left( \phi_{A}(x_1) + j\phi_{B}(x_1) \right) \left( \phi_{A}(x_2) + j\phi_{B}(x_2) \right) \] (16)

This may be expressed as

\[
\begin{pmatrix}
W_1 \\
W_2
\end{pmatrix} = W_B \left( \frac{\phi_{A,B}(x_1)}{\phi_{A,B}(x_2)} \right) e^{j\alpha_1} \left( \frac{\phi_{A,B}(x_1)}{\phi_{A,B}(x_2)} \right) e^{j\alpha_2} \] (17)

\( \alpha_1 \) and \( \alpha_2 \) are the phase shift between real and imaginary part of the deformation mode shape A and B, at the position of transducers. Now a travelling wave can be obtained by imposing these phases on the right and the left transducer, or by imposing a phase shift \( \alpha \) between the vibration amplitude of transducers with \( \alpha = \alpha_1 - \alpha_2 \) accordingly to Fig.3(b). The vibration amplitude of travelling wave can be controlled directly by the vibration amplitude of transducers, then eq.17 becomes:

\[
\begin{pmatrix}
W_1 \\
W_2
\end{pmatrix} = W_A \left( \frac{\phi_{A,B}(x_1)}{\phi_{A,B}(x_2)} \right) e^{j\alpha} \] (18)

A rotating reference frame related to the frequency of the vibration wave is introduced, using a Langevin transducer. The complex notation of the equation of motion about a vibration mode, and the decoupling according to two-axis allow a double independent closed loop control to regulate the real and imaginary parts of the vibration amplitude and its relative phase at any frequency, it acts directly on the amplitude of the supply voltage as a classical electromagnetic machines.
The voltage $V_\phi$ is used to control the vibration amplitude $W_d$, while $V_d$ is used to control the vibration amplitude $W_q$, through a regulator $C(s)$ for each closed loop [8]. This control does not depend on frequency accordingly to Fig.9.

![Fig. 9. schematic diagram control of a Langevin transducer](image)

The vibration amplitude in rotating frame can be determined by

$$\mathbf{W} = W_d + jW_q$$

(19)

$$\mathbf{W} = W e^{j\alpha}$$

(20)

where $W = \sqrt{W_d^2 + W_q^2}$ and $\alpha = \arctan \frac{W_d}{W_q}$ are respectively the magnitude and the argument of $\mathbf{W}$. Then the vibration amplitude of the transducers in rotating frame are done by:

$$\begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \begin{bmatrix} W_{d1} + jW_{q1} \\ W_{d2} + jW_{q2} \end{bmatrix}$$

(21)

by substitution the identified modes eq.14 in eq.18 leads to

$$\begin{bmatrix} W_{d1} + jW_{q1} \\ W_{d2} + jW_{q2} \end{bmatrix} = W_A \begin{bmatrix} 1.24e^{-j108^\circ} \\ 1.19 \end{bmatrix}$$

(22)

IV. RESULTS AND DISCUSSIONS

A. Travelling wave generation

The generation of a travelling wave was possible by excitation with vibration amplitude. The frequency was in the range between the two resonance modes, 26373 Hz and 29369 Hz. Accordingly to eq.22 a travelling wave can be obtained by imposing approximatively the same vibration amplitude in each transducer shifted with 108°.

In this paper a travelling wave was performed with a vibration amplitude of 0.5 µm, corresponding to vibration mode $W_A$ of 0.4 µm. Figure.10 depicts the evolution of the vibration amplitude of each transducer in rotating frame eq.22. For a reference step amplitude $W_{1_{ref}}$ from 0 to 0.5 µm, with $\alpha_1 = 108^\circ$ for the first actuator (the red curve), ie a step reference $W_{d1_{ref}} = 0$ to 0.47 µm, and a step of $W_{q1_{ref}} = 0$ to $-0.15$ µm. And for a reference step amplitude $W_{2_{ref}}$ from 0 to 0.5 µm with $\alpha_2 = 0^\circ$ for the second actuator (the blue curve), ie a step reference $W_{d2_{ref}} = 0$ to 0.5 µm, and a step of $W_{q2_{ref}} = 0$ µm.

![Fig. 10. Measured and simulated vibration amplitude of transducers](image)

A pure travelling wave is obtained with standing wave ratio (SWR) nearly equal to 1 at each points of the beam as shown in Fig.11, even if we change the frequency between two neighboured neutral modes shapes [26373 29369] Hz. It should be noted that, when the beam is excited near the frequency of one vibration mode, the SWR is distancing from 1.

![Fig. 11. Complex amplitude along the beam with $SWR \approx 1$](image)

This complex amplitude is obtained by measuring the vibration amplitude and its phase shift with the voltage supply of the first actuator [5]. During the test, a normal force is applied on the beam surface, this force or this load is rejected as a perturbation, and the vibration amplitude of this travelling wave is kept constant along the beam.

B. Direction change of the travelling wave

The direction of travelling wave can be inverted by changing the phase difference of the first transducer from 108° to -108°. Figure13 depicts the evolution of $W_d$ and $W_q$ as a function of time with a step phase shift from from 108° to -108°, with vibration amplitude of 0.5 µm, while the vibration amplitude of the second actuator is kept to 0.5 µm with $\alpha = 0^\circ$, ie a step reference $W_{1_{ref}} = 0.5e^{j108^\circ}$ µm to $0.5e^{-j108^\circ}$ µm.
A perfect travelling wave in the opposite direction is obtained in 10 ms accordingly the steady state.

V. CONCLUSION

In this paper, a new method is presented to generate a travelling wave, with excitation in vibration amplitude. A closed loop control is applied to each transducer to achieve this thanks to vector control method. A travelling wave has been produced and controlled in both direction and amplitude, with an optimum standing wave ratio.

ACKNOWLEDGMENT

This work has been carried out within the framework of the project StimTac of IRCICA (institut de recherche sur les composants logiciels et matériel pour la communication avancé), and the Project Mint of Inria.

REFERENCES


Measurement and Analysis of a Circular Wedge Acoustic Waveguide using a PZT Sensor

Tai-Ho Yu
Department of Electronic Engineering
National United University
Miaoli 36063, Taiwan ROC
yth@nuu.edu.tw

Abstract—This study investigated the propagation of flexural waves along the outer edge of a circular cylindrical wedge, as well as their phase velocities and the corresponding mode displacement. In this study, the dispersion curves were determined using the bi-dimensional finite element method (Bi-d FEM), as derived through the separation of variables and the Hamilton principle. According to a calculation of modal displacement, the maximum deformation appeared at the outer edge of the wedge tip. Additionally, the dispersion curves were measured using a laser-induced guided wave experiment, in which a PZT sensor measurement scheme and a two-dimensional fast Fourier transform (2D-FFT) method were used. According to the 2D-FFT calculation results of B-scan data from the laser-generated flexural waves at the circular cylindrical wedge tip, most of the measured signals belonged to the first and second modes.

Keywords—Bi-dimensional finite element method (bi-d FEM), circular cylindrical wedge-like acoustic waveguide, knife-edge technique, two-dimensional fast Fourier transform (2D-FFT).

I. INTRODUCTION


A wedge wave is a guided wave propagating along the tip of a wedge, with energy confined near the apex. Among various modes regarding patterns of particle motions, the antisymmetric flexural (ASF) modes have drawn most of the attention from researchers. With regard to the case of a linear wedge, the influences of the apex angle or of truncation on the characteristics of wedge waves have also been reported. With regard to the case a circular wedge, however, there have been no reports thus far. The objective of the current research was to explore the behavior of guided waves propagating along the tips of circular wedge acoustic waveguide by using a PZT sensor.

II. NUMERICAL ANALYSIS

A bi-dimensional finite element method (Bi-d FEM) was adopted to determine the resonant vibration modes and phase velocities of the circumferentially flexural waves. With regard to the polar coordinate system, the whole structure of the circular wedge acoustic waveguide was divided among several 2-dimensional discrete four-node isoparametric (Q4) elements. By separating the variables, the elastic displacements \( u \) (for the global structure) and \( D \) (for elements) are represented as

\[
\begin{align*}
\mathbf{u} &= \begin{bmatrix} u_r(r, x) \\ u_\theta(r, x) \\ u_\theta(r, x)e^{i\theta} \end{bmatrix} \quad \text{ND}
\end{align*}
\]

where the circumferential mode number \( n \) could be a non-integer for traveling waves, but it must be an integer for standing waves. The matrix \( N \) refers to interpolation functions in conjunction to each nodal displacement and electric displacement. It is noted that \( u_\theta \) has a 90° phase lag in comparison to the other two elastic displacement components. Using Hamilton’s principle, a system of equations for wedge waves traveling along the circular wedge acoustic waveguide can be calculated as follows

\[
\mathbf{K} \mathbf{D} + \mathbf{M} \mathbf{D} = 0
\]

where \( \mathbf{M} \) and \( \mathbf{K} \) are the global mass matrix and the global elastic stiffness matrices. If all the field variables are time-harmonic, Eq. (2) can result in a dispersion equation or frequency equation as follows

\[
\det (\mathbf{K} - \omega^2 \mathbf{M}) = 0
\]

The resonant modes of flexural waves can be calculated once the angular frequency \( \omega = 2\pi f \) and circumferential wave number \( k = n/R \) in Eq. (3) are determined. The \( R \) value is the radius of the circular wedge acoustic waveguide.
III. SIMULATION RESULTS

The prototype of the circular cylindrical wedge-like acoustic waveguide analyzed in this study had a vertex angle $\phi = 15^\circ$, a radius $r_o = 15$ mm, and a wedge height $h = 42.86$ mm. The waveguide was manufactured using 99.7% pure Titanium (Ti, density $\rho = 4.85$ g/cm$^3$, Elastic modulus $E = 102.0$ GPa, Poisson ratio $\nu = 0.3$) and a precision lathe turning process. The phase velocity and modal displacement of the flexural acoustic waves were calculated according to the bi-d FEM. The sectional view of the wedge was meshed into 254 Q4 elements and 284 nodes, with the elements divided more densely toward the vertex of the wedge. Each node featured three degrees of freedom, and the six nodes on the bottom were fixed.

Fig. 1 illustrates the dispersion curve of the bi-d FEM numerical solution as calculated by FORTRAN. Fig. 2 shows the mode shapes of the flexural waves of the circular cylindrical wedge with mode number $N = 1$–5.

Fig. 1. Simulated dispersion curves of a circular wedge.

![Dispersion Curve](image1)

Fig. 2. Mode shapes of a circular wedge ($N = 1$–5).

![Mode Shapes](image2)

Fig. 3. Schematic of the measurement process.

IV. EXPERIMENTAL RESULTS

Fig. 3 depicts the framework of the laser ultrasonic experiment. Four mirrors were mounted on an electric rotating platform (turn table). The platform was rotated using an NI-7344 motion control card (National Instrument Corp., TX, USA). A high-energy Nd:YAG normal incidence pulsed laser (GCR-100, Spectra-Physics Quanta-Ray Inc., San Jose, CA, USA, 200 mJ, $\lambda = 1064$ nm) was focused on various scanning points along the outer edge of the wedge tip, subsequently creating laser-generated flexural waves transmitted from the wedge tip of the acoustic waveguide. A small piece of PZT disk (PZT-5H, d31 poling direction) was used as the detecting sensor, and flexural waves were determined at the outer edge of the wedge tip by using a digital storage Oscilloscope LeCroy WS42Xs (LeCroy Corp., New York, USA). To reduce the noise interference, each data group was evaluated 10 times on average. The sampling frequency was 20 MS/s, and each data length was 1024 points. With each 0.1° turn of the rotating platform, a fraction of the waveform data was captured; 256 data were evaluated.

Fig. 4 depicts the experimental setup, and Fig. 5 shows a photograph of the circular wedge acoustic waveguide measurement setup. Fig. 6 shows the scanned image of a circular wedge acoustic waveguide. Fig. 7 shows a comparison of the c-k dispersion curve as determined by the simulation and by the measurement results.

Fig. 4. Experimental setup.

![Experimental Setup](image3)

Fig. 5. Photograph of the experimental setup.

![Photograph](image4)
V. CONCLUSIONS

By applying the bi-d FEM, this study investigated the phase velocities and resonance mode of flexural waves in a circular wedge acoustic waveguide. The experimental values of the flexural wave phase velocities were determined using laser ultrasonic measurement technology and 2D-FFT calculation, and the accuracy of the numerical calculation was confirmed. According to the displacement of the resonance mode, the guided waves at the tip of the circular cylindrical wedge were flexural waves. The greatest displacement occurred at the wedge tip, indicating that the energy of the flexural wave was concentrated on the wedge tip.

ACKNOWLEDGMENT

The author would like to thank the National Science Council of Taiwan, ROC, for financially supporting this research under contract NSC 103-2221-E-239-027.

REFERENCES


Characterization and Temperature Sensor Application of Ca$_3$TaGa$_3$Si$_2$O$_{14}$ Crystals

Hongfei Zu, Huiyan Wu, Qing-Ming Wang
Department of Mechanical Engineering & Materials Science
University of Pittsburgh
Pittsburgh, USA
qiw4@pitt.edu

Quanming Lin, Yanqing Zheng
Shanghai Institute of Ceramics
Chinese Academy of Sciences
Shanghai, China

Abstract—In this paper, the elastic, piezoelectric, and dielectric constants of Ca$_3$TaGa$_3$Si$_2$O$_{14}$ (CTGS) single crystals were fully characterized from room temperature to 800°C according to IEEE standard methods. The sensitivity and stability of the temperature sensors based on CTGS crystals were also investigated. One of the elastic stiffness constants $c_{11}$ shows the highest temperature sensitivity. The results of the repeated measurements of X-cut square-plate resonator sample indicate its excellent stability through the entire measurement temperature range.

Keywords—high-temperature crystals; high-temperature sensors

I. INTRODUCTION

Due to the increasing needs for advanced sensor technology which can be applied in extreme environments, the high-temperature sensors have won great attention in aerospace, auto-motive, and energy industries in the past few years [1-2]. The sensors used in these fields are sometimes required to be able to accurately and stably detect or monitor certain desired signals under the high temperatures up to or greater than 1000°C with a sufficiently long lifetime around 100,000 h [3-4]. Given the applications of the conventional sensing materials, such as Quartz and Gallium Orthophosphate (GaPO$_4$), are restricted by their phase transitions [5-6], the study of novel materials without phase transition has become a highly valued hotspot.

The langasite family crystals, which exhibit no phase transitions prior to their melting points (1300–1500°C) [2], quickly came to the fore among the newly investigated high-temperature materials, owing to their excellent electromechanical properties and temperature behaviors [7-8]. As a member of langasite family crystals, Ca$_3$TaGa$_3$Si$_2$O$_{14}$ (CTGS) has demonstrated the great promising as one of the best high temperature acoustic sensors by its most stable performance. Additionally, the effective piezoelectric constant, electromechanical coupling coefficient, and relative permittivity of CTGS are reported to be 4.6 pC/N, 12%, and 18.2 [9-10], respectively, which are all superior to quartz’s.

Though numerous research works have been carried out to evaluate the high-temperature performance of CTGS crystal, the full characterization of the dielectric, piezoelectric, and elastic constants and the investigation about its sensitivity and stability under high temperatures are currently not available yet from literature. In our recent study [11], we have completely characterized all elastic, dielectric and piezoelectric constants of the CTGS crystal resonators with different cuts. The results have provided the basic materials parameters for device design fabrication and applications. Here in this paper, we will summarize some very important results of the CTGS resonators, which have demonstrated that CTGS crystal resonators are very promising for high-temperature sensing applications. With appropriate crystal cuts, CTGS crystal resonators can provide excellent high-temperature sensitivity, as well as excellent high temperature stability.

In this work, the temperature dependence in the range from ambient temperature to 800°C of all elastic, dielectric, and piezoelectric coefficients of CTGS resonators are investigated according to IEEE standard methods [12]. Based on the sensitivity and stability of crystal resonators with different cuts in the wide high-temperature range, the optimal orientation for the high-temperature sensors has been determined.

II. EXPERIMENTAL

A. Sample Preparation and Measurements

The CTGS crystals were grown by the Czochralski (CZ) pulling technique. The powders of four high purity raw materials: CaCO$_3$, Ta$_2$O$_5$, Ga$_2$O$_3$, and SiO$_2$, were first mixed in stoichiometric ratio and sintered at 1250°C for 24 h, and then loaded into an Iridium (Ir) crucible to melt at 1350°C. Subsequently, the melted polycrystalline materials were pulled by a CTGS crystal with (110) direction to form the desired single crystals at a rate of 0.5-1 mm/h with the rotation of 5-15 rpm in an automatic diameter control (ADC) CZ furnace. The as-grown crystals were eventually cooled down to room temperature at a rate of 30 °C/h to avoid the possible cracking.

To obtain all of the elastic, piezoelectric, and dielectric constants requires measuring the properties of seven kinds of carefully oriented crystal resonators which can be divided into two categories: square plates (X-cut, Y-cut, and Z-cut), and rectangular bars (XY-cut, (XYt)45°, (XYt)-30°, and (XYt)-85°). The notations of the sample-cuts follow the IEEE standard [12] and the schematic of the seven kinds of resonators is shown in Fig. 1. The reason for selecting these resonators as test objects will be explained in detail later.
The $e_{ij}$ and $c_{ij}$ matrices are similar to the $d_{ij}$ and $s_{ij}$ matrices, respectively, except $e_{26}=e_{11}$, $c_{66}=c_{65}=c_{14}$, and $c_{66}=(c_{11}-c_{12})/2$.

First, the admittance (Y-) spectra of the four bar-shaped resonators (XY, (XYt)45°, (XYt)-30°, and (XYt)-85° cuts) were measured. Three elastic compliances ($s_{11}^{E}$, $s_{33}^{E}$, and $s_{44}^{E}$), and one combination ($2s_{13}^{E}+s_{44}^{E}$) can be obtained by the following equations:

$$
s_{11}^{E} = \frac{1}{4\rho(f_{r}-f_{0})^2}
$$

$$
s_{12}^{E}(\theta) = s_{11}^{E}\cos^2\theta + s_{33}^{E}\sin^2\theta - 2s_{13}^{E}\cos\theta\sin\theta \\
+ (2s_{13}^{E} + s_{44}^{E})\cos^2\theta\sin^2\theta
$$

where $\rho$ (≈4.61 g/cm$^3$) is the density of the crystals, $l$ is the length of the bar-samples, and $f_r$ is the resonant frequency of the resonators.

Then through the measurements of the impedance (Z-) or admittance (Y-) spectra of Y-cut and Z-cut square-plate resonators, $e_{66}^{E}$ and $c_{44}^{E}$ could be determined according to:

$$
c_{66}^{E} = 4\rho(f_{r-Y})^2; c_{44}^{E} = 4\rho(f_{r-Z})^2
$$

where $t$ is the thickness of the plates, and $f_{r-y}$ and $f_{r-z}$ are resonant frequencies of Y-cut and Z-cut vibrators, respectively.

Because the elastic compliance matrix $s$ and stiffness matrix $c$ are reciprocal: $c_{ij}=s_{ji}^{-1}$, we can get the following relations:

$$
2c_{11} = \frac{s_{11}^{E} + s_{12}^{E}}{\alpha}; 2c_{12} = \frac{s_{33}^{E} - s_{44}^{E}}{\beta}; c_{13} = \frac{-s_{33}^{E}}{\alpha},
$$

$$
c_{14} = \frac{s_{33}^{E} + s_{44}^{E}}{\beta}; c_{33} = \frac{s_{11}^{E} + s_{12}^{E}}{\alpha}; c_{44} = \frac{s_{11}^{E} - s_{12}^{E}}{\beta},
$$

$$
c_{66}^{E} = c_{11}^{E} - c_{12}^{E} = \frac{s_{44}^{E}}{2\beta}.
$$

where $\alpha = s_{33}(s_{11} + s_{12}) - 2s_{13}^{E}$;

$$
\beta = s_{44}(s_{11} - s_{12}) - 2s_{13}^{E}.
$$

Finally, by substituting the values of $s_{11}^{E}$, $s_{33}^{E}$, $s_{14}^{E}$, $2s_{13}^{E}+s_{44}^{E}$, $c_{66}^{E}$ and $c_{44}^{E}$ into Eqs. (7), all the remaining $s_{ij}^{E}$ and $c_{ij}^{E}$ can be obtained.

By capacitance measurements of X-cut and Z-cut square-plate resonators at low 2 kHz, the relative dielectric constants can be determined:

$$
e_{11,33}^{I} = \frac{C_{11,33}^{I}-1}{A}, e_{11,33}^{F} = \frac{C_{11,33}^{F}}{e_{0}},
$$

$$
e_0 = 8.85\times10^{-12} F/m
$$

B. IEEE Methods for the Determination of the Constants

CTGS crystal is belong to the point group 32, so the dielectric, piezoelectric, and elastic constants matrices are as follows:

$$
\varepsilon = \begin{bmatrix}
\varepsilon_{11} & 0 & 0 \\
0 & \varepsilon_{11} & 0 \\
0 & 0 & \varepsilon_{33}
\end{bmatrix}
$$

$$
d = \begin{bmatrix}
d_{11} & -d_{11} & 0 & d_{14} & 0 & 0 \\
0 & 0 & 0 & 0 & -d_{14} & -2d_{11}
\end{bmatrix}
$$

$$
s = \begin{bmatrix}
s_{11} & s_{12} & s_{13} & s_{14} & 0 & 0 \\
s_{12} & s_{11} & s_{13} & -s_{44} & 0 & 0 \\
s_{13} & s_{13} & s_{33} & 0 & 0 & 0 \\
s_{14} & -s_{44} & 0 & s_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & s_{44} & 2s_{14} \\
0 & 0 & 0 & 0 & 2s_{14} & 2(s_{11} - s_{12})
\end{bmatrix}
$$

Fig. 1. Schematic of the CTGS sample-cuts required for the complete determination of the elastic, piezoelectric, and dielectric constants.
where $C$, $t$ and $A$ are the low frequency capacitances (measured at 2 kHz), thicknesses, and effective areas of the samples, respectively.

As for the two piezoelectric coefficients, they can be obtained by the admittance measurements of any two of the four rectangular bar-resonators through the following equations:

$$
\frac{k_{12}^2(\theta) - 1}{k_{12}^2(\theta)} = \frac{\pi}{2} \left[ \frac{f_r(\theta)}{f_a(\theta)} \right];
$$
$$
d_{12}^L(\theta) = k_{12}^2(\theta) e_{11}^{r, s} e_{22}^{r, s};
$$
$$
d_{12}^E(\theta) = -d_{11}(1 + \cos 2\theta) + d_{14} \sin 2\theta;
$$
$$
e_{ij} = d_{ji} e_{ij}^E
$$

where $k_{12}$ is the electromechanical coupling coefficient, and $f_r$ and $f_a$ are the corresponding resonant and anti-resonant frequencies, respectively.

C. Temperature Sensor Application and the Investigation of the Sensitivity and Stability

At the beginning of delving into the “sensitivity” issue, it is necessary to determine which parameter should be exploited to describe it. Since the resonant frequencies of the crystal resonators are the most concerned parameters for acoustic wave resonator sensors, and the elastic constants are the primary parameters directly related to the resonant frequencies [referring to Eqs. (4), (5), and (6)], the temperature dependence of elastic constants is chosen to represent the temperature sensitivity of the resonators. In addition, considering that the linearity of the temperature dependence of elastic constants will determine the applicability of the resonator sensors, high temperature dependence linearity is always desirable. We will only consider the temperature dependence of the elastic constants with data fitting linearity, obtained from the experimental data set (denoted by “R”), higher than 0.9. In addition, to eliminate the influence caused by the difference of the absolute values between various elastic components, it is reasonable to employ their relative changes rather than the absolute values.

To evaluate the stability of the CTGS resonator sensors, the repeated measurements of the selected resonator sample (according to the sensitivity and linearity) were performed four times with one-month interval in the duration of three months. The results will, to some extent, indicate whether the sensors can work repeatedly and stably at high temperature conditions.

III. RESULTS AND DISCUSSIONS

A. Elastic, Piezoelectric, and Dielectric Constants

According to the previously summarized IEEE method, all the elastic, piezoelectric, and dielectric constants of CTGS crystals from room temperature to 800°C were determined. And the results at three selected temperature points (21°C, 400°C, and 800°C) are shown in Table I. For more details and information about these materials parameters and their temperature dependence, please refer to the paper to be published [11].

B. Sensitivity and Stability of the Temperature Sensors

The results of all the elastic constants were linearly fitted, and four of them possess excellent linearity verse temperature ($R > 0.9$): $s_{11}^{E}$, $s_{66}^{E}$, $c_{11}^{P}$, and $c_{66}^{P}$. The temperature dependence of $s_{11}^{P}$ and $s_{66}^{P}$ are shown in Fig. 2. Fig. 3 shows the linear relationship of $c_{11}^{E}$ and $c_{66}^{E}$, verse temperature. As can be seen, the linear regression R values of $s_{11}^{P}$, $s_{66}^{P}$, $c_{11}^{P}$, and $c_{66}^{P}$, verse temperature are 0.9881, 0.9961, 0.9093, and 0.9946, respectively. However, their temperature sensitivities, which are represented by the slopes, are $8.8 \times 10^{-4}$, $1.9 \times 10^{-3}$, $1.0 \times 10^{-2}$, and $3.6 \times 10^{-3}$, respectively, indicating that $c_{11}^{E}$ exhibits the highest temperature sensitivity.

$c_{11}^{E}$ can be calculated by Eqs. (4)-(7), and also can be directly determined through the impedance measurements ($R$-X, resistance and reactance) of X-cut square-plate resonators. The impedance ($R$-X) of one randomly selected X-cut square-plate resonator was repeatedly measured four times with one-month interval in the duration of three months, and the results are graphically represented in Fig. 4. It is obvious that the impedance essentially stays unchanged during the entire course of tests, indicating the excellent high-temperature stability of X-cut CTGS crystal resonator.

| TABLE I. RELATIVE DIELECTRIC CONSTANTS, PIEZOELECTRIC COEFFICIENTS ($d_{ij}: pC/N$, $e_{ij}: C/m^2$), AND ELASTIC CONSTANTS ($s_{ij}^{E}: pm^2/N$, $c_{ij}^{E}: 10^9 N/m^2$) OF CTGS CRYSTAL AT THREE SELECTED TEMPERATURE POINTS |
|-----------------|------|------|------|
| Temperature     | 21°C | 400°C| 800°C|
| $e_{11}^{S}$    | 22.33| 25.43| 43.31|
| $e_{23}^{S}$    | 23.42| 25.25| 41.16|
| $d_{11}$        | -5.35| -5.18| -6.41|
| $d_{15}$        | 11.72| 11.19| 12.54|
| $e_{11}$        | -0.49| -0.46| -0.55|
| $e_{14}$        | 0.57 | 0.52 | 0.54 |
| $s_{11}^{E}$    | 8.72 | 8.97 | 9.40 |
| $s_{23}$        | 6.13 | 6.33 | 6.89 |
| $s_{15}$        | 20.61| 21.59| 23.41|
| $s_{16}$        | 22.05| 22.89| 23.58|
| $c_{11}^{E}$    | -2.31| -2.47| -2.40|
| $c_{14}$        | -1.97| -2.16| -2.37|
| $c_{15}$        | 0.092| 0.10 | 0.12 |
| $c_{16}$        | 142.6| 143.44| 135.6|
| $c_{23}$        | 203.3| 204.6| 189.3|
| $c_{24}$        | 48.53| 46.33| 42.73|
| $c_{26}$        | 45.35| 43.70| 42.41|
| $c_{12}$        | 51.91| 56.04| 50.74|
| $c_{13}$        | 62.52| 68.13| 64.14|
| $c_{14}$        | -0.41| -0.40| -0.40|
IV. CONCLUSION

In summary, the elastic constants, piezoelectric coefficients, and dielectric constants and their temperature dependences of CTGS single crystal resonators have been investigated from room temperature to 800°C. Seven kinds of oriented crystal resonators: X-cut, Y-cut, and Z-cut square-plate, and XY, (XYt)45°, (XYt)-30°, and (XYt)-85° cuts rectangular-bar resonator samples were designed and fabricated for complete characterization of the materials parameters according to IEEE standard methods. Moreover, the temperature sensor applications of CTGS were also studied. The elastic constants with larger than 0.9 linear regression R value verse temperature are chosen to define the sensitivity of the resonator sensors, among which $c_{11}$ shows high temperature sensitivity. Finally, an X-cut square-plate resonator sample was repeatedly measured four times in three months to test the stability of the resonator sensor, and an excellent stability was obtained.

ACKNOWLEDGMENT

The author would like to acknowledge the financial support by the US National Science Foundation (NSF) under award No. ECCS-0925716 for this work.

REFERENCES

Improvement in tracking loop threshold of high dynamic GNSS receiver by installation of crystal oscillator on gyroscopic mounting

Maryam Abedi, Tian Jin
School of Electronic and Information, Beihang University
Beijing 100191, China
E-mail: abedi_maryam@buaa.edu.cn

Abstract— In high dynamic GNSS receivers, replica carrier is polluted by dynamic loads which modulated on it and generate clock drift as frequency and phase error. These dynamic loads consist of steady state acceleration, sinusoidal vibration, mechanical and acoustic random vibrations. As steady state load exists all over the trajectory and causes great phase deviation, therefor phase lock is lost and the sensitive phase lock detector distinguishes it and transits back to FLL, therefore the dominant tracking loop in high dynamic receivers is FLL. According to the calculations, all dynamic loads cause some amount of frequency deviation ($f_\text{ref}$) or jitter ($3\sigma_FLL$), it is also worth remembering that among them random vibration causes great frequency jitter and impacts FLL significantly.

A gyroscopic mounting is introduced to install crystal oscillator on PCB. This mounting affects dynamic loads-induced disturbances and reduces the probability of signal loss especially for crystal oscillators with g-sensitivity angle $\varphi \leq 30^\circ$.

In the case of steady state load, gyro mounting besides reduction of frequency deviation, shifts high values from most probable g-sensitivity angles $\varphi$ (i.e. $\varphi \leq 30^\circ$) to low probable ones. In the case of sinusoidal vibrations, gyro is able to reduce induced frequency jitter especially for $\varphi \leq 30^\circ$. As well as Gyro shows its significant effects on random vibration-induced frequency jitter as the main disturbance source of FLL and reduces the probability of signal loss, especially for crystals with g-sensitivity angle $\varphi \leq 30^\circ$. In this way, gyro mounting improves the tracking loop threshold in high dynamic GNSS receivers.

Keywords— High Dynamic GNSS Receiver, GPS, Gyro, Mounting, Tracking Loop, FLL, PLL, Random Vibration, Sinusoidal Vibrations, Steady State Acceleration, Acoustic Pressure Load, Launch Vehicle.

I. INTRODUCTION

In high dynamic GNSS receiver, there are two groups errors which affect signal tracking process, first comes from input signal which is polluted on the channel between satellite and receiver by thermal noise and dynamic state of host vehicle (i.e. velocity, acceleration and jerk), second comes from replica carrier which is polluted by oscillator inherent error (Allan deviation), dynamic state of host vehicle and vibration loads.

This paper focused on the second group of errors and tries to improve signal tracking of high dynamic receivers by reduction of these errors.

PLL tracking threshold:

\[
3\sigma_{PLL} \leq \text{Phase pull-in range of the PLL discriminator}/4 = 180^\circ/4 \\
\rightarrow 3\sigma_{PLL} \leq 45^\circ \text{ & } 1\sigma_{PLL} \leq 15^\circ; \quad [1]
\]

\[
3\sigma_{PLL} = \theta_{\text{PLL}} + 3\sigma_{A} = \theta_{\text{PLL}} + 3\sqrt{\sigma_{PLL}^2 + \sigma_{vPLL}^2 + \sigma_{A}^2};
\]

FLL tracking threshold:

\[
3\sigma_{FLL} \leq \text{Frequency pull-in range of the FLL discriminator}/4 = 1/4T (T=20\text{ ms}) \rightarrow 3\sigma_{FLL} \leq 12.5 (\text{Hz}) \text{ & } 1\sigma_{FLL} \leq 4.17 (\text{Hz}); \quad [1]
\]

\[
3\sigma_{FLL} = f_{\text{PLL}} + 3\sigma_{A} = f_{\text{PLL}} + 3\sqrt{\sigma_{FLL}^2 + \sigma_{vFLL}^2 + f_{\text{A}}^2};
\]

Incoming signal errors: $\theta_{\text{PLL}}$=dynamic stress phase error, $f_{\text{PLL}}$=dynamic stress frequency error, $\sigma_{PLL}$= thermal noise phase jitter, $\sigma_{vPLL}$= thermal noise frequency jitter.

Replica carrier errors: $\theta_{\text{PLL}}$=dynamic stress phase error, $f_{\text{PLL}}$= dynamic stress frequency error, $\sigma_{PLL}$= vibration-induced frequency jitter, $\sigma_{vPLL}$= vibration-induced frequency jitter, $\sigma_{A}$= Allan deviation-induced phase jitter, $f_{\text{A}}$= Allan deviation-induced frequency jitter.

II. HIGH DYNAMIC GNSS RECEIVER

![Fig. 1. (left) GPS receiver installed on centaur forward adaptor of launch vehicle AtlasV [2], (middle) position of centaur forward adaptor in launch vehicle [3], (right) total view of launch vehicle [3].](image)

To do this study assumed that GNSS receiver installed on launch vehicle as high dynamic GNSS receiver. There are external dynamic loads applied to host vehicle as long as moves on its trajectory. Its response to these loads applies to equipment installed on it, e.g. GNSS receiver. This response depends to the position of equipment on the vehicle. As GNSS receiver is installed on the centaur forward adaptor inside fairing (Fig.1), [2] and [4], where other electronic equipment installed therefore that is enough to focus on the response of
this part of launch vehicle to external loads.

To do numerical analysis assumed that high dynamic host vehicle is launch vehicle Arian and GNSS receiver installed on it is a GPS receiver equipped with an OCXO with $\Gamma=10^{-9}$, Q-factor=$10^{11}$, uses $L_1$ carrier=$154 \times f_0$, $f_0=10.23$ MHz.

III. DYNAMIC LOADS

Fig. 2. Clock drift sources applied on crystal oscillator on high dynamic receiver; (up-left) steady state acceleration as deviation source; (up-right) sinusoidal vibration as jitter source (down) random vibration as jitter-source.

Dynamic loads applied on crystal oscillator and cause frequency and phase errors on replica carrier are:

1- **Steady State Acceleration**: Time variant Acceleration of host vehicle on the trajectory on either longitudinal (trust) or lateral directions (maneuver). [5]

2- **Sinusoidal Vibrations (2-100 Hz)**: A sweep of all the sinusoidal excitations of frequency 2 to 100 (Hz) with different magnitude on longitudinal and lateral directions. [5-7]

3- **Mechanical & Acoustic Random Vibration**: Random Vibration is a bandlimited noise which is a combination of all the excitations of 20-200 (Hz) for mechanical ones and 20-10000 (Hz) for acoustic ones. It follows the Gaussian distribution in time domain [4-6], [8-12].

IV. DYNAMIC LOADS SOURCES

The external loads which affect launch vehicle and generate the dynamic loads which cause the disturbances on replica carrier could be divided into [3] and [5]:

**Trust** is the propulsion which is generated by engines minus drag forces in the trajectory appears as steady state acceleration.

**Rotating devices** such as engines generate sinusoidal vibrations which exist during powered flight and it is in maximum state during atmospheric flight and at the points of start and shut down of engines.

**Sound pressure load** generated by plumes, boundary layer turbulence, air separation and shock waves. This load appears as acoustic random vibration.

**Separation** of boosters, engines and fairings which generate shock load on launch vehicle, appears as mechanical random vibration. [3]

V. GYRO MOUNTING

Fig. 3.(left) gyro-mounting (middle) dynamic load applied on oscillator installed on gyro (right) gyro on PCB.

A gyroscopic mounting (Fig. 3) is introduced to install the crystal oscillator on PCB [13]. It gives the freedom to crystal to rotate freely around roll, pitch and yaw. Gyro response is different for each dynamic load.

**Steady State Acceleration**: Since this load is time variant, gyro rotates as acceleration can be perpendicular to crystal surface in any given moment.

**Vibration Loads**: Gyro rotates as the $A_{zp}$ of these loads can be perpendicular to crystal surface.

**Combined Load**: Since at the most parts of host vehicle trajectory, the combination of dynamic loads applied to oscillator, thus gyro rotates as the resultant load can be perpendicular to crystal surface in any given moment.

VI. FREQUENCY AND PHASE DEVIATION CAUSED BY DYNAMIC LOADS

**Frequency Error**: $\Delta f=154 f_0 \Gamma A=154 f_0 \Gamma A \cos \alpha \cos \beta$ (Hz); (5)

**Phase Error**: $\Delta \phi = 2 \pi \int \Delta f \ dt=2 \pi \int 154 f_0 \Gamma A \cos \alpha \cos \beta \ dt$ (rad); (6)

Where: $0<|\beta|<\pi$, $0<|\alpha|<\pi$; $A$: dynamic load, $u$: angle between load $A$ and g-sensitivity $\Gamma$, $\beta$: angle between pages through load $A$ and g-sensitivity $\Gamma$.

VII. TRAJECTORY AND SIGNAL LOSS

In high dynamic GNSS receiver, the Probability of signal loss in each step of flight is different. For ease of understanding, Dynamic loads and their sources are shown on the trajectory of Arian5 in Fig.4, Fig.5 and Fig.6 [4-5], [8-9], [14-21]. This data can be used to estimate where the probability of tracking loss is the highest. Then we can find the main reasons of signal loss and try to reduce them and their impacts on replica carrier or to improve tracking loop by making it more robust against these loads.

To make decision, it is necessary to analyze the impact of each dynamic load on replica carrier, then combine the loads according to these figures, compare the disturbances on it, before and after using proposed gyro mounting. Then it is possible to decide about the most probable signal tracking loss.
VIII. ANALYSIS OF DYNAMIC LOADS IMPACTS ON REPLICA CARRIER AND PROOF OF GYRO EFFICIENCY

A. Steady State Acceleration (g)

For launch vehicle Ariane5, the longitudinal steady state load on all over the trajectory (t=0-1500s) is shown in Fig.6 [5]. As density of air and mass of vehicle change on each step of flight, therefore this load is time variant acceleration.

In high dynamic GNSS receiver, this load in turn causes frequency and phase deviation on replica carrier.

As shown in Fig.1, GPS receiver installed on the Centaur forward adaptor of launch vehicle by an angle (not horizontal) [2], [3], and this angle is different from one launch vehicle to the other. Therefore in order to obtain the max gyro effect, this angle assumed as 90° (vertically installation) in calculations.

1) gyro-mounting Effect on Steady State Load-induced disturbances on replica carrier

Since steady state load exists all over the trajectory and causes great phase deviation, therefor phase lock is lost and the sensitive phase lock detector distinguishes it and transits back to FLL [1], that is why the dominant tracking loop in high dynamic receivers is FLL.

As seen in Fig.9, steady state load-induced frequency deviation in some parts of trajectory is greater than $1\sigma_{FLL}$, especially in atmospheric flight when according to Fig.6 a combination of all disturbances applied to oscillator and probability of signal loss is the highest. Whatever in this case gyro-mounting presents its great effects and aside from reduction of deviation, shifts large values from most probable cases when ($\varphi < 30°$) to low probable ones ($\varphi > 30°$) [13].

B. Sinusoidal Vibrations (2-100 Hz)

It is a sweep of all the sinusoidal excitations of frequency 2 to 200 (Hz) with different amplitude on longitudinal and lateral directions as shown for ARIAN5 in Fig.7 [5].

This dynamic load causes frequency and phase jitter on replica carrier and affects the tracking loop.

1) gyro-mounting Effect on sinusoidal vibration induced disturbances on replica carrier

Sinusoidal Vibrations cause large phase jitter (Fig.10). However as described, in the case of high dynamic receiver, the tracking loop is FLL and fortunately the sinusoidal-induced frequency jitter is below $1\sigma_{FLL}$ threshold before using gyro, and gyro reduces it significantly, especially for most probable $\varphi \leq 30°$.

C. Mechanical Random Vibration (20-2000 Hz)

Mechanical Random Vibration is a bandlimited (20-2000 Hz) noise with Gaussian distribution which is defined by PSD (g^2/Hz) in frequency domain. PSD (g^2/Hz) of launch vehicle Ariane is as shown in Fig.8 [9]. In this spectrum the amplitudes greater than $3\sigma$ have been truncated, that is why a flat line seen on it.
Since Random Vibration is a combination of all the frequencies at the same time, therefore to calculate its impact on replica carrier, at first it is necessary to configure this load in time domain. According to the parseval’s law, $g_{rms}$ is equal to $\sigma$ of Random Vibration. Therefore Random Vibration could be indicated in time domain by calculation of $g_{rms}$ from PSD.

\[ g_{rms} = \sqrt{PSD \times f} \]  

\[ (7) \]

\[ \text{a) Calculation of } g_{rms} \]

\[ g_{rms} = \sqrt{PSD (g^2/Hz) \times f (Hz)} \]

\[ (7) \]

\[ \text{b) Effect of gyro-mounting on Random Vibration-induced disturbances on replica carrier} \]

Maximum gyro effects and maximum gyro drawbacks on random vibration-induced frequency jitter and consequently on signal tracking loop is as shown in Fig.11 and Fig.12.

In critical state ($\beta$=0 & $\xi$=90°), gyro-mounting is able to reduce the frequency jitter generated by vibrations with amplitudes $\leq 1\sigma$ (i.e. nearly 68% of the trajectory which is polluted by this load), for crystals with g-sensitivity angle $|\phi| \leq 30^\circ$, from values close to 3$\sigma_{FLL}$ threshold to nearly close to or below 1$\sigma_{FLL}$ threshold. In safe state ($\beta=90^\circ$ or $\xi-\phi=90^\circ$), gyro mounting induces some disturbances on oscillator output, but maximum value of this disturbance for vibrations with...
amplitudes ≤ 1σ is nearly equal to 1σ_FLL, therefore gyro does not cause signal loss on nearly 68% of the trajectory which is polluted by this load.

By this way, gyro-mounting totally reduces the probability of signal loss on all over the high dynamic receiver trajectory when random vibration applied on it.

![Fig.12. (left) Max effects and (right) max drawbacks of gyro-mounting on random vibration-induced frequency jitter, (left) on critical state when β=0, ξ=φ, (right) on safe state when β=90 or ξ−φ=90° and 0<φ<90°, Γ=10−9/g.](image)

**D. Acoustic Random Vibration (20-3000 Hz)**

Acoustic pressure load generates by different sources at each step of atmospheric flight as shown in Fig.5 for Arian5. Acoustic pressure affects crystal oscillator output by 2 ways (Fig.13) [3], [8], [22]:

- Acoustic load applied directly on any surfaces and generates Acoustic Random Vibration. In some electronic devices like crystal oscillator this random vibration converts to Δv and so-called Microphonic effect. This effect causes disturbances on the oscillator output as frequency and phase jitter.

- Acoustic load is converted to Mechanical Random Vibration on the plates such as PCBs and transfers from connectors to equipment installed on it (e.g. crystal oscillator). Therefore as thinner and broader PCB, stronger random vibration is generated, also as more flexible connection between oscillator and PCB, less random vibration transfers to crystal [22].

**1) Acoustic Random Vibration in time domain**

Acoustic Random Vibration is a bandlimited (20-2828 Hz) noise with Gaussian distribution which is defined by noise spectrum (dB-Hz). This spectrum for inside the fairing of launch vehicle Arian5, where GPS receiver installed, is shown in Fig.14 (up) [5]. Each horizontal line is the sound pressure level (dB) for specified frequency band in octave.

The Acoustic random vibration like mechanical one is a combination of all the frequencies at the same time, therefore to calculate its impact on replica carrier; at first it is necessary to configure this load in time domain (Fig.14, down). The amplitude of this load in time domain is sound pressure (kPa) and 1σ of this vibration is equal to Prms.

![Fig. 14. Acoustic Random Vibration inside the fairing of Arian5 (up) Spectrum (dB-Hz); (down) in time domain](image)

**a) Calculation of Prms**

$$OASPL = 20 \log \left( \frac{P_{rms}}{P_{ref}} \right)$$  \hspace{1cm} (11)

$OASPL= 139.5$ (dB); $f= 20–2828$ (Hz); $P_{ref}=20$ ($\mu$Pa),

$P_{rms}=189$ (Pa); $(OASPL =$Overall Acoustic Sound Pressure Level)

$1\sigma=P_{rms}=0.189$ (KPa); $2\sigma=0.378$ (KPa); $3\sigma=0.567$ (KPa);        \hspace{1cm} (12)

**b) Effect of gyro-mounting on Acoustic Random Vibration-induced disturbances on replica carrier L1**

The calculation method of Acoustic Random Vibration-induced disturbances are the same as Mechanical Random Vibration mentioned above. The only difference is that acceleration must be calculated from Acoustic Pressure Load by consideration of mass, area and acoustic absorption coefficient [3] of crystal blank.

$$A (g) = P \times \alpha \times k / 9.8;$$  \hspace{1cm} (13)

Where: $A=$acceleration applied by Acoustic Load; $P=$ amplitude of Acoustic Random Vibration in time domain, $-4\sigma<P<4\sigma$; $\alpha =$Sound absorption coefficient of crystal; $k=$Surface Area /Mass of crystal;

**c) Conclusion**

According to the empirical experiments the main signal loss occurs during atmospheric flight, in this part of trajectory all
the 3 dynamic loads affect replica carrier (Fig.6), especially acoustic load, it affects crystal oscillator output by two ways and according to the calculations random vibration always causes great frequency jitter on replica carrier. Moreover as described on all over the trajectory of high dynamic receiver, tracking loop automatically transits back to FLL. Therefore in order to prevent signal loss during ascent flight, it is necessary to reduce the influence of this load on replica carrier in a way that total frequency deviation could be less than 3σ_{FLL} threshold. It can be solved by using gyro-mounting. It is proved numerically that gyro has significant effect to reduce random vibration-induced disturbances especially for random vibration with amplitude less than 1σ which occurs in 68.8% of cases. Aside from reduction of the disturbances generated directly on the oscillator, Gyro monting by making the flexible connection between oscillator and PCB reduces the disturbances transfer from it significantly.

CONCLUSION
In this paper, a gyroscopic mounting proposed to install crystal oscillator of high dynamic GNSS receiver on it, in order to improve signal tracking process.

The analysis results of using this mounting in order to improve FLL tracking loop threshold as dominant loop of high dynamic receiver are summarized in table I and table II.

### TABLE I. SUMMARIZED RESULTS OF GYRO-MOUNTING EFFECTS ON STEADY STATE AND SINUSOIDAL VIBRATION-INDUCED FREQUENCY ERRORS.

<table>
<thead>
<tr>
<th>LOAD</th>
<th>Max frequency errors (Hz) F≈30°, φ=0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed oscillator</td>
</tr>
<tr>
<td>Steady State</td>
<td>6.61σ_{FLL}</td>
</tr>
<tr>
<td>Sinusoidal Vibration</td>
<td>2 when: ƒ=2</td>
</tr>
<tr>
<td></td>
<td>&amp; φ=1°</td>
</tr>
</tbody>
</table>

### TABLE II. SUMMARIZED RESULTS OF GYRO-MOUNTING EFFECTS ON MECHANICAL RANDOM VIBRATION-INDUCED FREQUENCY JITTER φ≈30°

<table>
<thead>
<tr>
<th>Amplitude of RV</th>
<th>Mechanical Random Vibration- induced frequency jitter (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max Gyro effect</td>
</tr>
<tr>
<td></td>
<td>ƒ≈90°, φ=90°</td>
</tr>
<tr>
<td>1σ (68%)</td>
<td>11.51σ_{FLL}</td>
</tr>
<tr>
<td>2σ (95%)</td>
<td>23.03σ_{FLL}</td>
</tr>
<tr>
<td>3σ (99.8%)</td>
<td>34.54σ_{FLL}</td>
</tr>
</tbody>
</table>

When random vibration applied on high dynamic GNSS receiver, the probability of signal loss depends to the magnitude and angular orientation of either random vibration or crystal g-sensitivity vector. When crystal oscillator in this receiver is installed on gyro, this mounting shows some positive effects or some drawbacks depend to angular orientation of random vibration and g-sensitivity vector of crystal. But totally gyro reduces the probability of signal loss on all over the high dynamic receiver trajectory where random vibration applied on it (Fig.12 and table II).

REFERENCES

Feasibility Study of Proximity Sensing by using a Conventional Airborne Transducer

Ken Yamada, and Shu Agatsuma
Department of Electronic Eng., Tohoku Gakuin Univ., Tagajo, Japan
E-mail: k-yamada@mail.tohoku-gakuin.ac.jp

Abstract— Ultrasonic proximity sensors utilizing non-radiant (evanescent) acoustic fields created in the vicinity of piezoelectric vibrators were proposed. When an object is brought into the evanescent field, electric admittance of the vibrator varies depending on the vibrator-to-object distance. In former reports, the air-film damping effect occurred between the sensing plate attached to the length-extensional mode vibrator and the test-object plate was studied. In this study, the sensing system is constructed by using a commercially-available airborne transducer. Distance dependent variation in the electric admittance level such as that observed in the former study has been confirmed.

Keywords—proximity sensing; airborne transducer; mono-morph-type bending plate

I. INTRODUCTION

Ultrasonic proximity sensors utilizing non-radiant (evanescent) acoustic fields created in the vicinity of piezoelectric vibrators were proposed [1]-[4]. When an object is brought into the evanescent field, electric admittance of the vibrator varies depending on the vibrator-to-object distance. In former reports [3],[4], the air-film damping effect occurred between the sensing plate attached to the length-extensional mode vibrator and the test-object plate was studied. In this configuration, lateral flow of the viscous fluid sandwiched in between the parallel plates undergoes resistive force that will act as a damper for the motion of vibration. Because the degree of damping varies depending on the gap width, it is applicable for detection of the gap width by observing the change in electrical properties of the piezoelectric vibrator. In this system, a 140° rotated Y-cut LiNbO₃ bar operating in the length-extensional mode at 80.9 kHz was employed for driving the sensor plate. Although it showed good sensing properties, it would be more convenient if the system of the same kind could be constructed using a conventional airborne transducer. In this study, the sensing system is constructed by using a commercially-available airborne transducer. Feasibility of using the conventional airborne transducer for proximity sensing will be discussed.

II. TRANSDUCER CONFIGURATION AND MEASUREMENT SETUP

The experimental setup for the proximity sensing is shown in Fig. 1. An acrylic plate of 5 mm thickness is approached to the face of the transducer using a pulse-motor stage. The transducer is a commercially-available one (R40-16, NIPPON CERAMIC, Japan) composed of a mono-morph-type bending plate with a radiation cone attached to it and housed in a metal case, such as shown in Fig. 2. The diameter and the height of the housing are 16.2 mm and 12.2 mm, respectively. The electric admittance characteristic around the main resonance is shown in Fig. 3. The resonance (operation) frequency is 48.9 kHz.

![Fig. 1 Measurement setup for proximity sensing by a conventional airborne transducer.](image1)

![Fig. 2 Cut view and picture of the utilized airborne transducer.](image2)
III. EXPERIMENTAL RESULTS

The variation in the peak value of the electric admittance at the resonance on the distance \( d \) between the transducer front surface and the target plate is observed by an impedance analyzer (IM3570, HIOKI E.E. Corp., Japan). The result in an extended distance range is shown in Fig. 4(a). Distance dependent variation in the electric admittance level such as that obtained in the former study [3],[4] has been confirmed. The periodical variation is due to the resonance between the transducer and the plate, and the spatial pitch corresponds to a half-wavelength in air (3.5 mm). Figure 4(b) shows the variation in the small distance range where \( d \) is less than 5 mm. Although the variation is not smooth enough, decrease in the \( |Y| \) value on the distance \( d \) is observed.

![Fig. 3 Admittance characteristic of the airborne transducer around the resonance (operation) frequency.](image)

![Fig. 4 Variation in \( |Y| \) value at resonance on the distance \( d \) in the extended range (a) and in the small distance range (b).](image)

![Fig. 5 Variation in \( |Y| \) value at resonance on the distance \( d \) in the extended range (a) and in the expanded small-distance range (b) when the target plate is tilted by 10° with respect to the transducer surface.](image)

IV. CONCLUSIONS

Feasibility of using a conventional airborne transducer for proximity sensing has been studied. Distance dependent variation in the electric admittance level at resonance has been confirmed. However, marked variation in the near region such as that obtained by using the air-film damping effect [3],[4] has not been observed. Although the dynamic range and the smoothness in the variation of the admittance level are not comparable to those obtained in the former study, the present system might be used as an easy way of proximity sensing.

References


Interrogation of Orthogonal Frequency Coded SAW Sensors Using the USRP

Department of Electrical Engineering and Computer Science
University of Central Florida, Orlando, FL. 32816-2450
E-Mail: James.Humphries@knights.ucf.edu

Abstract—The universal software radio peripheral (USRP) is a versatile software defined radio (SDR) platform, developed by Ettus Research™, which is intended for a wide variety of applications ranging from communication links to RADAR. We have investigated another application of the USRP by implementing a transceiver capable of interrogating passive, wireless surface acoustic wave (SAW) sensors centered at 915MHz. Interrogation of wideband orthogonal frequency coded (OFC) SAW sensors imposes strict requirements on the timing and synchronization of the transceiver. In the standard mode of operation, samples are generated and streamed between the USRP and host computer, introducing latency and bandwidth limitations due to the sampling bus. To achieve the performance required for this application, the USRP FPGA has been modified to introduce new functionality. Extraction of the sensor temperature is accomplished with a custom matched filter correlator. The system is capable of interrogating multiple sensors and can quickly reconfigure the USRP. Demonstration of the USRP wireless sensor system is achieved by interrogating wireless SAW OFC sensors at 915MHz and extracting the sensor temperature.

I. INTRODUCTION

Interrogation of passive, wireless sensor tags pose many unique challenges, particularly in a research environment. As sensor designs evolve, so too must the wireless sensor interrogation system (passive tag reader). Typical interrogator design requires many discrete components, many of which must be replaced or reconfigured manually to be compatible with the sensor specifications. One solution is to utilize a software defined radio (SDR) platform to implement an interrogation system. The universal software radio peripheral (USRP), developed by Ettus Research™, is an SDR that is available as a commercial off-the-shelf (COTS) unit. The USRP has found a variety of applications ranging from RADAR to GSM base stations [1]–[3]. The USRP covers the frequency bands of interest for SAW sensors (50MHz - 3GHz) and can be rapidly reconfigured by programming the gain, center frequency, and bandwidth (among other specifications) in software, rather than modifying hardware. This SDR solution eliminates many challenges encountered with custom radio interrogator design.

This work is an enhancement upon previous work using the USRP N200 as a prototype system to interrogate SAW sensors [4]. Although the N200 showed promising results, the limited bandwidth made wideband SAW orthogonal frequency coded (OFC) sensor design difficult. Additionally, a software backend for rapid data capture and temperature extraction had not been developed for the USRP. This paper presents results using the USRP B200. Improvements have been made to both the FPGA design and host computer software that allows the full 56MHz bandwidth to be utilized. A software correlation package, previously developed at the University of Central Florida (UCF), has been ported from MATLAB to Python, allowing the temperature of the wireless SAW sensors to be extracted from data obtained from the USRP. Two sensors are interrogated simultaneously by the USRP and the temperature for each sensor is extracted from the received data.

II. USRP & EXPERIMENTAL SETUP

This section documents the USRP and host computer setup used for these experiments. Many configurations of USRP and host computer are possible; this section serves to document a successful setup for the USRP platform.

A. USRP

The USRP chosen for this experiment is the B200. The B200 is a fully integrated SDR, utilizing the Analog AD9361 direct conversion transceiver and Xilinx Spartan 6 FPGA. The B200 frontend frequency coverage (70MHz - 6GHz) is ideal for the sensors designed in this experiment at 915MHz. The frontend maximum transmitter output power is +10dBm. The B200 also has a maximum real-time bandwidth of 56MHz, which is improved from the USRP model used in previous efforts. The streaming bus is USB3.0, enabling rapid transfer of samples between the USRP and a host computer.

B. Host Computer & Software

Software radio performance can be limited by the ability of the host computer to stream and process samples. The typical mode of operation continuously streams data from the SDR to a host computer. While narrowband applications introduce little overhead for the streaming bus, high bandwidth applications, such as this experiment with wideband SAW sensors, may be difficult with the amount of bandwidth needed. Slower processing speed, for example, will introduce a maximum limit on the amount of data (sample rate) that can be streamed from the USRP in real-time. This section documents the computer and software used during these experiments, as performance may be affected by different setups. Table I lists the host computer hardware and software used.
III. FPGA MODIFICATIONS

In order to relax the USB streaming requirements of the USRP for this high bandwidth application, modifications were made to the stock FPGA design to implement custom functionality. The stock FPGA code is freely available online so that custom DSP modules can be implemented on the FPGA, while including the stock DSP [5]. This section details the modification made to the FPGA to maximize the system bandwidth, synchronize the transmit and receive chains, and improve received data integrity (eliminate dropped samples over USB due to the real-time streaming bottleneck).

A. Overview of FPGA Modifications

Both the transmit and receive sections of the FPGA were modified with custom modules. On the transmit chain, an interrogation signal generator was implemented that replaces the stock transmitter DSP. The signal generator outputs a linear frequency sweep (chirp) with a desired bandwidth and time length. On the receive chain, a receive buffer was implemented that stores samples for a specified duration after the transmit signal has finished. The samples are buffered in RAM at the full 56MHz sample rate and then read out of RAM at a rate much slower than the sampling frequency so that the USB bus does not become saturated and drop samples. This removes the limitations imposed by the real-time streaming operation of the USRP, which is not required for this application. 512 I/Q samples are stored for each interrogation cycle which corresponds to 9.15μs per data sweep. The FPGA only buffers a single data set of 512 samples for a given interrogation cycle. The received waveform samples are output to the host PC for processing and extraction of sensor data. A simplified block diagram illustrates the FPGA layout in Fig. 1. A timing diagram is also given in Fig. 2 to detail the interrogation cycle of the FPGA.

B. Transmit Chain FPGA Modifications

Transmit sample streaming between the USRP and host computer has been eliminated by implementing an interrogation signal generator on the FPGA. This greatly reduces the host computer processing overhead by eliminating its responsibility to generate and stream transmit samples. All of the stock transmit DSP algorithms on the FPGA were replaced with a module that outputs predefined samples to yield a desired frequency spectrum and interrogation time length. The module implemented outputs a linear FM chirp with time length of 1μs and bandwidth of 35MHz. The transmit module is in a waiting state until the receive module indicates that the receive process has finished (Flag: rx_done = 1), upon which the interrogation cycle will begin. A state diagram of the interrogation signal generation module is shown in Fig. 3. A plot comparing the predicted spectrum versus the spectrum output from the USRP is shown in Fig. 4.

The interrogation signal samples are calculated in MATLAB and then programmed directly into the FPGA using two’s compliment binary format. Both the In-Phase (I) and Quadrature (Q) samples are generated and output in parallel from the interrogation module. The parallel samples are input to a double data rate (DDR) module which serializes the data (I on positive clock edge and Q on negative clock edge). The AD9361 is configured by the USRP onboard microcontroller to accept data in a serial format.

C. Receive Chain FPGA Modifications

An FPGA receive module was implemented that synchronizes with the transmit state, defines a listen window, and buffers a single interrogation sweep in FPGA memory (RAM). The receiver module is active when the transmit module indicates that the interrogation pulse has finished (Flag: tx_done = 1). The module is a state machine that cycles through states as the receiver receives data. A state diagram of the module is given in Fig. 5. A list of the states and a description of each state is given in Table II as well. The receiver saving a total of 512 samples in the RAM sample buffer (9.15μs listen window at 56MHz sampling rate). Once the sample buffer is full, the data points are read from memory to the host computer at a rate of 2MHz. Fig. 6 is an oscilloscope measurement showing (a) a single interrogation pulse and (b)
multiple interrogation pulses measuring the pulse repetition interval (PRI). The PRI is controlled by the transmit sample strobe (not used for transmit in this design) and can be set by the user depending on the host computer performance. The default operation would stream samples at the user specified sampling rate, here a 56MHz rate would be required. The host computer can easily keep up with this reduced rate of data transfer, eliminating corrupted sweep data.

IV. HOST SOFTWARE & POST PROCESSING

Software running on the host computer has been implemented in Python to serve multiple purposes. First, the software creates an interface between the host computer and the USRP using the USRP Hardware Drivers (UHD) and GNU Radio. This interface programs the USRP (center frequency, gain, bandwidth, etc) and collects the sensor response data. Second, a matched filter correlator (previously developed at the
University of Central Florida (UCF) using MATLAB) has been ported to Python to extract the SAW sensor information, in this case extracting temperature. A block diagram showing the relationship between each portion of the software is illustrated in Fig. 7.

The interface between the USRP and host computer is made using the USRP Hardware Drivers (UHD) and GNU Radio. The USRP Source/Sink blocks are used in GNU Radio to program the USRP (center frequency, clock rate, Tx/Rx gain, and input/output transceiver ports). The main program asks for GNU Radio to store a specified number of samples and then stop the flowgraph (HEAD block). The received data is stored in a memory vector (VECTOR SINK block) that can be accessed by the main program to extract temperature.

A software matched filter correlator has been previously developed at UCF using MATLAB [6]. To achieve the maximum performance with the USRP, this software has been ported to Python. This allows a single program to directly access data from the USRP and process that data to extract temperature (other measurements such as strain and gas detection are planned for the Python port). A set of matched filters is generated for each sensor in the system, where each matched filter corresponds to a temperature (the user specifies the minimum and maximum temperatures as well as the temperature resolution). The matched filters are scaled using a frequency scaling factor (FSF), based on calibrated sensor parameters. The received sensor signal is correlated with each matched filter. The matched filter which yields the maximum correlation peaks corresponds to the temperature of the current sensor being processed.

V. INTERROGATION OF SAW TEMPERATURE SENSORS

Integration of the USRP, FPGA modifications, and correlation software formed a complete interrogation system. A single wireless SAW sensor is tested first with a reference thermocouple to assess the temperature system accuracy. Testing of the full system is also discussed by interrogating multiple SAW sensors simultaneously and extracted each sensor’s temperature. The SAW sensor temperature precision is also considered by demonstrating averaging of multiple data sweeps to improve the signal-to-noise ratio (SNR). First, a small array of SAW sensors were designed that satisfied the modified USRP operational parameters (bandwidth and listen window).

A. SAW Sensor Design

A set of SAW sensors was designed specifically for the USRP system to fit within the system bandwidth and listen window. The devices are fabricated on YZ-LiNbO3. The SAW devices are centered at 915MHz and have bandwidths less than the interrogation signal bandwidth (35MHz) to allow some margin for the center frequency to drift with temperature.

Two SAW devices have been fabricated for these experiments. Orthogonal frequency coding (OFC) is used to identify each sensor [7]. The first device, labeled usrp-m1-d1, has an initial delay time of 2µs, 3 OFC chips, and a bandwidth of 30MHz. The second device, labeled usrp-m1-d3, has an initial delay time of 3µs, 5 OFC chips, and a bandwidth of 26MHz. The usrp-m1-d3 device is compatible with the 915MHz ISM band. The usrp-m1-d1 device uses a standard Bragg reflector design for each of the three OFC chips. The usrp-m1-d3 device, on the other hand, utilizes a withdrawal weighted reflector design to optimize the device response [8]. This reduces distortion of the time responses due to intra-electrode ringing, which are typical in long Bragg reflectors and take longer to decay in time [9]. This design trades reduced time domain distortion for longer chip lengths to make up the device code; effectively reducing the amount devices that can be placed in a given time window. A time domain plot of each sensor is shown in Fig. 8. The double transit of usrp-m1-d1 is visible in the time domain plot beginning at ≈ 4µs. Future device designs will need to suppress this double transit as it can cause interference with other sensor responses, which
reduces the SNR [10]. Techniques for reducing the multi-transit response include unidirectional or withdrawal weighted transducer structures [11], [12].

B. Temperature Measurements

Two antennas were used, one each for the transmit and receive ports of the USRP. Each Tx/Rx antenna is a folded dipole with approximately 2dBi of gain. Each sensor was also attached to a folded dipole antenna of approximately 2dBi gain. A photograph of the sensor and antenna is shown in Fig. 9. The sensors were interrogated by the USRP transceiver and the temperature was extracted by processing the received data with the software correlator on the host computer.

The first test was a single wireless SAW sensor with a wired reference thermocouple. The sensors were cooled (cold nitrogen gas) and heated (electric heat gun) randomly for a few minutes. After heating and cooling, the temperatures were left to return to room temperature. The measured temperatures of both the SAW sensor and thermocouple are shown in Fig. 10. For the second test, both SAW sensors (usrp-m1-d1 and usrp-m1-d3) were placed in the interrogator field-of-view. Again, the sensors were heated and cooled randomly and then returned to room temperature. The extracted temperatures for both sensors during the test are shown in Fig. 11.

C. Temperature Extraction Precision

The extracted temperature precision of the SAW sensor can be improved by averaging multiple data sweeps before performing the correlation routine to extract temperature. This is due to the SNR improvement by averaging out random noise from the environment and system. A high SNR (>20dB) has been shown in recent work to give a standard deviation of

![Graph](image_url)
measurements better than 0.5°C by using a VNA [13]. A single sensor was placed at a wireless range of 15cm (6 inches) from the USRP antennas and the temperature was held constant. The USRP TX gain was set to 70dB (-10dBm peak output power) and RX gain was set to 30dB. 100 measurements were taken for averages varying between 1 and 1000. Fig. 12 shows the extracted temperature for these measurements as the number of averages was increased. The standard deviation of each test was also calculated and is shown in Fig. 13. The system is capable of receiving 1000 data sweeps, averaging them, and extracting the temperature in approximately 1 second.

VI. CONCLUSION

This paper presented the application of the USRP software radio as a wireless SAW sensor interrogator. The USRP FPGA has been modified to introduce new functionality and unlock the full bandwidth potential of the SDR. The modified USRP is capable of simultaneously interrogating multiple passive, wireless SAW OFC sensors and extracting temperature during post-processing on the host computer. It has been demonstrated that the SNR and extracted temperature precision is improved using the USRP as a coherent integration receiver with the developed correlator parameter extraction software. This design can be applied to other USRP models for higher performance (greater bandwidth or output power) applications or for applications requiring a stand-alone (embedded platform) interrogator. The USRP B200 has been demonstrated as an efficient interrogation system for passive, wireless SAW OFC sensors.

REFERENCES

SH-SAW–Based Sensor for Heavy Metal Ion Detection

Zeinab Ramshani, Binu B. Narakathu, Avuthu S. G. Reddy, Massood Z. Atashbar
Department of Electrical and Computer Engineering
Western Michigan University
Michigan, USA
zeinab.ramshani@wmich.edu

Jared T. Wabeke, Sherine O. Obare
Department of Chemistry
Western Michigan University
Michigan, USA

Abstract— In this study, a shear horizontal mode surface acoustic wave (SH-SAW) sensor, was designed and fabricated for the detection of heavy metals. The SH-SAW sensor was photolithographically fabricated by patterning gold (Au) interdigitated electrodes (IDE) and reflectors on the surface of a 64° YX-LiNbO₃ based piezoelectric substrate. A flow cell, with a reservoir volume of 3 µl, which employs inlet and outlet ports for the microfluidic chamber and polydimethylsiloxane (PDMS) based microfluidic channels, was also designed and fabricated using acrylic material. Phenol and naphtho[2,3-a]dipyrido[3,2-h:2′3′-f] phenazine-5,18-dione (QDPPZ) were employed as the sensing layers for mercury nitrate (Hg(NO₃)₂) and nickel nitrate (Ni(NO₃)₂).

Keywords—microfluidics; mercury; nickel; SH-SAW sensor

I. INTRODUCTION

Heavy metal detection has been receiving an increasing concern due to its adverse effects in aquatic and atmospheric ecosystems, even at the micro- or nano-molar concentration levels [1,2]. In case of aquatic systems, this eco- and phytotoxicity on vital microorganisms is influenced by environmental factors [3]. Among the various type of heavy metals, mercury (Hg) is the second most toxic heavy metal in the planet and nickel (Ni) has been seeing a rising demand in the surface coating and jewelry industries. Both these elements are often released from natural and manmade sources, and have require complex labeling methods. Therefore, it is necessary to develop sensing systems with easy-to-use, cost effective, highly sensitive and rapid detection techniques. Surface acoustic wave (SAW) sensors, which have been employed in various sensing applications, are an optimum choice for the detection of Hg²⁺ and Ni²⁺ due to its relatively small size, high resonant frequency, low power consumption and compatibility with CMOS technology [15,16]. Moreover, SAW sensors operating in the shear horizontal mode cause parallel particle displacement, which has a minimal energy loss in aqueous media [17,18]. Therefore, the development of detection systems, which incorporate shear horizontal surface acoustic wave (SH-SAW) sensors, for the highly sensitive detection of toxic heavy metals, such as Hg and Ni, in aqueous media is of utmost importance.

In this work, a 64° YX-LiNbO₃ based piezoelectric substrate with gold (Au) interdigitated electrodes (IDE) and reflectors was used to fabricate a SH-SAW sensor using photolithography techniques. The sensor was used along with a polydimethylsiloxane (PDMS) based microfluidic channel mounted in an acrylic flow cell. Phenol and naphtho[2,3-a]dipyrido[3,2-h:2′3′-f] phenazine-5,18-dione (QDPPZ) were used as sensitive layers for the chemical absorption of Hg²⁺ and Ni²⁺. The capability of the SH-SAW sensor for picomolar level detection of mercury (II) nitrate (Hg(NO₃)₂) and nickel (II) nitrate (Ni(NO₃)₂) was investigated by monitoring the SH-SAW sensor's frequency response (S21).

II. THEORY

Surface acoustic wave devices work based on piezoelectricity. Each fabricated SAW sensor has a particular resonant frequency (f₀), which is dependent on the width of the metalized IDEs (w) and SAW propagation velocity on the piezoelectric substrate (vₛ). and is mathematically given as Eq. (1) [19]

\[
f₀ = \frac{vₛ}{4w}
\]  

The electrical signal is applied to the input IDEs deposited on the surface of the piezoelectric substrate. The electrical signal transforms to a surface acoustic wave and travels through the piezoelectric substrate to reach the output IDEs and then converts back to an electrical signal. The absorbance of...
the different concentrations of Hg(NO$_3$)$_2$ and Ni(NO$_3$)$_2$ by the sensitive layers, which are coated on top of the piezoelectric substrate, causes mass loading effect to take place. This changes the SAW propagation velocity and thus the resultant center frequency [20]. This phenomenon can be used for the detection of the various analyte concentrations.

III. EXPERIMENTAL

A. Chemicals and Sample Preparation

1,10-phenanthroline, potassium bromide, ethanol, methylene chloride and conc. sulfuric acid were purchased from the Sigma Aldrich Company. 1,2-diaminoantraquinone was purchased from Alfa Aesar. All solvents were reagent grade and used without further purification. Hg(NO$_3$)$_2$ and Ni(NO$_3$)$_2$, in crystalline form (purchased from the Sigma Aldrich Company) were used to prepare 1 pM, 100 pM, 1 nM, 100 nM and 1 $\mu$M concentrations of Hg(NO$_3$)$_2$ and Ni(NO$_3$)$_2$ by mixing with deionized (DI) water. All test analytes were stored at $-20$ ºC in 10 ml aliquots before use. Tubing (inner diameter - 0.01”; outer diameter – 0.0625”) and tube connection accessories (from Upchurch Scientific) were used for the sample transfer in the flow cell.

B. SH-SAW Sensor Fabrication

The SH-SAW sensor was photolithographically fabricated by patterning the Au IDEs on the surface of a 64º YX-LiNbO$_3$ based piezoelectric substrate. It consisted of eight pairs of input and output IDEs, 120 reflectors on the outer side of input and output IDEs as well as 20 reflectors in between the IDEs. Metal sputtering technique was used to deposit the 0.1 $\mu$m thick Au IDEs with electrode width and gap of 10 $\mu$m, for a two port resonator SAW sensor configuration. With a designed wavelength of 40 $\mu$m and SAW propagation velocity of 4474 m/s on the 64º YX-LiNbO$_3$ piezoelectric substrate, the resultant resonant frequency will be 111.8 MHz and the measured resonant frequency of the fabricated sensor was 112 MHz [21]. The photograph of the fabricated SH-SAW sensor, with overall device dimensions of 11 x 12 mm$^2$ is shown in Fig. 1.

C. Synthesis of the Chemical Sensing Layer

The chemical sensor QDPPZ was synthesized using a two step process. In the first step, concentrated sulfuric acid (H$_2$SO$_4$, 20 mL) and concentrated nitric acid (HNO$_3$, 10 mL) were added dropwise to a mixture of 1,10-phenanthrolione (1.00 g, 5.56 mmol) in the presence of potassium bromide (KBr) (5.95 g, 50 mmol) at 0 ºC. The solution was refluxed for 2 hours and then cooled to room temperature, yielding a brown, oily product. The contents of the flask were diluted with 400 mL of deionized water and neutralized with sodium bicarbonate (NaHCO$_3$), yielding a clear yellow solution. The product was extracted with methylene chloride and dried over anhydrous MgSO$_4$. The solvents were removed using a rotary evaporator, resulting in a yellow solid. The product was purified by recrystallization from methanol. The average yield of the product, 1,10-phenanthrolione-5,6-dione was (1.11 g, 5.31 mmol) which was calculated to be 95%. In the second step, 1,10-phenanthrolione-5,6-dione (0.50 g, 2.38 mmol) was refluxed in ethanol for 15 min. 1,2-diaminoantraquinone (0.981 g, 2.38 mmol) was then added, resulting in a purple solution, and the solution was refluxed for 4 hrs. The dark purple product was collected using vacuum filtration, washed with methanol and concentrated in vacuum. The reaction yield was 80%. The products were characterized by $^1$H and $^{13}$C NMR and found to be pure. Phenol was deposited on top of the sensor surface as the sensitive layer for Hg$^{2+}$ detection and QDPPZ for Ni$^{2+}$ [22-25].

D. Microfluidic Flow Cell Fabrication

An acrylic flow cell was designed in AutoCAD™ and CNC machined, with overall device dimension of $70 \times 50 \times 52$ mm (w/l/h), (Fig. 2). PDMS was used to fabricate a microfluidic flow channel, with dimensions of $710 \times 6800 \times 710$ $\mu$m (w/l/h) and a total channel volume of $\approx 3.4$ $\mu$L. Two sets of inlet and outlet ports were integrated into the flow cell for aiding the flow of the test analyte through the PDMS microfluidic channel. An axially magnetized set of Neodymium magnets (0.25" diameter; 0.375" thickness; 13,200 gauss magnetic strength) from K&J Magnets, Inc was used for the effective closing of the flow cell, which results in the tight sealing of the PDMS microfluidic flow channel around the sensing area of the SH-SAW sensor. Two sets of epoxy potted spring probes with SubMiniature version A (SMA) cables were mounted into the flow cell to probe electric signals from the SAW devices. Each set of spring probes consisted of four pins with spacing of 2.5 mm.

Fig. 1. Photolithographically fabricated shear horizontal mode surface acoustic wave (SH-SAW) sensor.
E. Experiment Setup

The experiment setup is shown in Fig. 3. The SH-SAW sensor was placed in the sensor groove of the flow cell. Calibration for the wires and probes was done before the measurements. Before use and at the end of each experiment, the sensor was cleaned with acetone, and then blow-dried with pressurized air. The measurements were performed at a constant room temperature (25 °C) since any changes in the temperature will affect the SAW velocity and attenuation. Initially, the reference signal was established using DI water. Then, varying concentrations of Hg(NO$_3$)$_2$ and Ni(NO$_3$)$_2$ (1 pM, 100 pM, 1 nM, 100 nM and 1 µM), were injected into the flow cell at a flow rate of 10 µl/min using a KD Scientific (KDS210P) programmable syringe pump. The phase response (S$_{21}$) of the SH-SAW sensor was monitored using an Agilent 4395A network analyzer. System control, data acquisition and post processing of the network analyzer measurements were performed using a LabView™ based application.

IV. RESULTS AND DISCUSSIONS

The phase response (S$_{21}$) of the SH-SAW sensor was first tested towards different concentrations of Hg(NO$_3$)$_2$. Fig. 4 shows the SH-SAW sensor resonant frequency shift for the varying concentrations of Hg(NO$_3$)$_2$ solution, when compared with DI water. Frequency shift of 184.83±23 kHz, 378.33±62 kHz, 458.33±80 kHz, 600.00±96 kHz and 748.33±116 kHz was observed for the 1 pM, 100 pM, 1 nM, 100 nM and 1 µM concentrations of Hg(NO$_3$)$_2$ solution, respectively.

Then, varying concentrations of Ni(NO$_3$)$_2$ was injected onto the SH-SAW sensor and the phase response (S$_{21}$) of the SH-SAW sensor was monitored. Fig. 5 shows the SH-SAW sensor resonant frequency shift for the varying concentrations of Ni(NO$_3$)$_2$ solution, when compared with DI water. Frequency shift of 273.33±60 kHz, 441.66±50 kHz, 566.66±49 kHz, 673.33±43 kHz and 806.66±57 kHz was observed for the 1 pM, 100 pM, 1 nM, 100 nM and 1 µM concentrations of Ni(NO$_3$)$_2$ solution, respectively.

The results obtained demonstrated that phenol and QDPPZ can be employed as sensitive layers for Hg(NO$_3$)$_2$ and Ni(NO$_3$)$_2$, respectively. The chemical absorption phenomenon, provided by these sensitive layers, allow the sensing system to detect the analytes concentration variations due to the mass loading effect in the SH-SAW sensor, which results in the center frequency shift. Therefore, the SH-SAW sensor is capable of detecting pico molar level concentrations of Hg(NO$_3$)$_2$ and Ni(NO$_3$)$_2$, while the United States Environmental Protection Agency (EPA) maximum allowable level for Hg(NO$_3$)$_2$ and Ni(NO$_3$)$_2$ in drinking water is reported to be 10 nM and 0.3 µM, respectively [26,27].
Fig. 5. Frequency shift in SH-SAW sensor phase response (S11) towards varying concentrations of Ni(NO3)2.

V. CONCLUSION

In this project, a SH-SAW sensor was successfully designed and fabricated, with Au IDEs, on a 64° YX-LiNbO3 piezoelectric substrate. An acrylic flow cell, which employs PDMS based microfluidic channels was also fabricated. The feasibility of using the fabricated device for detecting varying concentrations of Hg(NO3)2 and Ni(NO3)2 were investigated by studying the phase response (S11) of the SH-SAW sensor. The measured phase response (S11) of the SH-SAW sensor demonstrated a 184 kHz and 273 kHz frequency shift for the measured phase response (S21) of the SH-SAW sensor mounted on a PCB is part of our future work.

REFERENCES


The Study of Beidou Timing Receiver Delay Calibration

Hongbo Wang*† Hang Yi*† Shengkang Zhang*† Haifeng Wang*† Fan Shi*† Huaiying Shang*† Yujie Yang† Jun Ge† and Zhiqi Li‡

*Beijing Institute of Radio Metrology and Measurement, Beijing China 100854
†Science and Technology on Metrology and Calibration Laboratory, Beijing China 100854
‡Xidian University, Xi’an China 710071

Abstract—This paper gives a test method of Beidou timing receiver delay. A Beidou System simulator and a time interval counter were used in the experiments. During the process, the most important step is to calibrate the delay of the simulator. The uncertainty of this method is analyzed to be less than 1.5ns. Using this method, some typical Chinese commercial timing receivers were tested, and the results are shown in figures. The receiver delay variations with temperature have been studied.

Keywords—Beidou; Receiver; Calibration; Test

I. INTRODUCTION

Beidou Satellite Navigation System (BDS) has been in service for Asia-Pacific Region since December 2012. The role of BDS is becoming more and more important in China, especially in military applications and government security departments. Timing is an important application for BDS. Since Chinese receiver technology has been developed for only a few years, as a part of many vital systems, the domestic BDS receivers need to be evaluated and calibrated. This paper gives the method of BDS timing receiver delay calibration. After the calibration, BDS receiver can generate accurate time. Using this method, some typical Chinese commercial BDS receivers were tested.

The performance of BDS timing receiver is affected by many factors, including orbit error, satellite clock error, atmospheric (troposphere and ionosphere) delay, antenna phase center, multipath, receiver thermal noise, etc. The method in this paper only focuses on the internal delay calibration of the BDS timing receiver.

The GPS timing receivers at many UTC laboratories all over the world are mostly calibrated by comparing with a standard receiver delivered from International Bureau of Weights and Measures (BIPM). Using this method, the uncertainty still remains at 5ns [1]. But in China, until now, the standard BDS receiver has not been put forward. Like the absolute calibration of GPS receiver [2] [3], a BDS simulator is used to calibrate the BDS timing receiver.

II. BDS RECEIVER CALIBRATION

Supposing the BDS simulator’s internal delay is already known, the internal delay of the BDS receiver could be measured, as Fig.1 shows. The 10MHz from a cesium atomic clock is input to the BDS simulator and the time interval counter as external reference. The BDS simulator outputs the simulated BDS satellite signal to the receiver directly, simultaneously generates 1PPS signal to the time interval counter. The BDS timing receiver is set to operate in timing mode, and outputs its 1PPS signal to the time interval counter. Recording the time intervals between the two 1PPS signals, remove the difference of the two 1PPS cable delays and the internal delay of the simulator, the delay of the BDS receiver under test could be gained.

According to GNSS observation equation, the pseudorange PR could be expressed as

\[ PR = R + c (\Delta t_{rx} - \Delta t_{sat} + b_{rx} + b_{sat} + TtC + n) + r_{iono} + r_{trop} + r_m \]

where:
- \( R \) is the real range,
- \( \Delta t_{rx} \) is the delay of receiver clock,
- \( \Delta t_{sat} \) is the delay of satellite clock,
- \( b_{rx} \) is the receiver delay,
- \( b_{sat} \) is the satellite delay,
- \( TtC \) is 1PPS signal tick to code offset of the simulator,
- \( n \) is the noise,
- \( r_{iono} \) is the ionospheric delay,
- \( r_{trop} \) is the tropospheric delay,
- \( r_m \) is the multipath.

Fig. 1. The diagram of BDS timing receiver calibration with BDS simulator

978-1-4799-8866-2/15/$31.00 ©2015 IEEE
During the test procedure, the simulator is set to operate in a static scenario, excluding all the error sources. In order to simplify calculation, when all the errors from the simulator are set to zero, the delay of receiver and additional noise can be computed with the following equation:

$$bias_{rx} + n = \frac{PR - R}{c} - TtC$$  \hspace{1cm} (2)

Generally, the pseudorange $PR$ could be recorded by PC from the receiver and the real range is set from the simulator. But for some simple receiver, the pseudorange message could not be output, so the first item on the right in (2), can be measured with time interval counter. The key step is to get the $TtC$ value, which is the internal delay of the BDS simulator.

III. BDS SIMULATOR CALIBRATION

In order to calibrate the BDS receiver, the internal delay of the BDS simulator itself has to be calibrated first.

The BDS receiver calibration bench is equipped with a BDS signal simulator NSS8000, developed and manufactured by Satellite Navigation Research Center in National University of Defense Technology. This simulator supports all the BDS signals, besides GPS/GLONASS/GALILEO systems, can supply 18 channels for BDS. Users can choose any GNSS system, frequency, satellite signal and set any simulated scenario. It shows a good performance, the output signals are precise and stable.

Fig. 2 gives the connection relationship of the simulator calibration. Input the 10MHz reference frequency of a cesium atomic clock to the BDS simulator and high-speed sampling oscilloscope. Set the simulator in high power output mode and operate in the “zero value” scenario, so that the signal propagation range is zero and without any error sources. Let the simulator output a single channel BPSK signal in B1 frequency. Both the BDS signal and the 1PPS signal generated from the BDS simulator are fixed in the phase to the 10MHz reference frequency. Use high-speed sampling oscilloscope (Tektronix MSO70804C, 25GS/s) to capture and store the offset between the code phase zero-crossing point of the BDS signal and the 1PPS signal.

It is inadequate to directly give the offset as the calibration value of the simulator, since we need to verify whether the observed zero-crossing point is the start of the code. Many zero-crossing points could be seen from the oscilloscope, the data from the simulator is consecutive, so the beginning of the code should be found out. Therefore, the BDS code and the 1PPS signal are recorded with the oscilloscope over a period of several code lengths. A Matlab program is developed to generate the same code and makes correlation with the stored code. The correlation peak marks the start of the code. The offset between this correlation peak and the recorded 1PPS signal represents the simulator 1PPS to code offset, including the cable and connector delay.

After one channel is calibrated, change another channel to do the same work, until every channel is calibrated. During the channel change, a small random offset may occur. Therefore, the calibration has to repeat many times for each channel, take an average value as the internal delay of the BDS simulator.

IV. BDS RECEIVER CALIBRATION RESULT

The same static scenario is used for each receiver, with a standard BDS constellation including five geostationary satellites, five inclined geostationary satellites and four medium earth orbit satellites. Four Chinese commercial BDS receivers were calibrated. Since the receivers under test are still developing, the manufacturers and the type of the receivers here are anonymous. The receiver internal delays are shown in Table I.

<table>
<thead>
<tr>
<th>Receiver Index</th>
<th>GNSS System and Frequency</th>
<th>Calibration Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GPS L1/L2 BDS B1/B2</td>
<td>486.7 ± 2.3</td>
</tr>
<tr>
<td>2</td>
<td>GPS L1/L2 BDS B1/B2</td>
<td>-444.0 ± 17.8</td>
</tr>
<tr>
<td>3</td>
<td>GPS L1 BDS B1</td>
<td>447.7 ± 11.1</td>
</tr>
</tbody>
</table>

Fig. 3. The BDS simulator delay test result by oscilloscope
These receivers are not only for BDS, they could also receive GPS signals. But during the calibration, the simulator only transmitted BDS signal, so the result can reflect their BDS timing performances. The delay of all the tested receivers could be adjusted by command with resolution of 100ns, and the results are the factory default values. Fig. 4 shows the delay data of the four receivers.

Table II lists all the error sources in the calibration and corresponding estimated values. The total uncertainty is 1.3ns.

VI. DELAY VARIATION WITH TEMPERATURE EXPERIMENT

Generally, the delay of a receiver will vary with the temperature [9]. With the BDS simulator, the delay variation with temperature has been studied.

The test process is the same as the BDS receiver delay calibration, except put the receiver into a climate temperature chamber, as Fig. 5 shows.

The first BDS receiver of the four is used to make the experiment. Set the climate temperature chamber in different temperatures, from -40°C to +60°C. Every 10°C make a measurement with the same one hour static scenario from the simulator, so that there are 11 temperature points in total. First, let chamber temperature decrease to -40°C from the room temperature. Once the temperature is stable on -40°C, make the one hour test. Then increase the chamber temperature by 10°C till +60°C.
During the entire test, the temperature of the lab is stable on 22°C. The cable in the temperature chamber is less than 20cm, so that its delay variation can be neglected.

Fig. 6 shows the result of the temperature experiment. As the chamber temperature increases from -40°C to +60°C, the internal delay of the BDS receiver increases about 2.0ns. A linear fit to the data in Fig. 6 yields a coefficient of 14ps/°C.

![Fig. 6. The four BDS receiver calibration results](image)

VII. CONCLUSION

Before the calibration of BDS receiver, several GPS receivers are calibrated in our laboratory, including Javad TRE-G3T, Novatel OEMV-III and Septentrio PolaRx2. These GPS receivers show very good precision and stability. Comparing with these GPS receivers, the Chinese BDS receivers have a lot to improve. This is the first time the BDS receivers are calibrated in our laboratory. The uncertainty of calibration is estimated to 1.3ns with the simulator. Using this method, the dependence of receiver delay on temperature is studied.

Because not all the calibration laboratories have BDS simulator, it is recommended to put forward a standard BDS receiver like BIPM. Comparing with the standard receiver will facilitate the calibration work. However, the BDS antenna and cable calibration is not contained in our research. As they are necessary components of a BDS timing system, this work should be done in the future.

We thank the technical team from Satellite Navigation Research Center who support us for the operation of the BDS simulator NSS8000.

REFERENCES

Developing of one time link calibrator with GNSS at NIM

LIANG Kun, ZHANG Aimin, YANG Zhiqiang, WANG Weibo
Division of Time and Frequency Metrology
National Institute of Metrology(NIM)
Beijing, China
liangk@nim.ac.cn

YANG Hang
School of Electronics and Information Engineering
Beijing JiaoTong University
Beijing, China

Abstract—Since 2014 NIM has also started to design and develop one kind of homemade calibration system for the time link calibration on the basis of NIMTFGNSS-2 receiver. The system has been constructed preliminarily and calibrated, and in the near future, with this calibrator we might implement the calibration campaigns of some TAI links under the APMP(Asia Pacific Metrology Programme) scheme.

Keywords—time transfer; GNSS; time link calibration

I. INTRODUCTION

Time link calibration is the premise of time transfer. At present, we can have two kinds of time links in TAI(Tempes Atomique International) corporation in summary, such as, GNSS(Global Navigation Satellite System) based links and TWSTFT(Two Way Satellite Time and Frequency Transfer) based links. GNSS based links can be calibrated using the methods of differential calibration (golden receiver method)[1] with the uncertainty of about 5 ns and using the whole link calibration with uncertainty of less than 2 ns, even 1 ns. Since 2001, BIPM has always implemented many GNSS time link calibration campaigns all over the world using the first method mostly because the method is easier to use. TWSTFT based links can be calibrated using two way mobile station with the uncertainty of about 1 ns, however the mobile station is hard to get for use of the calibration and the calibrated GNSS based links can be used for the alignment and calibration of TWSTFT based links and calibration for GNSS based links could be transferred to TWSTFT based links.

Base on the whole link calibration for GNSS based links, some successful calibration campaigns have been implemented by BIPM(Bureau International des Poids et Mesures), ROA(Royal Institute and Observatory of the Armada, Spain) PTB(Physikalisch-Technische Bundesanstalt, Germany), NIM(National Institute of Metrology, China) and LNE-SYRTE(OP, Observatory of Paris, France) and so on. PTB and BIPM has developed the homemade precise mobile calibration setups for the time link calibration as described in [2] and [3]. Moreover, BIPM has started to draw up the new guideline for GNSS link calibration and assigned several NMIs including NIM as the group 1 laboratories to implement the possibility of calibration of group 2 laboratories in the local RMO(Regional Metrology Organization) that might give some assist to BIPM.

II. FIRST NIM CALIBRATOR

The first NIM calibrator is IMEU that is a NIMTFGNSS-1 receiver, based on the GNSS module of Topcon company, homemade in 2008. IMEU is for portable using and operated in operation mode 1, where time and frequency of the receiver is directly synchronized to the local reference time scale by an internal phase lock loop and with which the receiver noise performance is better. Inside the receiver, we have the GNSS module located in one temperature chamber so that its ambient temperature can be controlled within the stability of less than 0.3(1 sigma) degrees, which would decrease the measurement uncertainty of the receiver due to the variation of the environmental temperature. IMEU was evaluated in the performances as a traveling receiver for the link and equipment calibration in [4].

It was the first time for it to be taken as a traveling receiver for the link calibration between NIM and PTB since 2010 and we got the consistent result in [5] with BIPM, which has been verified referenced to that by BIPM in 2009.

Since 2010, we have started to build the new timekeeping system at our new campus(Changping campus). Until 2012, we have operated for more than one year to test the performance of the new UTC(NIM) at Changping campus. Since Oct 14th 2012, we have changed our UTC(NIM) to the new time scale UTC(NIM1) implemented at the new campus. In order to contribute to the realization of TAI with low uncertainty, a calibration of the time transfer link between NIM and PTB is needed. IMPR receiver (Septentrio PolaRx2eTR) located at the old campus has been calibrated since the end of 2009 by BIPM and we used it as the reference receiver to calibrate IMEU with the differential calibration. Shortly After that, IMEU was moved to the new campus and calibrated the two GNSS time and frequency transfer receivers including IMEJ(Dicom GTR50) and BJNM(Septentrio PolaRx3eTR) in [6]. In June 2014, the time transfer links and IMEU of NIM have just been successfully calibrated again by BIPM as depicted in [7] and the results can agree well with those acquired by self-calibration of NIM in 2012.

In 2014, NIM sent IMEU to BSNC for the calibration tour. During the calibration, we have implemented the differential calibration between IMEU and BSNC GPS time transfer receivers, and INTDLY of one of BSNC receivers has been...
acquired as 54.2 ns that is 1.9 ns different from that by us this time in 2014.

III. PRESENT NIM CALIBRATOR

Afterwards, we modified our self-developed GNSS time and frequency transfer receiver NIMTFGNSS-1 to NIMTFGNSS-2, and at the same time for our G1 mission that we should be responsible for G2 lab calibration under APMP scheme we plan to create our new calibrator.

A. New GNSS Time and Frequency Transfer Receiver: NIMTFGNSS-2

During the time link or the equipment calibration, the most important thing is to measure the internal reference delay of IMEU manually, which is not so convenient for the user and will make some uncertainty due to the difference of the measurement procedures among the different people. So on the basis of NIMTFGNSS-1 receiver, we add one time interval counter inside the receiver for the real-time measurement and correction of the internal reference delay of the receiver to construct NIMTFGNSS-2 receiver. As well, we expand the possibility of adding one Rubidium clock to construct one UTC(NIM) disciplined oscillator(NIM), i.e. one time and frequency standard with high accuracy and precision.

There are two power switches on the receiver as described in Fig. 1 and 2. One is power switch for the hardware on the back panel of the receiver and the other is the power switch for the operation system on the front panel of the receiver.

This time and frequency transfer receiver NIMTFGNSS-2 that is the key part of the calibrator is simple to use as the traveling receiver. The front and back views are described in Fig. 1 and 2. We just need to connect the cables as Fig. 2 indicates and power on. The time and frequency references for the receiver are 1PPS and 10 MHz and the 10 MHz signal must be sinusoidal with a peak-to-peak amplitude ranging from 0.5VP-P to 3.0VP-P on 50Ω.

B. New Calibrator

At present, our new calibrator and its scheme are separately shown in Fig. 5 and 6.
As Fig. 6 demonstrated, the calibrator includes several parts which function description is as follows.

1) GNSS time and frequency transfer receiver(type: NIMTFGNSS-2): it is the main part of the calibration, which is responsible for the calibration measurements as traveling receiver. As well, we reserve one position of two euros inside the rack for GTR51 receiver, so that we can expand our system to two GNSS time and frequency transfer receivers for verification of the results and better reliability.

2) Time interval counter(SR620): it is assembled here for the user to measure the external reference delays of the calibrator and the user receiver at the user laboratory, so that the systematic error of SR620 could be cancelled out if the user uses it.

3) PDA and FDA: they are for the distribution of frequency and 1PPS signals to meet multiple use of the time and frequency signals inside the calibrator.

4) The compositive KVM(Keyboard, Video, Mouse) switcher is used for the interaction interface between PC and people.

There is one note that one 5 MHz frequency signal is needed from the G2 laboratory for the calibrator, and then one 10 MHz signal will be generated and be feed into the receiver by a frequency doubler inside the calibrator.

IV. EXPERIMENT SET-UPS AND OPERATING PROCEDURES

During the calibration campaign, NIM calibrator and the local system should be connected as Fig. 7 shows during the calibration.

The user had better take into account the multipath environment (reflecting surfaces) and masks when selecting the location of its antenna. It should be oriented according to manufacturer specs. The provided antenna cable must be used during the calibration campaign. The receiver should be in the same environment as the local receivers. The input frequency and the 1PPS signals must be derived from the same reference as for the local GNSS transfer system and the cables provided with the traveling equipment should be used for transporting the 1PPS and the frequency.

In the calibration campaign, several measurements need be finished as follows. The reference delays need to be measured by SR620 located at NIM calibrator at the beginning and end of the visit of the calibrator and the local receiver. Daily CGGTTS files of two receivers(NIMTFGNSS-2 and lab GNSS transfer system) should be logged for four to seven successive whole days. Daily Rinex files with a 30s data interval of two receivers(NIMTFGNSS-2 and lab GNSS transfer system) should be logged for four to seven successive whole days, if the local GNSS transfer system can generate Rinex files. If possible, the antenna cable delay of the local systems should be measured.

The precise positions of the local and traveling antennas should be known with an uncertainty of order (at most) a few cm. If the Rinex files of the local GNSS transfer system are available, the precise position of it can be determined by the solution analysis. And the Rinex files of NIM calibrator are able to be available.

Use these experiment set-ups, the time link calibration and the equipment calibration could be implemented at the same time.
V. CALIBRATION OF NIM CALIBRATOR

For the calibration of NIM calibrator, we use differential calibration referenced to IMEJ(Dicom GTR50) that is calibrated by BIPM in 2014. The detailed procedures can be found in [1] and CCD results are as follows and the calibration values for P3 code and C/A code have been calculated.

VI. CONCLUSIONS AND PERSPECTIVES

NIM has self-developed its second generation GNSS time and frequency transfer receiver NIMTFGNSS-2 and based on it one kind of homemade calibration system for the GNSS time link and the GNSS equipment calibration has been designed and realized. We have finished the construction of the calibrator and calibration of the calibration referenced to IMEJ calibrated by BIPM in 2014. With NIM calibrator, we could assist BIPM for the calibration campaigns of some TAI links under the APMP(Asia Pacific Metrology Programme) scheme.

REFERENCES

Discovery of Persistent Ionospheric Frequency Shifts of a few Herz and Impact on Time and Frequency Transfer

Michael J Underhill
Underhill Research Ltd
Lingfield, UK
e-mail: mike@underhill.co.uk

Abstract—Unexpected frequency shifts and occasional splitting of short-wave Standard Frequency Signals into components with persistent frequency shifts of up to a Herz or so have been ‘discovered’ using SDR receivers with milliHerz resolution. Such spurious signals have also been observed on AM carrier frequencies and DRM signals between ~1MHz and 25MHz. The frequency shifts are in general downwards like ‘Hubble’ shifts but at times can be upwards, depending on the time of day and direction of the path around the earth. Shifted components in general correlate with the existence of an ionospheric propagation path. At some times an unwanted frequency shifted component can be stronger than the original carrier signal and then the Standard Frequency Signal cannot be used as an accurate reference frequency. There are also problems with inaccurate pirate standard frequency signal transmissions on some frequencies. Small short term, diurnal and seasonal frequency variations are being investigated as a possible method of detecting gravity waves.

Keywords—Standard Frequency Signals; new ionospheric effects; SDR milliHz measurements

I. INTRODUCTION

New discoveries are made possible by advances in measurement techniques. The SDR (Software Defined Radio) is proving to be a major advance. It provides spectrum analysis at low cost with milliHerz resolution bandwidth capability for long recording times. SDRs facilitate several new applications and avenues for new research [1-5]. The three different SDR receivers used for the results given in this paper are shown in Fig. 1 [2,3].

Fig. 1. Stack of three Software Defined Radio (SDR) receivers. At top: RFSpace SDR-IQ, 0 to 30MHz, with TCXO reference. Middle: WinRadio ‘Excalibur’, 0 to 50MHz, with drift correctable crystal reference oscillator. At bottom: RFSpace SDR-IP 0 to 34MHz with OCXO reference.

The first indication of persistent ionospheric frequency shifts was seen in waterfall spectrum plots of 10kHz wide DRM (Digital Radio Mondiale) signals as shown in Fig. 2 and Fig. 3. The striations are found to be because there is a frequency shift that occurs for each ionospheric reflection. A double hop path gives double the frequency shift of a single hop path. The ‘striations’ are interference patterns between double hop and single hop components. The Ionospheric layer height is 15km times the number of striations in the 10kHz DRM spectrum width [4].

Fig. 2. 10kHz wide DRM spectrum from Issoudun France. Ionospheric (Hubble) shift of 0.35 Hz for layer height of 10.8×15 = 162km = 2×162= 324km path difference for one ionospheric reflection. The smaller two-hop double ionospheric bounce signal beats with the single hop stronger signal to cause the waterfall beat pattern of the 10kHz wide DRM signal and sideband beat pattern. Note negative frequency shift of this pattern at 1437 UTC. The shift is ~0.088ppm.

Fig. 3. As above but 0800 UTC. Note positive frequency shift of 0.22Hz or ~0.55ppm.

The two maximum frequency shifts shown in Figures 2 and 3 have a mean value that is negative indicating a ‘red-shift’ overall. However for this due south transmitter the shift is ‘diurnal’, positive in the morning and negative in the evening.

Fig. 3 shows a three day right to left frequency spectrum waterfall of 10MHz Standard Frequency signals as received at Lingfield, Surrey, UK. Such spurious signals have been observed on AM carrier frequencies and DRM signals between ~1MHz and 25MHz [2]. They correlate with the existence of
an Ionospheric propagation path. At some times the dependent signal can be stronger than the original carrier signal. Then the Standard Frequency Signal cannot in any way be used as an accurate reference frequency. The error can be 1ppm.

A. 10MHz Frequency Standard Components after Ionospheric Propagation – 19 sec Record

![Image](an10MHz_frequency_standard_components.png)

**Fig. 4.** On WinRadio Excalibur display: 10MHz Signals received at Lingfield, Surrey, UK, at 2147 UTC on 2 June 2014 from WWV Fort Collins USA (4700 mile NW path), BPM China (5200 miles NE path), PPE Rio de Janeiro, Brazil (5700 miles SW), WWVH Kekaha, Hawaii (7200 miles NNW) and Italian ‘pirate’ time signal located in Tuscany. Split into four main components, at +6, -6, -12, and -48Hz. Top left: 19 second reverse (upwards) waterfall with 1Hz resolution bandwidth. Snapshot at top right. 0 to 30MHz spectrum at bottom. 50 and 60Hz ‘hum’ sideband are visible.

An unexpected discovery was of the observation of the persistent existence of a dependent ‘spurious’ signal at an offset of ~10Hz below the 10MHz carrier frequency. The corresponding Doppler velocity is 300m/s or about 1000kph. For a one day duration the distance covered would be 24000 km if this was a Doppler frequency shift. Such velocities and distances do not correlate with any known physical process. It later was established that this was a ‘pirate’ transmission.

II. SPECTRUM RECORDS FROM SDR-IP AND SDR-IQ

In this section a selection of SDR-IP and SDR-IQ narrow band spectrum time waterfall records of various short-wave standard frequency transmissions is presented. The waterfalls are from right to left with durations of one to three days. Spectrum analyzer FFT resolution bandwidths of a few milliHerz have been used. Such records do not appear to have been published previously.

In each case the intensity colour coding appears as a scale at the top of the waterfall. Most plots have used a scale factor of 5dB/div as noted in a setting box at the bottom left of the pictures. The sampling rate, FFT size and resolution bandwidth are also displayed adjacently. The centre frequency and the span are displayed at the bottom centre. The waterfall frequency scale is displayed vertically on the left.

In the following the description and discussion for each record is included in the text of the figure captions.

A. Three Day 10MHz Record

![Image](an3_day_10MHz_record.png)

**Fig. 5.** Three day (to 01 Jan 2015) right to left waterfall record of 50 Hz span at 10MHz of mixture of carriers from WWV Fort Collins USA (4700 mile NW path), BPM China (5200 miles NE path), PPE Rio de Janeiro, Brazil (5700 miles SW), WWVH Kekaha, Hawaii (7200 miles NNW). From RFSpace SDR-IP receiver using SpectraVue software with 15mHz resolution. Lower trace at ~10Hz low is from Italian ‘pirate’ time signal.

B. Three Day 10MHz Record With Auroral Effects?

![Image](an3_day_10MHz_record_with_auroral_effects.png)

**Fig. 6.** Three day (up to 0919 UTC on 23 Jan 2015) right to left waterfall record of 50 Hz vertical span at 10MHz of mixture of carriers as in Fig 3. Plot from RFSpace SDR-IP receiver using SpectraVue software with 15mHz resolution. Enhanced propagation, probably auroral, with increased frequency fluctuations can be seen on the middle of the three days in Fig 3. The ‘pirate’ signal component at ~10Hz below 10MHz appears to have narrower bandwidth than the main components at 10MHz. The shorter 1400km path length appears to correlate with a narrower received carrier width. Ionosphere spreading tends to be greater for longer paths. Although a high Q ‘parametrically’ excited ‘whispering gallery’ mode in the ionosphere surrounding the earth could cause some filtering and narrowing of the carrier width?

C. One Day 10MHz Record

![Image](an1_day_10MHz_record.png)

**Fig. 7.** Figure 4. One day 10MHz record up to 28 Dec 2014 at 0102 UTC. 50Hz span. Note signal splitting around 0600 UTC (one third of the way from left) and appearance of Italian pirate signal just before this at about 10Hz low.
D. One Day 15MHz Record

Fig. 8. Figure 5. One day 15MHz record ending at 0744 UTC, 29 Mar 2015, 300Hz span. WWV, BPM and lower trace at ~100Hz excited by Italian pirate signal believed to be in Tuscany.

E. One Day 5MHz Record

Fig. 9. Consecutive overlapping one day 5MHz-right to-left records up to 0011 UTC on 10 Mar 2015 with 20Hz vertical span. Diurnal frequency shifts of up to 2Hz or 0.4ppm can be seen. Diurnal differences in spreading/dispersion can also be seen.

F. Three Day 1680 kHz Record with Auroral Effects?

Fig. 10. Figure 7. Three day USA to UK propagation on 1680 kHz up to 0019 UTC on 17 Mar 2015. 4mHz resolution bandwidth used. The SDR-IQ reference oscillator temperature drift can be seen on the main signal. But variable excitation of upper and lower ‘sideband modes’ can be seen, with spacings of about 0.4Hz. Solar flare on 11 March may have caused auroral effects on third day on right. The frequency drift can in principle be corrected digitally as discussed in [3].

III. DETECTION OF GRAVITY WAVES?

There are three physics processes that can change the frequency of a photon or a radio-wave: (a) Doppler shift which can be positive (approaching) or negative (receding) as the the relative motion between the source and the observer; (b) The Hubble Red-Shift, which here is separated from any Doppler motion and is thus in general negative [6,7], and (c) the Gravitational Red-Shift, which makes atomic clocks run slower in regions of high gravitational potential [7,8].

Reference [7] explores the possibility that the Hubble red-shift is not a Doppler shift due to the Universe expanding after the big bang, but is a steady loss of photon energy $E$ and hence a reduction in photon frequency $f$, on account of $E=hf$. The Hubble fractional red-shift $\delta f/f$ in terms of the Hubble Ratio $H_R$ ($7.31 \times 10^{-17}/m$) for a distance $d$ in metres is given by:

$$\delta f/f = H_d$$

As an example a frequency of 10MHz travelling for 5000km in free space will experience a frequency reduction of 366 $\mu$Hz, which is about a thousand times less than observed for the Ionospheric path measurements given here. The unanswered and un-calibrated question is how and to what degree the ionosphere has a red-shift greater than the assumed Hubble shift of free space?

What is also possible is that the observed red-shifts are related to the effect of the earth’s gravitational field rather than the free space gravitational potential [8]. It is well known that the frequency of atomic clocks is ‘red-shifted’ by the gravitational field [9]. But we cannot yet be sure that this is the cause of the red-shifts so far observed. More measurements are needed.

What is more evident on the records taken so far are the small $<$1Hz shorter term frequency fluctuations. These can be seen to vary over the course of a day. The speculation is that components of these fluctuations may be from variations in the solar system gravitational fields or the cosmic/galactic gravitational potential. If so we have a possible mechanism for detecting gravity waves!

We can therefore propose an arrangement for detecting gravitational waves that differs significantly from what has been tried so far [10]. The idea is use a stable laser source rather than a radio frequency source propagating over a reasonably long path in free space. The path could be between satellites in space or a reflected path to the moon, or to a geostationary satellite, and back the earth.

If the Hubble shift is a continuous process not primarily based on Doppler shifts of the expanding universe can estimate the fractional frequency shift to the
The signal processing should take account that gravity waves have been predicted to exist as fluctuations in the frequency spectrum $10^{-16} < f < 10^4$ Hz. The time for light to travel the path to the moon and back is ~2.56 seconds corresponding to a frequency of 0.39 Hz. Thus suitable phase coding of the laser at a few kHz can allow the laser beam to be operated as a linear phased array antenna with directional capability for gravity waves above about 0.4Hz. At frequencies below this, samples taken from the phase coded records at longer intervals can allow Synthetic Aperture Radar (SAR) type of beam-forming at lower frequencies than 0.4Hz by virtue of transverse movement of the laser path relative to the direction of the source of the gravity waves.

IV. CONCLUSIONS

The results given in this paper are a small selection of records collected over the last one or two years. Some of the conclusions have been derived from records not presented here.

The ionosphere can frequency shift carrier components of frequencies between 1 and 25MHz. These effects can persist for hours and repeat daily. Shifts can be up to 1ppm or more. When this occurs short wave frequency standards can become unusable. Occasionally a carrier can split into two or more components.

Frequency shifts are mainly down as for Hubble redshifts. But they can be up depending on the dominant direction of propagation around the earth.

Shorter paths in general show less frequency spreading or dispersion than for longer paths around the earth.

But when Ionospheric Round the World Echo (RWE) conditions exist there can be filtering and unstable injection locking of round-the-world ‘whispering gallery’ ionospheric modes. The instantaneous bandwidth can appear improved by the filtering action, but the frequency excursion can become larger.

The hum sidebands of some of the transmitted frequency standards demonstrate the same sort of fluctuations that have been investigated for oscillators [1,2].

It is thought that some of the short term and diurnal variations may be due to fluctuations in gravitation potential/field over the propagation path. This may provide a new technique for detecting gravity waves. A proposal for a detection scheme has been made. It is proposed that this should be investigated further.

The impact of what has been discovered on timing errors has not yet been sufficiently measured or assessed for any substantive conclusions to be drawn for these.

REFERENCES

Research on Time and Frequency Transfer based on BeiDou Common View

Hang Yi*,† Hongbo Wang*,† Shengkang Zhang*,† Haifeng Wang*,† Fan Shi*,† Huaiying Shang*,† Jun Ge† Yujie Yang† and Zhiqi Li‡
*Beijing Institute of Radio Metrology and Measurement, Beijing China 100854
†Science and Technology on Metrology and Calibration Laboratory, Beijing China 100854
‡Xidian University, Xi’an China 710071

Abstract—With the development of BeiDou(BD) Navigation Satellite system, BD can be another choice for remote precise time and frequency transfer. A strict common view test using BD Navigation Satellite System is carried out in this paper. Since there are GEO, IGSO and MEO satellites in BD Navigation Satellite system, common view based on GEO, IGSO and MEO are discussed in this paper separately. Finally, we give a different weight to each satellite according to its elevation, and get a weighted-average common-view result which is better than 5 ns.

Keywords—time transfer; BeiDou; common view;

I. INTRODUCTION

BD Navigation Satellite System is the global navigation satellite system developed by China independently. At present, there are already 5 GEO, 5 IGSO and 4 MEO BD navigation satellites in orbit which can provide stable and continuous navigation services for the Asia-Pacific region. Since BD Navigation Satellite System is going to provide the global navigation service in 2020, time and frequency transfer using BD Navigation Satellite System will be another choice in the remote precise time and frequency transfer field.

Common view method is one of the main methods for remote precise time and frequency transfer [1]. At present, the application of common view is mainly based on GPS. In this paper, the common view method using BD Navigation Satellite System is researched.

Considering that GEO and IGSO have longer time in view than MEO, it is easier to realize strict common view using BD Navigation Satellite System. As a result, we discussed the common view method using GEO, IGSO and MEO separately at first. Then, all the satellites in view are used in our common view time and frequency transfer. We give a different weight to each satellite according to its elevation and get a weighted average common view result which is much better.

This paper is organized as below. At first of this paper, the common view time transfer method using BD and our common view experiment are introduced in section II. And then, our data process method is introduced in section III. Next, the common-view results are given in section IV. Finally, in section V, it is about our future research.

II. COMMON VIEW TIME TRANSFER BASED ON BEIDOU

A. The common view principle based on BD

The common view principle based on BD is shown in fig 1. The same as the GPS common view, station A and station B are two time reference station. Each station has its own atomic clock. They provide 10MHz and 1PPS signal to the two receivers. The receiver output the code phase observation, the carrier phase observation and the navigation broadcast ephemeris. Using these messages, we can calculate the clock error between the receiver and the satellite clock on board. Then, the clock error data are exchanged through the internet and we can get the clock error between the two stations [2] [3].

Assuming that δτA, δτB is the clock error between the BD satellite and station A and B separately. Then, we have:

\[
\delta \tau_A = \tau_A - \tau_{sat} \\
\delta \tau_B = \tau_B - \tau_{sat}
\]

\[
\tau_A, \tau_B, \tau_{sat} \text{ is the time of station A, station B and BD satellite separately.}
\]

Then the clock error between station A and station B can be written as

\[
\delta \tau_{AB} = \delta \tau_A + \delta \tau_B
\]
\[ \delta \tau_{AB} = \tau_A - \tau_B = \delta \tau_A - \delta \tau_B \quad (3) \]

**B. Common view experiment based on BD**

Actually, the observation data of the BD receiver is matched with the local 1PPS of the receiver. The clock error we calculate by the observation data is the clock error between the local clock of BD receiver and the satellite clock. In order to calculate the clock error between the local clock of the time reference station and the satellite clock, the receiver must satisfy the characteristics as follows, 1) external reference input; 2) local 1PPS output; 3) raw observation data output. Then, we should give the 10MHz signal of the time reference station to the BD receiver, and compare the 1PPS signal of the local clock with the 1PPS output signal of the BD receiver. The relationship is shown in fig 2.

**III. DATA PROCESS**

In this section, the whole data process is introduced in detail. The data process method is as follows.

- Use empirical model to calculate the troposphere delay.
- Use weight average method to get better results.

![Data Process Flow](image)

**A. Pseudo range smooth**

In this paper, we use the former 20 epochs data to smooth the current epoch data. The pseudo range of epoch 1 can be estimated through the code phase data and carrier phase data of epoch k. Assume that \( \rho_{s1k} \) is the pseudo range of epoch 1, it can be expressed by

\[ \rho_{s1k} = \rho_k - \lambda (\phi_k - \phi_1) \quad (4) \]

The average of \( \rho_{s1k} \) can be taken as the smoothed pseudo range of epoch 1.

\[ \rho_{s1} = \frac{1}{K} \sum \rho_{s1k} \quad (5) \]

where, k=1, 2, 3 .. K, and in our experiment, K=20.

**B. Geometric distance**

The position of the receiver is exactly known and we have to calculate the position of BD satellite. In order to get position of BD satellite we have to calculate the clock offset between the send time of BD signal and TOE. The send time (BD Time, BDT) of BD signal is calculated as follows [4].

\[ t = t_{SV} - \Delta t_{SV} \quad (6) \]

\[ \Delta t_{SV} = a_0 + a_1 (t - t_{SC}) + a_2 (t - t_{SC})^2 + \Delta t_r \quad (7) \]

\[ \Delta t_r = F \cdot e \cdot A^{1/2} \sin(E_k) \quad (8) \]

where \( t \) is the send time (BDT) of BD signal, \( t_{SV} \) is the send time (BD satellite time) of BD signal, \( \Delta t_{SV} \) is the clock error of the satellite clock, \( \Delta t_r \) is the relativistic correction.

**C. Clock error of BD Satellites**

The clock error of BD Satellites can be calculated by equation (7).
D. Ionosphere Delay

The BD receiver we used is a dual-frequency receiver which can track both B1 and B2 signal. The double frequency data can be used to calculate the ionosphere delay.

\[ I_1 = f_1^2 * (f_1^2 - f_2^2) * (\rho_2 - \rho_1) \]  
(9)

\[ I_2 = f_2^2 * (f_1^2 - f_2^2) * (\rho_2 - \rho_1) \]  
(10)

E. Troposphere Delay

Troposphere delay is calculated through an empirical model in our experiment as follows.

\[ T = \frac{2.47}{\sin \theta + 0.0121} \]  
(11)

where \( \theta \) is the elevation of the satellite.

F. Weight Average

The value of the weight of each satellite is calculated according to its elevation and it will be set to zero when the elevation is lower than 20 degree.

\[ \omega_i = \theta_i / \sum \theta_i \quad (\theta_i > 20) \]  
(12)

where \( \omega_i \) is the weight of satellite \( i \), \( \theta_i \) is the elevation of satellite \( i \).

IV. RESULTS AND ANALYSES

The results of the zero baseline experiment are shown in this section. And the common view interval is 1s.

A. Common view results based on GEO IGSO and MEO

In BD navigation satellite system, satellites of PRN1~5 are GEO, satellites of PRN6~10 are IGSO and the others are MEO. Our results are shown in fig5~8. The y label is the common view error and x is time in the figures. In order to show clearly, the system error (about 0.00055975s) has been removed. Fig5 and fig 6 are the common view results of GEO and fig.5 is the result of a GEO with a higher elevation. We can see that for the GEO with a high elevation the peak-to-peak jitter is 8ns, for the GEO with a low elevation is 15ns. Obviously, the higher the elevation is, the better the common view result is. Fig. 7~8 is the common view result of the IGSO and MEO separately. We can see that the common view error is up to 60ns when the elevation is too low. Compared with the GEO, common view results using IGSO and MEO are discontinuous. Also, we can see that the mean value of the common view error using different satellites is a little different which needs some more research.
B. Weight average results

The two receivers of the two time reference station can track several same BD satellites at the same time. According to equation (12), we give each satellite a different weight and the result is shown in fig. 9. We can see that the final common view result is better than 5ns.

![Common view result (weight average)](image)

Fig. 9. Common view result (weight average)

V. FURTHER RESEARCH

We are going to perform the long baseline experiment to get some more data to test the capability of BD Navigation Satellite System in time and frequency transfer field. TWSTFT will be a good way to check our method [5] [6]. Besides, BD carrier phase time transfer will also be researched.

ACKNOWLEDGMENT

We thank all the members in science and technology on metrology and calibration laboratory for their help in our research and experiment.

REFERENCES

Preparing ACES-PHARAO data analysis

Frédéric Meynadier*, Pacôme Delva*, Christophe le Poncin-Laffite*, Christine Guerlin†, Philippe Laurent* and Peter Wolf*

* LNE-SYRTE, Observatoire de Paris, CNRS, UPMC, LNE,
61 avenue de l’Observatoire 75014 Paris, France
Email: Frederic.Meynadier@obspm.fr
† Laboratoire Kastler-Brossel, ENS, CNRS, UPMC,
24 rue Lohmond, 75005 Paris, France

Abstract—The Atomic Clocks Ensemble in Space (ACES-PHARAO mission [1]) will realize, on board of the International Space Station, a time scale of very high stability and accuracy, connected to a dedicated microwave link (MWL) for comparison to ground clocks. This will allow to perform high performance time transfers between distant ground stations, as well as fundamental physics tests, such as measuring the gravitational redshift with unprecedented accuracy, and search for a violation of the Lorentz local invariance.

Our team at SYRTE is currently developing a dedicated software for analyzing the data retrieved by the MWL. It will be a crucial component of the processing chain, allowing to reach the nominal scientific performances, thus complementing the quick-look data provided by the ground segment. During ACES flight, our software will run in a data processing center that will be set up in SYRTE, and process data as requested by IWG (ACES Investigator Working Group) request.

We will present the current status of the data analysis software and its expected capabilities, mostly based on data generated by our own data simulator, but also some preliminary results of comparisons with actual test data from the instruments. Our simulator itself will also be presented, as it plays a key role in the validation of the data analysis software, and also allows to assert the sensitivity of ACES data to various theoretical signatures.

II. GLOBAL ORGANIZATION OF ACES GROUND SEGMENT FOR MWL DATA

ACES MWL Ground stations will be dispatched at several participating metrology institutes around the world. Each ground station will be connected to a 100 MHz signal and PPS provided by the institute, against which local clock can be measured. As the ISS orbits around the earth, it becomes visible for one (or more) station: each ground station in sight then starts to transmit and receive time comparison data. One data point is obtained every 80ms (approximately) on ground and in space during the pass.

Once the pass is finished (typically 300-500 seconds), received raw data is automatically transferred to the central Data Processing Center for the Colombus module in CADMOS (CNES), where it is stored in an archive together with various housekeeping data. Then several steps of data processing are needed to get to scientifically usable data, i.e. desynchronisation measurement between the ground and the space timescale. The various components of the processing chain are presented on fig. 1.

The MWL core data consists in:

- one-way time transfer data for uplink Ku-band signal (approx. 13.5 GHz, collected in space)
• one-way time transfer data for downlink Ku-band signal (approx. 14.7 GHz, collected at the ground station)

• one-way time transfer data for downlink S-band signal (approx. 2.2 GHz, collected at ground station).

Combining the first two signals allows to perform a two-way time transfer, while combining the last two signals allows to determine the delay due to the ionosphere traversal (as it is a dispersive medium, measuring the differential delay between 2.2 and 14.7 GHz allows to calculate precisely what is its effect).

Additionally, time transfer data is determined either by determining the phase of the incoming carrier (dubbed "carrier" data), or by matching an encoded pattern at 100Mchips rate (dubbed "code" data). Code data is unambiguous, but coarse (20 ps resolution). Carrier data can theoretically achieve 1ps resolution, but suffers from phase ambiguity (any measurement will only be able to give a modulo-2π phase). Combining both measurements correctly is key to optimal performance of the MWL.

Auxiliary data needed for data processing will also be contributed to the archive : instrumental calibrations, ISS orbitography, ground station positions and atmospheric parameters, etc.

III. TWO-WAY MICROWAVE LINK

The basic principle of such a link has been presented in Delva (2012) [4]. In summary : sending a timestamp from clock A to clock B while sending a timestamp the other way, from clock B to clock A, allows (at the first order) to cancel the signal’s propagation time between the two clocks, thus allowing the calculation of the desynchronisation without precise knowledge of this time of flight. However, we remain sensitive to the variation of this value between the uplink and downlink measurement : ISS orbitography and ground station positionning allows us to model this effect to a certain level.

An even greater accuracy can be achieved by using the so-called Λ configuration : in this configuration, the downlink signal leaves the space clock just when the uplink signal reaches it, thus minimizing the impact of our uncertainty on the ISS position. Figures 2 and 3 illustrate this configuration.

However, calculating the correct date to which we should interpolate the signals requires previous knowledge of various delays, including a first approximation of the time of flight (as accurate as 1µs, see [5]). We therefore have to iterate : first we assume that clocks are synchronized, which allows us to derive a first estimate of the different delays, which in turn allows to reach a first estimate of the effective desynchronisation. Then we reinject those values in the data processing to reach a better accuracy. Our experiments show that three iterations of the loop allow to recover any desynchronisation susceptible to happen during the mission (i.e less than a few 10 ms), two being sufficient in most cases.

IV. PROCESSING AND ANALYSIS SOFTWARE

Software development for this project started several years ago, with an initial objective of producing a prototype for the operational software that will run within the ground segment. It was also foreseen to use this software to test alternative analysis during ACES flight, as a flexible tool to inject any effect we would like to test. We chose the Python language [6] together with numpy libraries (numerical python, [7]) to have a flexible yet fast enough processing chain. The Matplotlib library [8] is also heavily used for data representation.

It has been decided in 2014 by ESA that our software would become an integral part of the operational data processing chain, and that SYRTE would host a Data Processing Center for this purpose. This imposes new constraints on future software development, as robustness and automation become not only desired, but needed. However this should not change significantly the development path of the current software : we estimate that the current architecture will eventually achieve all these goals. CPU load should not be a concern either, as a 500s pass is currently processed in 170s on one core i7@2.7GHz, with the possibility of processing multiple passes in parallel.

V. SIMULATION SOFTWARE

Independently from the processing software, we are developing a simulation software that allows to produce realistic data. The goals are :

• To accompany the development of our processing software, by producing a large range of well-
controlled data (from very simple scenarios, e.g. "constant range, no atmosphere, no clock drift", to realistic ones).

- To provide a test bench for the capabilities of the software once it is finished, i.e. check that a given effect will be visible or not in the end.
- To provide data samples for other actors of the ground segment.

In order to achieve the first goal, the simulation has to be highly modular and allow to switch on or off any effect that may be encountered. A wide range of relative motions between the stations should also be possible. Also, we have taken care too separate as much as possible the development of the processing and the development of the simulation software: they use different languages (simulation uses Matlab [9]), are written by different people, and adopt different approaches.

VI. CURRENT STATUS

Figure 4 presents the processing chain as a flowchart. The following sections describe the development status of the modules represented there.

Figure 5 is the output of the tool we have designed for validating our software during its coding: as much as possible, each effect is isolated during the process and compared to its theoretical value (i.e. the value we can calculate from simulation’s input parameters). It allowed us to check, after the inclusion of each new feature in the code, how other modules were affected and, in case of observed regressions, find the guilty module.

A. Preprocessing (1st line of fig 5)

Our data analysis algorithm expects input data to be very similar to pseudo-ranges (as in GNSS time transfer). However the data provided by MWL telemetry is quite different (see [4]), therefore we had to implement a pre-processing module that converts the hardware data into those pseudo-ranges. This required a large amount of work and numerous exchanges with the instrument makers, as it was crucial not to loose any performance at this step.

A side benefit of this module is to isolate the core of the processing software from most variations within the hardware’s output format, or changes in the raw data structure: this minimized the impact of some of the changes that were deemed necessary throughout the hardware construction.

A spread of 20ps for each individual code data point, as well as a 20 ps global offset which varies from one dataset to another. Those performances are nominal: technically, the data that is retrieved from MWL terminal consists in integer counter data, sampling time intervals at 100.1953125 MHz. Raw data therefore as $\approx 10$ ns resolution, but the measurement principle relies on a beatnote, so the effective resolution is 10 ns multiplied by the ratio between the beatnote and the signal’s frequency, i.e. $\approx 10 \cdot 10^{-9} \times 195 \cdot 10^{3}/100 \cdot 10^6 = 19.5$ ps. Our global offset determination is also limited by this uncertainty.

For carrier data, the calculation gives $\approx 10 \cdot 10^{-9} \times 729 \cdot 10^3/14 \cdot 10^9 = 0.5$ ps. The observed spread for carrier data is in line with this result. However carrier data exhibits large offsets ($\approx 50$ ps for $f_1$, $\approx -15$ps for $f_2$ in this example). This is due to the fact that carrier offset determination can not be determined exactly as for the code. We are currently working on this and expect offset values to be inferior or equal to the spread in the final version of the software for the carrier, as it is with the code.

B. Tropospheric delays (2nd line of fig 5)

Tropospheric delay is a (mainly) non-dispersive effect that can reach several hundred nanosecond on each signal. Two-way measurement cancel it to a large extent, but we want to evaluate it as well as possible at higher orders. Some dispersive, second order effects have been studied by Hobiger, Piester and Baron (2013) [10] and will be implemented in the future.

Current model uses Saastamoinen formula [11] to derive it from local atmospheric parameters, both in the simulation and the processing software. It is therefore merely a self test of interpolation routines and basic geometry match. During flight, atmospheric parameters estimation will anyway be quite coarse and probably not completely representative of integrated values along the signal’s path.

C. Ionospheric delays (3rd, 4th and 5th lines of fig 5)

These are due to the dispersive nature of the ionosphere, and therefore affect differently frequencies $f_1$ and $f_2$. We separated those in 2 terms, one that is proportional to $1/f^2$, the other that is proportional to $1/f^3$. Slight differences may be seen in the latter case: these have been identified as results of the difference between the model used for Earth’s magnetic field in the simulation and in the processing software, and remain largely below detection threshold.

Ionospheric delays imply to calculate the STEC (Slanted Total Electron Content) as a by-product. This is a measurement of the density of electrons, integrated along the line of sight, and its "normalized" counterpart (vertical TEC) is routinely monitored for GNSS. The STEC residuals line shows the uncertainty on STEC determination, which is evaluated from code data of the downlink signals: as STEC values are of the order of $10^{18}$, the observed spread of $10^{16}$ (corresponding to the 20 ps spread on the pseudo-range determination) represents roughly 1%, which is enough for satisfactory calculation of the ionospheric delays.

D. Geometrical time of flight and range (6th and 7th lines of fig 5)

We define Geometrical time of flight as a "classical, atmosphere-free" coordinate time interval between emission and reception of a given signal: this is simply the distance between emission point and reception point, divided by $c$. As both stations move with respect to GCRF, this calculation is not completely straightforward and needs iterations or Taylor series developments to converge.

Range is the instantaneous distance between the two stations at a given coordinate time. This is calculated from the ISS orbitography files and the ground station coordinates, after conversion to GCRF.


\[ T_m = \text{Range} + \text{Tropo. delay} \]

This is an internal consistency test : \( T_m \) is one of the observables that is coming out of the hardware, this graph shows how well we recover it. Spread is 10 ns for \( f_2 \), which is expected because \( T_m \) is basically an integer counter value, incrementing at roughly 100 MHz. For \( f_1 \) the spread is much lower because this signal is interpolated for achieving the \( \Lambda \) configuration, thus averaging on several point to find the interpolated value.

\[ \text{F. Desynchronisation and associated time deviation (9th and 10th lines of fig 5)} \]

This is the end result of the processing : apply two-way measurement method, compensating for all other effects, and recover the clock desynchronisation. In this example, the initial desynchronisation between clocks was 0.1 ms, and both frequencies drift because of gravitational potential.

Residual spreads are consistent with those of the pseudo ranges (first line), which is expected. The offset for carrier-phase data is the mean between the offsets on pseudo-ranges, it is a direct consequence of the offsets on pseudo-ranges which should disappear once those are removed.

Apart from this offset, no visible effects are noticed throughout the pass, showing that all other effects were correctly removed.

\[ \text{VII. Conclusion} \]

As ACES launch is getting closer (2017), it is crucial to be ready to process the data that will be collected during flight. Although some areas still require refinements, the core software we have prepared is in good shape for this duty and should be ready in time for inclusion into the processing chain.

The following months should see the completion of hardware calibration : we will analyze the corresponding data in order to implement the "calibration" module that is still undefined. Conversely, our simulated data has been sent for processing through existing ground segment pipelines.

\[ \text{REFERENCES} \]

Fig. 5. Typical output from a comparison between the input data (i.e. theoretical values calculated by the simulation), and intermediate/final results calculated by the processing software. In this example, the simulated data includes a keplerian orbit for the ISS, earth rotation, atmospheric delays with variable parameters (troposphere + ionosphere), initial desynchronisation between the clock equal to 0.1 ms, frequency drifts due to gravitational potential for ground and space. Each point is the difference between the expected data at this date (known from the simulation’s scenario), and the data actually calculated by the processing software for the same coordinate time. Ideally we should therefore get values close to zero + noise. Blue points refer to uplink (f1) signals, green points refer to downlink (f2) signals, and red points refer to desynchronisation (which combines both signals). Left column contains result for carrier-phase data, and right column for code-phase data. See text for a description of the lines.
Abstract—Using fiber based frequency dissemination system, we are building Beijing regional time and frequency network. Currently, Tsinghua University (THU), the Changing site of National Institute of Metrology (NIM-Changping) and Beijing Institute of Radio Measurement (BRIM) have been linked and synchronized via the fiber network. The frequency signals of three hydrogen masers, two placed in NIM, one in BRIM are transferred to Tsinghua University, and compared with the local hydrogen maser. By this remote comparison, the absolute frequency stability of each clock as well as the correlations between the two clocks in NIM are measured and studied.

Keywords—frequency transfer; frequency stability; three-corner-hat; frequency network, correlations

I. INTRODUCTION

Atomic clocks in the same place always affected by common environments, such as temperature, pressure, gravity and so on. Consequently, for a group of clocks placed in the same environment, certain correlation always exists between them, but it cannot be easily measured [1-3]. Thanks to the fiber based frequency dissemination techniques developed recently [4-12], we can give an estimation of this correlation through comparing them with the clocks at remote locations.

In this paper we introduced our study on the absolute frequency stability measurement and correlations of the hydrogen masers. The results show that in long averaging time (>10000s), there exists strong correlations between clocks in the same place.

II. CORRELATIONS BETWEEN CLOCKS

The frequency stability of an atomic clock often represented by Allan Variance:
\[
\sigma^2(\tau) = \frac{1}{2M} \sum_{i=1}^{M} (\Delta y_i)^2,
\]  
where \(\Delta y_i\) is the difference between the \((i+1)\)th and ith of M fractional frequency values averaged over the measurement interval \(\tau\). \(\Delta y_i = y_{i+1} - y_i\).

In practice, the frequency stability of an atomic clock can be measured only by comparing its output with that of one or more other clocks. The result of any such measurement includes not only the instability of the clocks being measured, but also that of the reference clock or ensemble. Consider three clocks, A, B, and C. The frequency stability of clock A is to be measured, and clock B and C are served as reference. By simultaneously measure the frequency stability of each two clocks, we can get the result:
\[
\sigma^2_{ab} = \sigma^2_a + \sigma^2_b - \frac{1}{M} \sum_{i=1}^{M} (\Delta y_i^a \Delta y_i^b),
\]
\[
\sigma^2_{ac} = \sigma^2_a + \sigma^2_c - \frac{1}{M} \sum_{i=1}^{M} (\Delta y_i^a \Delta y_i^c),
\]
\[
\sigma^2_{bc} = \sigma^2_b + \sigma^2_c - \frac{1}{M} \sum_{i=1}^{M} (\Delta y_i^b \Delta y_i^c),
\]  
(2)

\(\sigma^2_{ab}\) is the total frequency stability (relative stability) between clock A and B, \(\sigma^2_a, \sigma^2_b\) is the independent stability (absolute stability) of clock A and B. The last term of each equation is the correlation term which means the correlation between every two clocks. Using traditional method, we can’t measure or solve this term independently. There are some methods supposed by several research groups to judge the strength of the correlation term [2,3]. While in most cases, we suppose there are no correlations between the clocks, the last term is considered to be zero. Then the frequency stability equation can be simplified as:
\[
\sigma^2_{ab} = \sigma^2_a + \sigma^2_b,
\]
\[
\sigma^2_{ac} = \sigma^2_a + \sigma^2_c,
\]
\[
\sigma^2_{bc} = \sigma^2_b + \sigma^2_c.
\]  
(3)

Equations (3) can be solved for \(\sigma^2_a\), a measure of the frequency stability of the individual clock A:
\[
\sigma^2_a = \frac{1}{2}(\sigma^2_{ab} + \sigma^2_{ac} - \sigma^2_{bc}).
\]  
(4)

Thanks to the fiber based frequency dissemination techniques developed recently, we can get the frequency stability of a clock in a more precision level.

III. COMPARISON OF FOUR REMOTE CLOCKS

From 2013, we start the program of Beijing regional time
and frequency network (Fig. 1) [4,12]. Based on this network and using three-corner-hat method, we simultaneously compared four hydrogen-masers placed in the Changping site of National Institute of Metrology (NIM-CP), THU and BIRM, respectively. Through the fiber links between these three places, the clocks A, B, D, are recovered in THU’ lab. Each of the transition link is set up with a fiber based stable frequency dissemination system. The typical transition stability of the fiber link is about $1 \times 10^{-12}$s, $3 \times 10^{-17}/10^5$s, which can fully satisfy the frequency dissemination stability requirements of these hydrogen masers [13] (Fig. 2).

![Figure 1. Beijing regional time and frequency network. The solid line is the synchronization link under used. The dash line is the synchronization link under construction. Symbol A, B, C, D are Hydrogen masers, while symbol E is a Sr optical clock.](image1)

There are two H-Masers (Maser A, Maser B) placed in the same room at NIM-CP. As we mentioned above, there exits certain correlations between these two clocks. Through three-corner-hat comparison between clocks A, C, D and clocks B, C, D, we can get the independent frequency stability of these four masers by using equation (4). In fact, as we divided the clocks into two groups, and each group contains clocks C and D, so the stability of clock C and D would have two results. The results of the same clock from different group would indicate the reliability of the measurement: the closer, the better.

Considering no correlations between clock A and B, the relative stability of these two clocks can be calculated as: $\sigma_{ab}^2(cal) = \sigma_2^2 + \sigma_3^2$. But in practice, the measured relative frequency stability $\sigma_{ab}^2(meas)$ of clocks A and B always contains the correlation term. By comparing the frequency stability of $\sigma_{ab}^2(cal)$ and $\sigma_{ab}^2(meas)$ at different averaging time, we can not only know the strength of the correlation term, but also the averaging time it play a part in the total relative frequency stability.

In summary, through stable frequency dissemination systems made by ourselves, we realized the absolute frequency stability measurement of the remote hydrogen masers. And furthermore, we studied the correlations between hydrogen masers placed in the same place. We find that there is strong correlations between these two clocks at long averaging time ($>10^4$s), which would be probably caused by the similar temperature and air pressure changing.

![Figure 2. The typical performance of the hydrogen maser and the stable fiber link. Through active noise compensation technique, the link’s frequency dissemination stability is about 1 order higher than the maser.](image2)

**ACKNOWLEDGMENT**

This work was supported by the National Key Scientific Instrument and Equipment Development Project (No. 2013YQ09094303) and the Beijing Higher Education Young Elite Teacher Project (No. YETP0088).

**REFERENCES**


LMJ Timing and fiducial System: overview of the global architecture and performances

V. Drouet, M. Prat, P. Raybaut, D. Sainte-Beuve
CEA-DAM, DIF, F-91297 Arpajon, France
vincent.drouet@cea.fr

Abstract—The Laser MegaJoule (LMJ) timing system has to synchronize 176 laser beams better than 40 ps rms to compress symmetrically the millimeter-size target in order to ignite the deuterium and tritium filled capsule despite the fact that the quadruplet laser sources are separated within the building by several hundreds of meters. This kind of performance is also required for fiducial pulses used to temporally mark laser and plasma diagnostics.

After the LMJ was officially commissioned on October 23rd 2014 with a first successful physics campaign, this article offers an overview of the final overall timing system architecture and its performances.

The latest results since our presentation during the previous EFTF conference edition [1] as part of ongoing studies on time drifts are also presented.

Keywords—LMJ; Timing system; picosecond time drift; CEA; Standard and High Precision Timing; Ultra-High Precision Timing; global architecture.

I. INTRODUCTION

Laser MegaJoule is a high-power laser facility located near Bordeaux, France (see Fig. 1) [2]. It is designed to achieve Inertial Confinement Fusion (ICF) of a deuterium-tritium target, in order to validate physical models of the behavior of matter in extreme conditions of temperature and pressure.

To achieve this, LMJ will focus the power of 176 laser beams onto the target, delivering 1.2 MJ @ 351nm in a time frame of 40 ps rms. In fact, the laser beams are directed against the inner walls of the gold hohlraum holding the microcapsule to produce intense X-rays. The radiation is trapped inside the cavity, where the temperature can reach 10 MK to 100 MK for about 100 ps. It is the X-rays that interact homogeneously with the microcapsule: this method is called “indirect drive”.

The implosion of the capsule will also generate particles and electro-magnetic fields that will put severe constraints to the operation of experimental data-collecting instruments and the quality of their results.

Laser MegaJoule is currently under exploitation, after it was officially commissioned on October 23rd 2014 with a first physics successful campaign.

II. THE LASER MEGAjOULE TIMING SYSTEM

A. Goals and challenges

Synchronization is necessary to ensure the ordered sequencing of Laser sources and LMJ diagnostics by generating and delivering correlated trigger signals. The most demanding experiences need to synchronize the 176 laser beams to better than 40 ps rms despite the fact that the quadruplet laser sources are separated within the building by several hundred of meters (facility dimensions: 300 m x 100 m, see figure 1).

This kind of performance is also required for fiducial electrical pulses, used to temporally mark laser and plasma diagnostics. Timing and fiducial systems designs are both based on very accurate electrical delay generators connected to a master clock through a time distribution network which can be optical or electrical depending on classes of performance:

- Standard and High Precision Timing system (1800 channels),
- Ultra-high Precision Timing System (60 channels),
- Electrical Fiducial System (200 channels).

To outline our challenge, we have to deploy a synchronization system that offers performance typically encountered in time-frequency laboratories (regarding jitter and drift in particular), over long distances (hundreds of meters) and in such quantity (> 2000 channels).
TABLE I. LMJ TIMING AND FIDUCIAL SYSTEMS REQUIREMENTS

<table>
<thead>
<tr>
<th>Classes</th>
<th>Delay range (peak to peak)</th>
<th>RMS jitter (peak to peak)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 second</td>
<td>&lt; 5 ns</td>
</tr>
<tr>
<td>High Precision</td>
<td>100 µs</td>
<td>&lt; 15 ps</td>
</tr>
<tr>
<td>Ultra-High Precision</td>
<td>100 ns</td>
<td>&lt; 5 ps</td>
</tr>
<tr>
<td>Fiducial</td>
<td>-</td>
<td>&lt; 5 ps</td>
</tr>
</tbody>
</table>

B. LMJ Timing Systems requirements

TABLE I. summarizes performance needed for the four classes of performances of LMJ Timing and Fiducial Systems.

Specifications of Standard and High Precision (SHP) Timing system are directly inherited from the synchronization equipment used on the Ligne d’Intégration Laser (LIL), which was the prototype of the LMJ. However, this performance is not sufficient in order to reach the LMJ 40 ps rms specification. That is why the Ultra-high Precision (UHP) Timing system was designed.

C. LMJ Timing Systems technologies

The global architecture of LMJ timing and fiducial system is divided in two types of technologies:

- **Standard and High Precision Timing system** is based on opto-electronic equipment: the master provides an optical digital clock to all slaves (general principle is described by [3]) through an Optical Time Network (OTN, see §.IV). This technology was developed to provide the LIL and is now well-controlled. However, some improvements have been integrated in LMJ synchronization slaves (e.g. auto-calibration of electronic vernier to guarantee delays accuracy requirements).

- **Ultra-high Precision and Fiducial systems** are based on the use of passive electronic components only which allows reaching very low jitter and time drift. An analogic electrical clock (high voltage pulses delivered with a frequency from single shot up to 1 kHz) is distributed to all slaves through a Coaxial Time Network. The feasibility of UHP system based on this technology has been demonstrated through a demonstrator in our laboratory. The principle and technology are presented by [4][5].

- **A unique master** is the link between these various synchronization systems by delivering perfectly synchronous signals.

III. PROSPECTS OF IMPROVEMENT THANKS TO DRIFTS MEASUREMENTS

We are leading some studies to analyze critical functions of our timing systems in order to improve their performances, particularly regarding time drift aspects.

This part illustrates one of the processes currently conducted applied to opto-electronic equipment of the High Precision class.

Fig. 2. Optimal configuration set-up for picosecond time drift measurements

A. Picosecond time drift: method & measurement

Thanks to a collaboration between CEA and Besançon Observatory (2013-2014) we have developed and qualified time drift measurement solutions for delays ranging from µs to tens of µs with an accuracy of 2 ps over 1 day for the most restrictive case (Ultra-high Precision class). The approach and the results we obtained are presented in [1].

Fig. 2 presents the optimal configuration set-up for picosecond time drift measurements we developed during the collaboration: the reference frequencies for the chronometer and the delay generator must be slightly shifted in order to avoid unexpected noise due to the non-linearity of the interpolators.

Fig. 3 presents the results obtained with the optimal configuration presented: we have demonstrated the **picosecond time drift measurement feasibility** for the Ultra-high Precision class.

Following this collaboration, we want to use this method of measurement for:

- The analysis of the drift of our existing equipment and the identification of key contributors,
- The implementation of compensation and / or corrections of these phenomena in order to improve their performances.

Fig. 3. Measurement obtained with the experimental set-up. The requirement of fluctuations below 2 ps peak to peak over 1 day is fulfilled.
B. Analysis and identification of main time drift contributors

1) Analyze and identify time drift contributors

The analysis of the results obtained by applying the time drifts measurement method to commercial devices (Master GFT3001 and GFT1004 slave units – very similar to LMJ synchronization master and slave currently under development) allows us to reveal the main contributors to temporal drifts of synchronization channels and the most sensitive parts of optoelectronic equipment.

2) Presentation of setup and configuration

Fig. 4 displays the experimental setup applied to these devices during 1 month and TABLE II. presents configuration setup main characteristics.

- Temperature in the lab is 21 ± 1°C (see Fig. 5),
- Nonstop measurement from 12/19/2014 to 01/20/2015,
- Devices under test:
  - Master unit: specific GFT3001,
  - Slaves units: 2x GFT1004 and 1x LIL slave.
- Two optical fiber length used to determine by comparison thermal effects:
  - 300 meters for two synchronization slaves (maximum length for optical fibers on the LMJ facility),
  - 5 meters for the third device.
- Observed delays are μs delays.

TABLE II. CONFIGURATION SETUP CHARACTERISTICS

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Temperature</th>
<th>Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTR 100 (4 classes)</td>
<td>TCD 842</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>&lt;0.01°C</td>
<td>5 ps</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>&lt;0.02°C</td>
<td>3 ps</td>
</tr>
<tr>
<td>Average</td>
<td>NO</td>
<td>YES (203)</td>
</tr>
<tr>
<td>Acquisition period</td>
<td>60 s</td>
<td>200 ms</td>
</tr>
<tr>
<td>Nb of acquisition</td>
<td>45 875</td>
<td>13 889.36</td>
</tr>
</tbody>
</table>

3) Results and analysis

Fig. 5 and TABLE III. present the results we obtained regarding delays and temperature variations.

We clearly observe a thermal effect on measured delays, that will be discuss below.

Fig. 5 presents the maximum time drift values obtained for three different sliding durations corresponding to the time drift requirements (cf. TABLE I.

In a first approach, we can explain major part of the time drift measurement gap observed between GFT1004 SN104 and GFT1004 SN100 channels by the difference of the optical fiber lengths (300 meters vs 5 meters). Actually, Thermal Coefficient of Delay (TCD) for this reference of optical fiber has been qualified at 7.8 ppm/°C. Numerical application for a ~300 meters optical (measured transit time @1550 nm is 1.98 μs) fiber leads to a 15.4 ps/°C effect, to convolve with a 2°C maximum thermal variation in the lab resulting in maximum of 30.8 ps gap between these two synchronization channels. This theoretical result is close to measurement results (difference of 26.1 ps during 24 hours).

Regarding the gap observed between GFT1004 and LIL precise slave synchronization channels, we can notice two significant points:

- Contrary to GFT1004 slave, LIL precise slave integrates polynomial corrections to compensate thermal variations,
- In LIL precise slave case, electronic components were specifically selected and sorted regarding their superior characteristics.

TABLE III. MAXIMUM TIME DRIFT RESULTS

<table>
<thead>
<tr>
<th>Observation period</th>
<th>GFT1004 SN104 (300 m opt. fiber)</th>
<th>GFT1004 SN100 (5 m opt. fiber)</th>
<th>LIL precise slave (300 m opt. fiber)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 hours</td>
<td>60.0 ps</td>
<td>33.9 ps</td>
<td>27.8 ps</td>
</tr>
<tr>
<td>7 days</td>
<td>80.6 ps</td>
<td>51.3 ps</td>
<td>41.4 ps</td>
</tr>
<tr>
<td>1 month</td>
<td>86.5 ps</td>
<td>55.2 ps</td>
<td>43.5 ps</td>
</tr>
</tbody>
</table>
4) Apply compensation / correction

Thanks to this analysis, we want to identify most critical functions of our devices and try to implement technical solutions in order to significantly reduce drift.

At this stage of our studies, the main contributor clearly identified is temperature variations; the next step is quantifying what would be the final drift when temperature effects are suppressed numerically.

First, we determinate if the correlation between temperature variations and time drift is equivalent to a linear behavior. For example, Fig. 6 presents the linear regression applied to the LIL precise slave T0 channel.

In this example, we obtain a R correlation coefficient equals to 0.94 (R² determination coefficient equals to 0.89 as shown on Fig. 6) – which is a trustable value in a first approach - and a thermal coefficient of this whole synchronization channel of 20 ps/°C.

In a second time, we use this basic linear model to predict the results we would obtain by suppressing thermal effects on delays stability. Fig. 7 presents the estimation of residual time drift after the thermal effects have been suppressed.

| TABLE IV. RESIDUAL TIME DRIFT VALUES (AFTER SUPPRESSION OF THERMAL EFFECTS) |
|------------------------|------------------------|------------------------|------------------------|------------------------|
| Observation Period | Residual Time | Thermal Correction | $\Delta t_{0}$ | $\Delta t_{1}$ | Residual Ref. | LIL |
|------------------------|------------------------|------------------------|------------------------|------------------------|
| 24 hours            | <2 ps                  | ND                      | 60.0 ps                 | 33.0 ps                 | 27.0 ps                 |
| 7 days               | <10 ps                 | YES                     | 33.0 ps                 | 13.0 ps                 | 20.0 ps                 |
| 1 month              | <50 ps                 | ND                      | 50.5 ps                 | 30.0 ps                 | 20.5 ps                 |

IV. PRESENTATION OF THE GLOBAL ARCHITECTURE

Fig. 8 shows the LMJ Timing and fiducial system global architecture intended to equip the facility in full configuration (22 Laser Chains). Outstanding points are listed here:

- **Master unit:**
  - An external reference clock can be used to improve long term stability,
  - A second unit ensures a hot redundancy: both units are permanently connected to the OTN through an optical coupler. Thus, the shift between them is almost instantaneous.

- **Optical Time Network** is composed of:
  - Single mode 24 way optic couplers : one coupler per Laser Chain in order to make easier integration phases,
  - Single mode optical fibers selected for their low thermal drift: Thermal Coefficient of Delay is around 8 ppm/°C.

- **Slave units** are declined into 5 definitions:
  - Standard Electrical Slave,
  - Standard Optical Slave,
  - Precise Slave (electrical only),
  - Ultra-high Precise Slave (electrical only),
  - Marker Slaves (electrical only).

For each type of slave, major specifications (jitter, number of channels and delay range) are mentioned in Fig. 8.
PETAL Timing system [6]:

- PETAL is a 10 multi-PW picosecond Laser (E > 3kJ ; I > 1020 W/cm² on target) integrated to LMJ facility,
- LMJ timing system must allow to manage PETAL synchronization devices in anticipation of coupled shots
- A LMJ fiducial channel is dedicated to mark PETAL laser diagnostic in order to establish a time correlation between these two lasers.

V. CONCLUSION

In a first time, this article offers a brief review of the main characteristics of the LMJ timing and fiducial system.

The use of the picosecond time drift measurement developed in 2014 in collaboration with the Besançon Observatory applied to our synchronization equipment allows us to identify the main contributors and imagine technical solutions to implement in order to improve performance of our timing systems.

We present this process applied to opto-electronic solution used for Standard and High Precision classes and detail some preliminary results.

In this study, our new goal is to improve the model of opto-electronic channel type synchronization facing a temperature variation to refine our estimates of thermal drift. The next step will be to compare these predictions with experimental results obtained on an opto-electronic prototype "insensible" to temperature changes currently under development. The results are expected in Q2 2015.

Finally the overall architecture of LMJ timing and fiducial system is presented and the main characteristics are detailed.

REFERENCES

Abstract - The Network Time Protocol (NTP) is commonly utilized to synchronize computer clocks in packet-switched, wide area networks (WANs) such as the public Internet. The delay asymmetry in WANs, often due to inconsistent routing and/or bandwidth saturation, is usually the dominant source of error. It typically limits NTP time transfer uncertainty to about one millisecond. This paper discusses the uncertainty of NTP time transfer when network asymmetry is largely eliminated. We performed NTP measurements over a local area network (LAN) when both the server and client are referenced to a common clock. Three variations of a LAN are tested, including a direct connection between the server and client with an Ethernet crossover cable. The elimination of network asymmetry reveals other uncertainty sources that serve as practical limitations for NTP time transfer, including client instability, asymmetry in network interface cards, and server instability.

Key words—local area network, network time protocol (NTP), time transfer

I. INTRODUCTION

The Network Time Protocol (NTP) [1, 2] has been utilized for several decades to synchronize computer clocks in packet-switched, variable latency networks. Client software for NTP is built-in to numerous operating systems and Internet devices, and NTP servers around the world now receive many billions of timing requests per day [3].

NTP is typically thought of as a “commodity” source of time, and not as a high accuracy vehicle for time transfer. This is primarily due to the fact that most NTP users utilize the public Internet, typically with the intent of synchronizing their computer clocks to the nearest integer second, an objective that is easily accomplished. Thus, the uncertainties of NTP time transfer are not often studied by time metrologists, and examples of published measurement data provided by timing laboratories are relatively rare. However, a few published studies have recently appeared [4, 5, 6, 7, 8], mostly focusing on the large delay asymmetries of the public Internet, although brief consideration of local area networks (LANs) was provided in [8]. This paper expands upon these earlier studies by focusing entirely on measurements where the delay asymmetry from the network itself has been largely eliminated (through the use of LANs), allowing us to study the other sources of uncertainty that limit NTP performance.

II. MEASUREMENT SYSTEM

The NTP measurements were performed by using a commercially-available NTP server (Symmetricom S350).* Client software developed at the National Institute of Standards and Technology (NIST) was used to initiate the NTP requests and to record the measurements. The server, client, and measurement method are described in the following sections.

A. Description of Server

The NTP server has sufficient bandwidth, according to the manufacturer’s specifications, to handle 7,000 requests per second. The server clock was a rubidium oscillator that was disciplined to signals from the Global Positioning System (GPS) satellites via a 12-channel L1 band receiver, with an antenna mounted outdoors on the roof. The accuracy of this clock was measured at NIST and found to be within 100 ns of Coordinated Universal Time (UTC).

The server has four dedicated and isolated Ethernet ports. One port is Gigabit Ethernet (not used in the experiment). The three ports that are used in the experiment are 10/100Base-T connections, capable of transmission of speeds of up to 100 megabits/s. The three ports were configured to automatically negotiate the transmission speed, and were used to connect to three variations of LANs, as discussed in Section III.

B. Description of Client

The client computer’s microprocessor was an Intel Quad-Core running at 3.4 GHz.* The client computer had 8 GB of random access memory (RAM) and ran a 64-bit version of the Microsoft Windows 7 operating system (OS).* The client software was referenced to an internal clock board with 0.1 µs (100 ns) resolution. The client clock was continuously synchronized by a 1 pps (pulse per second) output signal from the GPS clock in the server, allowing us to conduct a “common-clock” experiment. The cable connecting the server clock to the client clock had a calibrated (and compensated) delay of 0.017 µs (17 ns), which was smaller than the system’s measurement resolution. The client software, previously described in [5], had enough channels to sequentially measure up to 20 NTP servers. For this experiment, three of the available 20 channels were utilized to measure three different LAN configurations, as will be described in Section III.
To reduce the effects of client latency due to system management interrupts (SMIs), the computer’s basic input output system (BIOS) was configured to disable some of the power performance settings, including: SpeedStep, Cstates Control, and Turboboost. In addition, the on-motherboard network port was disabled, and all software services not necessary for this experiment were shut down or uninstalled.

Two Intel gigabit-capable PCIe dual-port network interface cards (NICs) were installed in the client computer and used for this experiment.* The NICs were configured as shown in Table I, to help reduce the variation in incoming and outgoing packet delays.

### TABLE I. NETWORK INTERFACE CARD CONFIGURATION

<table>
<thead>
<tr>
<th>Network Card Parameter</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interrupt Moderation Rate</td>
<td>Minimal</td>
</tr>
<tr>
<td>Flow Control</td>
<td>Disabled</td>
</tr>
<tr>
<td>Tx/Rx Buffers</td>
<td>512</td>
</tr>
<tr>
<td>Jumbo Packet</td>
<td>Disabled</td>
</tr>
<tr>
<td>Receive Side Scaling (RSS)</td>
<td>Enabled</td>
</tr>
<tr>
<td>Power Management</td>
<td>Disabled</td>
</tr>
<tr>
<td>Log Link State</td>
<td>Disabled</td>
</tr>
<tr>
<td>Wait for Link</td>
<td>Off</td>
</tr>
<tr>
<td>QoS Packet Scheduler</td>
<td>Disabled</td>
</tr>
<tr>
<td>Adaptive Interframe Spacing</td>
<td>Disabled</td>
</tr>
<tr>
<td>Large Send Offload</td>
<td>Enabled</td>
</tr>
</tbody>
</table>

C. Measurement Method

The client requested a timing packet from the server every 10 s. The request was made by sending a 48-byte packet via the user datagram protocol/Internet protocol (UDP/IP) to port 123. The last eight bytes of the packet included the time of the request, \( T_1 \), as obtained from the client clock (which is referenced to the same GPS clock as the server).

The server responded to the timing request by returning a data packet. The entire packet is decoded by the client software, including three 64-bit time stamps. These time stamps utilize 32 bits to represent integer seconds, and an additional 32 bits to represent fractional seconds, providing a resolution of \( 2^{-32} \) s (233 ps).

One of the time stamps returned by the server simply echoed back \( T_1 \), the time when the client made the request (measured by the client). Two other time stamps contained the time, \( T_2 \), when the request was received by the server, and the time, \( T_3 \), when the server transmitted its response. When the client received the packet, it again queried its clock board and recorded \( T_4 \), the time of the packet’s arrival. The time difference, \( TD \), between the server and client clocks was obtained using the standard NTP equation for clock offset [1],

\[
TD = \frac{(T_2 - T_1) + (T_3 - T_2)}{2}. \tag{1}
\]

Using these same four time stamps, the round trip delay between the client and server was calculated [1] as

\[
RT_{\text{Delay}} = (T_4 - T_1) - (T_3 - T_2), \tag{2}
\]

where the time interval required for the server to process the NTP request, \( T_3 - T_2 \), is subtracted from the round trip delay measured by the client clock. Therefore, variations in server processing time did not impact \( RT_{\text{Delay}} \). For the server utilized in this experiment, \( T_3 - T_2 \) was typically about 75 µs, but can periodically be much larger (exceeding 1 ms in some cases).

The results of both the time difference and round trip measurements are updated every 10 s on the client system’s display and recorded in a file.

Note that in our configuration, both the server and client were referenced to the same clock, so that \( TD \) in Eq. (1) should theoretically be 0. Any deviation from 0 was due to NTP time transfer uncertainties. Note also that the “divide by 2” in Eq. (1) assumes that the delay from the server to the client is equal to one half of the round trip delay. If this assumption were true, the delays in the path to and from the server would be equivalent and dividing by two would fully compensate for all delays. In practice, however, the incoming and outgoing delays are not equal, and this delay asymmetry contributes to the uncertainty of NTP time transfer.

### III. MEASUREMENT CONFIGURATION

The measurement system was configured to sequentially record readings from the same server (described in II.A) using three different LAN configurations. Three consecutive measurements, one from each LAN, were recorded in the same second at 10 s intervals. This process was repeated for a period of 20 days, from 02/28/2015 to 03/19/2015. The following sections describe the three LAN configurations and Fig. 1 provides a diagram.

![Fig. 1. LAN configurations for NTP measurements.](image)

\[ \text{GPS Antenna} \]

\[ \text{Crossover (direct connection)} \]

\[ \text{LAN A} \]

\[ \text{NTP Server} \]

\[ \text{Client Computer} \]

\[ \text{Addressable switch} \]

\[ \text{Subnet A} \]

\[ \text{Router} \]

\[ \text{Subnet B} \]

\[ \text{LAN B} \]

\[ \text{LAN C} \]
A. Direct connection to server via crossover cable

A network port of the client computer was connected directly to a port on the NTP server using a Category 6 (Cat 6) crossover cable with a length of 3 m. This represents the simplest possible LAN configuration. The server port had a unique Internet protocol (IP) address to guarantee that the packets were transferred via this connection.

B. “One-hop” local area network

A second network port of the client computer was connected to a LAN port of a 10/100 Mbps router (used here as an addressable network switch). An NTP server port, again with a unique IP address, was connected to another LAN port of the router. The connections in this link were made with two Cat 6 cables, each 4 m in length. This is considered to be a data link layer (layer 2) network path.

C. “Two-hop” local area network

A third network port of the client computer was connected to a LAN on a subnet different than that of the NTP server. The packet sent by the client had to travel through a layer 2 switch and through a router (layer 3) to reach the server subnet. The Cat 5 cables utilized in this LAN configuration were preexisting parts of the NIST network and their lengths are not known.

IV. MEASUREMENT RESULTS

The following sections show the measurement results, including the server-client time difference and round trip delay, for each of the three LAN configurations. As noted in Section III, a measurement was recorded from each LAN every 10 s for 20 days (172 800 data points per LAN).

A. LAN A Measurement Results

Figure 2 shows the results of the simplest possible LAN configuration, a direct connection between the server and client via a crossover cable (Section II.A). The average time difference for the entire measurement interval was 0.9 µs with a standard deviation (STDEV) of 9.4 µs. The average \( RT_{\text{delay}} \) was 215.1 µs with a STDEV of 17.9 us (approximately 2× the STDEV of the time difference, as expected).

Because there is no network asymmetry in this configuration, and very little delay (~0.02 µs round trip) in the crossover cable, nearly all of the round trip delay and all of the uncertainty of the time measurements can be attributed to other factors, as discussed in Section V.

B. LAN B Measurement Results

The LAN B results were slightly worse than LAN A, with the average time difference increasing to 1.6 µs and STDEV increasing to 13.7 µs (Fig. 3). The average \( RT_{\text{delay}} \) increased by about 30 µs, to 244.8 µs, with a STDEV of 28.0 us. Again, the STDEV of \( RT_{\text{delay}} \) was approximately 2× the STDEV of the time difference, as expected. The slight decrease in performance with respect to LAN A indicates that a small amount of network asymmetry was introduced by the additional hardware, probably because the delay through the addressable network switch is not the same in each direction.

C. LAN C Measurement Results

For this “two-hop” LAN configuration, the NTP packets travel to a different network layer and back through a router. This introduces some network asymmetry due to routing, and also due to packet delays, because NTP packets are now required to pass through a subnet that also carried traffic from other computers located in the building (Fig. 4).
The average time difference for LAN C increased to 6.2 µs with a STDEV of 21.5 µs. The average \( RT_{\text{Delay}} \) increased by less than 5 µs with respect to LAN B, to 249.5 µs. However, the STDEV of \( RT_{\text{Delay}} \) increased to 96.2 µs, more than 4× the STDEV of the time difference, due to periods when \( RT_{\text{Delay}} \) increased substantially, briefly exceeding 10 ms (Fig. 2). However, the impact in \( RT_{\text{Delay}} \) had on the time error was attenuated, suggesting that the packets travelling in both directions were delayed by nearly equal durations as they passed through the router.

Table II summarizes the measurement results. Because the average time difference is smaller than the standard deviation, the shaded columns in Table II represent the combined time transfer uncertainty (1 standard deviation (STDEV), uncertainty estimates were obtained with STDEV and the time difference is smaller than the standard deviation, as they passed through the router.

Table II. SUMMARY OF LAN MEASUREMENT RESULTS

<table>
<thead>
<tr>
<th>LAN</th>
<th>Server-Client Time Difference</th>
<th>Round Trip Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistics (all units are microseconds, µs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>STDEV</td>
</tr>
<tr>
<td>A</td>
<td>0.9</td>
<td>9.4</td>
</tr>
<tr>
<td>B</td>
<td>1.6</td>
<td>13.7</td>
</tr>
<tr>
<td>C</td>
<td>6.2</td>
<td>21.5</td>
</tr>
</tbody>
</table>

The time transfer uncertainty can be reduced by averaging over longer intervals. To illustrate this, Fig. 5 compares the TDEV of the results obtained with all three LAN configurations. The uncertainty of LAN C is larger than the other LANs for all averaging intervals, but the time transfer uncertainty (1 σ) for all three LAN configurations is less than 1 µs after approximately 10 hours of averaging.

V. FACTORS THAT LIMIT NTP TIME TRANSFER UNCERTAINTY WHEN NETWORK ASYMMETRY IS ELIMINATED

As previously noted, network asymmetry was eliminated in LAN A, but other factors still limit the uncertainty of NTP time transfer. The three primary limiting factors appear to be client instability, asymmetries in the network interface cards, and server instability, as discussed in the following sections.

A. Client Instability

The uncertainty of the round trip delay measurement in Eq. (2) is dependent upon the client’s ability to accurately record \( T_1 \) and \( T_4 \) by reading the time from its internal clock board. This is analogous to a stopwatch measurement made by a human operator who pushes a button at the beginning and end of the time interval they are measuring. Due to variations in human reaction time, the button will always be pushed early or late by some amount. However, if the start and stop delays are equal, they will not affect the measurement accuracy. In the case of NTP, the clock readings will always have some latency, meaning that the time will always be recorded after it occurs. If the latency of the \( T_1 \) reading equals the latency of the \( T_4 \) reading, then the measurement of \( RT_{\text{Delay}} \) will be correct. If the reading of \( T_1 \) has less latency than the reading of \( T_4 \), then \( RT_{\text{Delay}} \) will be underestimated. If the reading of \( T_4 \) has more latency than the reading of \( T_1 \), then \( RT_{\text{Delay}} \) will be overestimated. The uncertainty contributed to the calculation of TD in Eq. (1) is one half of the error in the \( RT_{\text{Delay}} \) measurement.

To determine the latency of the client used in this experiment, testing software was written to run in a tight loop on the client computer and to do consecutive clock reads, with no instructions executed in between the clock reads. It was found that on average, the client could read the clock about once every 6 µs. The STDEV was about 0.3 µs, meaning that client instability did not contribute significantly to the uncertainty results summarized in Table II.

Care was taken to configure the client computer parameters to make it more stable than its default “out of the box” condition. The BIOS settings listed in Section II.B were especially effective at reducing noise. Further improvement could be realized by utilizing a real-time OS (rather than a general purpose OS such as Windows 7*) and an SMI-free BIOS.

B. Network Interface Card Asymmetry and Server Instability

To attempt to determine the asymmetry of the client’s NIC, the latency testing software was modified to execute a UDP send command in between two successive clock reads. This software sent a standard 48-byte NTP packet through the NIC, but did not wait for a response from an NTP server. The UDP send operation added a delay of 10 µs, increasing the interval between clock readings to about 16 µs. The STDEV doubled, to about 0.6 µs, but was still an insignificant part of the uncertainty values summarized in Table II.
The latency testing software was again modified to wait for an NTP packet to return after the UDP send command was executed. This test is now a measurement of the full round-trip delay, and the server processing time, ($T_3 - T_2$), was subtracted per Eq. (2). Short tests of this UDP send/receive sequence resulted in a STDEV of ~15 µs, or similar to the result of 18 µs shown in Table II for the entire 20-day period.

If we assume that accounting for server processing time removes the instability of the server (or at least reduces it to a level similar to the sub-microsecond uncertainty of the client), then the large increase in STDEV from 0.6 µs to about 15 µs between the “UDP Send” and “UDP Send/Receive” tests suggests that the dominant source of uncertainty is due to delays introduced by either the NIC of the client or the NIC of the server when a packet is received. For example, delays in the packet received by the client would occur before the client records $T_4$. Delays introduced in the packet received by the server would occur before the server records $T_2$.

Our tests were unable to determine whether the client or server NIC was the dominant source of NTP time transfer uncertainty. However, the client NIC is a more likely candidate, as it was designed for general purpose network applications. The server NIC is more likely to be optimized (balanced) for time transfer applications.

VI. SUMMARY

We have measured NTP time transfer performance via three different LAN configurations. The results indicate that the measurement uncertainty (1 σ) of a single NTP timing request when utilized over the simplest possible LAN configuration is less than 10 µs; or at least two orders of magnitude smaller than the uncertainty of a single time request made via the public Internet.

Our measurements have shown that when network asymmetry is eliminated, other (much smaller) sources of uncertainty are revealed, including client instability, asymmetry in network interface cards, and server instability. The delays introduced by the network interface cards when receiving packets appear to be the dominant source of uncertainty. These uncertainties could be further reduced with hardware and software that is better optimized for time transfer, but they can perhaps be considered as the practical limit of NTP time transfer for most applications.

* The measurements were conducted using commercially-available hardware and software products that were made available to the authors. These products are mentioned for technical completeness, but this does not imply endorsement by NIST. Other products may work equally well or better.

REFERENCES

Precise Three-Channel Integrated Time Counter

Ryszard Szplet, Paweł Kwiatkowski, Zbigniew Jachna, Krzysztof Różyć
Department of Electronics
Military University of Technology
Warsaw, Poland
rszplet@wat.edu.pl

Abstract—We present a design, FPGA-based implementation and test results of a new three-channel time interval counter developed for a project called Legal Time Distribution System (LTDS). The main aim of the counter is to gather information about time drift of clocks involved into the LTDS, then to evaluate their stability, and finally to select the most stable one as a local clock. The counter provides a high measurement precision (< 15 ps) and wide range (> 1s) that are obtained by combining period counting with two-stage time interpolation. The time counter is implemented in a universally available and relatively cheap programmable device from family Spartan-6 (Xilinx).

Keywords—multichannel time counter, time-to-digital converter, two-stage interpolation, programmable device

I. INTRODUCTION

Various users in many countries are still more often obliged by local legislation for applying of a legal time in their activities. Typical way to achieve the legal time is based on synchronization of locally created time scales to national time scale generated by National Institute of Metrology (NIM). Quality of a local time scale and its ability for autonomous operation (unlocked to NIM) depends on the stability of clock sources involved. This stability can be verified through the measurement of clock’s drift performed with the use of precise time counter. We propose the three-channel, high-precision, integrated time counter for a system of local time scale based on three clocks, which is developed within LTDS EUREKA project.

Atomic clocks, which are commonly used as clock sources for creation of time scales, most typically generate 10 MHz sinusoidal waveform and 1PPS (Pulse Per Second) signal strictly related to the former one. If the clocks stability is evaluated by measurement of time relation between 1PPS signals, the involved time counter has to provide wide measurement range of at least 1 s and high precision at picosecond level.

Precise time interval measurements within a wide range are currently performed with the aid of interpolation methods that employ the classical counting method and inner time interpolation [1]. To evaluate the time interval between just two events the use of one of known interpolation methods in start-stop mode is enough [2]. However, if simultaneously observed events come from more than two sources, the continuous measurement based on the timestamps method is recommended [3].

II. MEASUREMENT METHOD

The timestamps measurement method employed in the time counter combines continuous period counting and two-stage in-period interpolation to create a very precise time scale (Fig. 1) that is common for all channels of the counter. Any input event is identified by the counter on the time scale and registered with a unique timestamp $TS$ [3]. To obtain consistent time scale for all observed events the scale is created by permanent counting edges of a reference clock signal in period counter. The occurrence of event causes registering the current state of the period counter. In such a way the coarse part of the timestamp is created, however the resolution of this measurement is rather poor and limited to the period of reference clock ($T_0$).

![Fig. 1. The principle of timestamps method.](image)

Further improvement in the resolution is obtained by involving two-stage time interpolation within a single period. In a first stage of interpolation the four phase clock is created by using both edges, rising and falling, of the two reference clock signals shifted in phase to each other by quarter of period ($T_0/4$). Such multiphase clock is used in the first interpolation stage (FIS in Fig. 2), which identifies between which edges of the clock an input event appears. The four phase clock gives fourfold improvement in resolution in comparison to simple timestamps method based just on a period counter. The result from the FIS is stored as $T_{fine1}$ Value.
The next measurement step is aimed to evaluate more precisely a relatively short time interval ($T_{\text{fine2}}$) between rising edge of the input event and the nearest edge of the four phase clock. This is performed in the second interpolation stage (SIS) where each phase segment of the four phase clock is accurately quantized.

Finally, the value of the example timestamp $T_{S1}$, related to the first input event (Fig. 1), is calculated as:

$$T_{S1} = (N-1) \cdot T_0 - (T_{\text{fine1}} + T_{\text{fine2}}).$$

The two-stage time interpolation allows to expand common time scale resolution twice, whereas the timestamps method gives possibility to measure time intervals between any registered events. Additionally, an improvement in resolution due to the use of two-stage interpolation causes increase in precision of time interval measurement, which can reach even picosecond level.

### III. COUNTER DESIGN

The designed time counter was implemented in a popular and low-cost Spartan-6 FPGA (Field Programmable Gate Arrays) device manufactured by Xilinx [4]. A block diagram of the time counter is shown in Fig. 2. It consists of three independent measurement channels. Each channel contains FIS and SIS modules, code converters, channel register with its synchronizer and code processor. As a timebase clock the signal of 500 MHz frequency is applied. It is generated outside its synchronizer and code processor. As a timebase clock the FIS and SIS modules, code converters, channel register with independent measurement channels. Each channel contains.

The occurrence of any input event or, in our application, any measured clock edge (CLK 1, CLK 2 or CLK 3) is recorded in the FIS. The FIS involves build-in two-phase clock generator that shifts reference clock by about 500 ps (quarter of period). Such delay is obtained using short delay line based on FPGA programmable logic elements. Then the set of double synchronizers operating with both clock edges, identifies clock phase segment in which the measured edge appeared. Synchronizer is used to virtually eliminate metastability effect that can occur between two asynchronous input signals (measured CLKs and reference clock) of channel register. Finally, the FIS produces two signals, i.e.: information about the measured value ($T_{\text{fine1}}$) and identified clock phase edge. The latter signal is then used in the SIS and synchronizer for further conversion phases.

In the SIS a time coding delay line is used [1]. It can be relatively easily implemented in Spartan-6 device employing dedicated carry chain structure [5]. Such structure is normally utilized for fast arithmetic operation. However, due to the short propagation time of subsequent delay elements involved in the chain and direct connections with related flip-flops, it fits well also to implement time coding delay lines. During a measurement process the measured edge of clock (e.g. CLK1) propagates through the carry chain. When the clock phase edge identified in FIS appears at the SIS input, the related flip-flops register current state of the chain. The last flip-flop, which registered high voltage level (or logic state ‘1’) informs about the length of time interval $T_{\text{fine2}}$. Such value is calculated as $T_{\text{fine2}} = \tau \cdot n$, where $n$ is the number of last flip-flop and $\tau$ denotes the propagation time of a single carry chain delay element (carry chain multiplexer in Spartan-6).

The common time scale is produced by period counter that works continuously. It is 30-bit binary counter that operating with 500 MHz clock frequency provides 1.07 s time scale range. After that time counter overflows and starts counting from the beginning. Taking into account that 1PPS signals are measured, the range of time scale created in a single cycle of the counter is wide enough. A current state of the period counter is rewritten to the buffer at each edge of the reference clock. The use of buffer improves timing margins for data saving to channel register.

The clock phase edge, identified in FIS as a subsequent after input event, is synchronized to the reference clock in the synchronizer, and then use for latching the buffer state in the channel register. Results from the channel register, FIS and SIS are then transferred to the code processor. The code processor corrects measuring data using actual transfer characteristic of interpolator collected in internal memory during calibration process and performed obligatory in the beginning of each measurement session. The operation of data correction, carried out on-the-fly in the measurement mode, evaluates data as a 12-bit fraction $T_i$ of the clock period $T_0 = 2$ ns. Finally, the stream of timestamps $T_i$ is formed as

$$T_i = T_0 \cdot (N + T_{\text{fine2}} / 2^{12})$$

where $k$ denotes the identity number of channel.
To achieve a high precision of timestamps the calibration procedure of interpolators must be executed repeatedly and frequently enough. There are two main steps in the time counter calibration procedure: (1) statistical code density test \[6\] - to identify accurate transfer characteristics of interpolators without influence of nonlinearity error, and (2) measurement of input channels time offset – to adjust the common time scale in each measurement channel.

All calibration tasks are performed under the control of FSM (finite-state machine) unit implemented in spare logical resources of the same FPGA device. The user of counter is allowed to start calibration and check the flag signals at any time.

A communication with the counter is provided via USB interface. We applied well-known interfacing circuit FT2232H (FTDI) as a converter of USB 2.0 Hi-Speed (480Mb/s) to internal FIFO memory. The main advantage of this solution is availability of the drivers for several operating systems, i.e. Windows, Linux, Mac OS X, and Android.

The control and diagnostic software executed by external computer may thoroughly control the measurement process and save the obtained data. The software, written for Windows, allows for example for: selecting the calibration procedure or the measurement mode, switching on/off selected measurement channel, setting input threshold voltage and signal polarization, and presenting measured results.

![Fig. 3. Snapshot of the control and diagnostic software.](image)

### IV. MEASUREMENT TESTS

The time counter precision was tested for time intervals within the range up to 1 s using two reference time interval generators: Model 745 from Berkeley Nucleonics Corporation (for time intervals up to 2 ms) and T5200U from Vigo Systems (for longer time intervals). The chosen test signal sources give the best precision (the lowest jitter of generated time intervals) in preselected ranges. All instruments, the time counter and time interval generators, were connected to separate rubidium standard generators FS725 (Stanford Research System) that provided the reference stable 10 MHz clock signal.

The first measurement channel (CH I) of the counter was used to register timestamps for START pulses and the second (CH II) and third (CH III) ones for STOP pulses. Each channel collected at least 1000 timestamps. The measured time intervals were obtained as a difference between timestamps from channels CH II or CH III, and related timestamps from channel CH I. The counter precision was calculated as a standard deviation of 1000 measurements and it is shown in Fig. 4.

![Fig. 4. Time counter precision.](image)

The described time counter provides the precision of about 15 ps for both tested channels configuration (CH I - CH II and CH I - CH III). The decrease of precision observed for time intervals within the range from 5 ms to 100 ms comes from a slightly higher timing jitter of generator T5200 than BNC745 used for shorter time intervals. The precision for time intervals longer than 100 ms is deteriorated by limited stability of the reference clock sources.

### V. CONCLUSION

A three-channel time interval counter implemented in a standard FPGA device is described. The principle of its operation is based on collecting electronic timestamps related to consecutive input events. Since these timestamps locate events at the common time scale very precisely, then they allow to calculate time intervals that elapse between freely selected events. The use of reference clock period counting provides wide measurement range (> 1 s), whereas involving the two-stage in-period interpolation results in a high measurement resolution (~19 ps) and precision (< 15 ps). Thanks to the use of programmable device the counter can be customized to match the user application at reasonable cost and short time.

### ACKNOWLEDGMENT

This work was supported by the Polish National Centre for Research and Development under contract no. E! 8727 LTDS/1/2014

### REFERENCES


A fiber link for the remote comparison of optical clocks and geodesy experiments

Clivati Cecilia, Calonico Davide, Frittelli Matteo, Mura Alberto, Levi Filippo
Physic Metrology Division
Istituto Nazionale di Ricerca Metrologica-INRIM
Turin, Italy
Email: c.clivati@inrim.it

Abstract—Differential comparisons of distant optical clocks with uncertainties in the $10^{-18}$ level linked by phase-stabilized optical fibers can be used to probe the Geoid potential at the centimeter sensitivity. Here, we present the realization and characterization of a coherent optical link that will enable gravitational potential measurements on a baseline of 90 km and at 1000 m difference in altitude. This is a proof-of-principle experiment that aims at exploring extensive interconnections between geodesy and frequency metrology.

Index Terms—coherent optical links; gravitational red-shift; comparison of optical clocks

I. INTRODUCTION

Optical clocks have demonstrated the $10^{-18}$ level of uncertainty [1], [2]. This is an important achievement for frequency metrology, as it paves the way for a re-definition of the SI second on optical transitions. However, such high accuracies, together with the capability to perform ultrastable frequency transfer over long optical fiber links [3]–[6] open a number of opportunities in geodesy as well.

Since the early experimental demonstration of the gravitational red-shift of clocks in different values of gravity potential, with a relative frequency sensitivity of $\Delta \nu / \nu = 10^{-16} / m$ [7], [8], all clocks involved in the International Atomic Time (TAI) generation apply a correction proportional to their elevation on the Geoid [9], [10]. The gravitational shift uncertainty ranges from $5 \times 10^{-18}$ to $20 \times 10^{-18}$, thus optical clocks have already been used to observe the gravitational red-shift over few tens of centimeters on a laboratory scale [11]. The quest now is for long-baseline and fast measurements of the gravitational potential; the short time needed to achieve a resolution at the 10 cm level makes optical clocks a useful tool to measure gravitational potential fast variations. On one hand, this could allow a better knowledge of the Geoid fast motion and an improvement of global models; on the other hand, this is a prerequisite for future timescales, as the gravitational red-shift significantly contributes to the accuracy of any clocks comparison.

Europe has a large variety of optical clocks and is developing long-baseline optical fiber links to compare them. On this basis, the European Metrology Research Programme (EMRP) “International Timescales on Optical Clocks” [12] has planned a proof-of-principle geodesy experiment, where two clocks located at about 1000 m difference in height will be compared on a baseline of 90 km. The experiment will involve the Yb lattice clock developed at INRIM [13] and a portable Sr lattice clock developed at PTB. The two will be first compared locally, then the PTB Sr clock will be moved, together with a portable optical comb developed at NPL [14], at the Laboratoire Souterraine de Modane (LSM). This laboratory is placed inside the Frejus mountain on the Italy-French border, at a geographical distance of 90 km and 1000 m higher. The frequency ratio $\nu_{\text{Sr}}/\nu_{\text{Yb}}$ will be measured in an indirect way, by comparing both clocks to an ultrastable transfer laser in the IR domain, disseminated via a phase-stabilized optical fiber. LSM main activity is in the field of particle physics; therefore, also a 100 MHz RF signal will be disseminated by INRIM through a second optical fiber, to ensure a proper frequency referencing of the clock setup and optical comb.

In the following section we will describe the experimental setup and the characterization of the two links.

II. OPTICAL FREQUENCY DISSEMINATION

The transfer laser (TL) is a 1542 nm, cavity-stabilized laser with a residual instability of $8 \times 10^{-15}$ at 1 s and a frequency drift typically $\leq 10^{-15}$/s [15]. This may affect the clocks comparison if the measurements in the two laboratories are not synchronized; in particular, a frequency bias $\delta = d \tau$ could arise, where $\tau$ is the delay in the measurements and $d$ is the TL frequency drift. To overcome this effect, we directly lock the TL to an hydrogen-maser by using an optical frequency comb. The control-loop is implemented via software, as the TL and the optical comb are in two different laboratories. The control bandwidth is 0.05 Hz, that is a compromise between an adequate rejection of the maser noise and a sufficiently tight locking; the correction is applied once per second with an Acousto-Optic Modulator (AOM). The residual drift is $\leq 10^{-19}$/s, which is negligible even for $\tau = 1$ s for the purpose of this experiment.

The optical fiber connecting INRIM to LSM is 150 km long and is implemented on a Dense- and Coarse-Wavelength-Division-Multiplexed (DWDM & CWDM) architecture: a single channel is dedicated to this experiment, while other channels are occupied by Internet traffic. 6 Optical Add/Drop Multiplexers have been used to route the coherent signal along the link. This infrastructure has been characterized by looping the fiber at LSM, in order to realize a 300-km link with both
The optical phase-stabilization is based on the typical Doppler noise-cancellation scheme and is shown in Fig. 1. The ultra-stable laser is split into two beams: one is used as a local oscillator to detect the fiber noise, the other is sent into the link. At the remote end, it is frequency-shifted by AOM2, then a small part is reflected back to INRIM along the same fiber and compared to the original beam on photodiode PD1. We track this beatnote with two independent voltage-controlled-oscillators, in order to improve the signal to noise ratio (SNR) and to properly detect the occurrence of cycles-slips. The additive phase noise introduced by the fiber in the double pass is detected and compensated with a Phase-Locked Loop (PLL) acting on AOM1. The original and the delivered radiation are then compared on PD2, in order to determine the link performance.

The total loss for this loop is 120 dB. 5 bidirectional Erbium Doped Fiber Amplifiers (b-EDFA) are used to partly compensate for it; the gain is critically set to minimize the stray reflections and Amplified Spontaneous Emission build-up, and is always limited to <18 dB. To guarantee a proper SNR on the control beatnote on PD1, we also use a Fiber Brillouin Amplifier (FBA) [16] at the remote link end, providing a gain of ~10 dB; this is limited by the high injection losses of the DWDM network, which result in a rather low coupled pump power (~2.5 mW). The b-EDFAs remote control is based on Ethernet connection provided by the network manager, whilst the FBA is directly controlled inside the laboratory.

The residual link loss amounts to 30 dB; in this configuration, it is always possible to maintain a SNR >27 dB in 100 kHz bandwidth on PD1, which guarantees a cycles-slips rate <10^{-4} /s.

The phase noise of the delivered optical carrier is shown in Fig. 2. A broad peak is noticed on the un-stabilized link at frequencies between 10 Hz and 20 Hz, and has been attributed to building vibrations and traffic [17]. The compensation bandwidth is set by the round-trip time and approaches 170 Hz for the looped link. When the phase-stabilization is engaged, the residual link noise achieves the ultimate limit imposed by the delay [18].

The phase-difference between the original and the delivered radiation has been counted on a dead-time free phase/frequency counter [19]; the resulting frequency instability is shown in Fig. 3. The graph shows the free-running link instability (red diamonds) and the stabilized-link instability, both when the counter is operated in the so-called Π-mode (blue circles) and averaging mode (black squares). The Π-mode acquisition is equivalent to a full-bandwidth sampling of the optical phase; this is impractical for most clocks comparisons involving fiber links, as the signal delivered across the optical fiber is often deteriorated by the delay-unsuppressed fiber noise up to a bandwidth of ~1 kHz [18]. A more suited solution is to filter the beatnote either in the electronic or in the digital domain. In fact, the averaging of the optical phase acts as a digital low-pass Lorentzian filter whose cutoff frequency depends on the averaging factor. Although sharper digital filters should be preferred for very noisy fibers [5], this is not the case for this optical link and averaging phase data leads to the same results we obtained by filtering full-bandwidth data on a 1 Hz measurement bandwidth.

The stabilized link instability achieves the 3 × 10^{-19} level in few hours of measurement and no frequency bias have been observed at this level. Hence, the optical link does not affect an optical clocks comparison for measurement times longer than 10 s.

III. Radio-frequency Dissemination

In order to properly operate the Sr clock apparatus (frequency shifters and counters) and the optical comb, high-
quality RF signals are needed at LSM. GPS-disciplined oscillators are not an option, being LSM a subterranean laboratory. Hence, we choose to deliver such references via a second optical fiber, in the same bundle but independent on the one used for the optical frequency dissemination. The specifications on instability for RF dissemination are less stringent than those for the optical signal dissemination, since none of the referenced instruments directly contributes to the total uncertainty of the clocks comparison.

The scheme used to deliver the 100 MHz RF signal is depicted in Fig. 4. The same optical path as for the optical dissemination has been used for the characterization of this loop. A free-running fiber laser is amplitude-modulated at 100 MHz with a Mach-Zehnder Electro-Optical Modulator and is injected into the optical fiber. At the remote end, the amplitude-modulation is detected on a fast photodiode, amplified and phase-compared to a commercial synthesizer at the same frequency, which is referenced to a 10 MHz Oven-Controlled Quartz oscillator (OCXO). The DC phase-error is integrated and directly feeds the OCXO on a bandwidth $<0.1$ Hz. This is chosen as a compromise between a proper rejection of the high-frequency noise from the fiber and a sufficiently tight

locking. Both the 10 MHz and the 100 MHz signals can be extracted from this system.

Fig. 5 shows the Allan deviation of the delivered signal as measured in the II-mode, full bandwidth, on a dead-time free phase/frequency counter [19]. On the short term it is dominated by the noise of the OCXO; at measurement times between 0.1 s and 1000 s the instability starts to be dominated by the detection noise of the optically-delivered RF; at longer averaging times the main source of instability are the optical path length variations due to environmental effects. No frequency bias have been observed at this level, and the achieved uncertainty fulfils the requirements for the RF references needed in this experiment.

IV. Conclusion

We have described the installation and characterization of the infrastructure that will be used for a proof-of-principle geodesy experiment. A Sr portable lattice clock and an optical comb will be placed at LSM, inside the Frejus mountain, and compared to an Yb clock at INRIM via a phase-stabilized optical fiber link. The residual instability of this link achieves the $3 \times 10^{-19}$ level in few hours of operation, which is compliant with the requirements for this experiment. Stable RF signals will also be delivered via an un-stabilized, amplitude-modulated optical signal travelling in a second optical fiber. The characterization of both links has been pursued by connecting the two parallel fibers at LSM, to form a loop with both ends at INRIM. The results here presented thus place an upper limit to the real performance of both links.

This link extends the present fiber backbone for optical frequency dissemination in Italy [5] and represent a possible way to access the European network for frequency metrology currently under development.

ACKNOWLEDGMENT

The authors thank the GARR Consortium and the TOP-IX Consortium for technical help with the fibers. This work was supported by the European Metrology Research Programme (EMRP) under SIB-55 ITOC. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.
REFERENCES


OPTIME - the system grows - a new 330 km line

L. Buczek, J. Kołodziej, P. Krehlik, M. Lipiński, Ł. Śliwczycyński
AGH University of Science and Technology
Krakow, Poland
mlipinsk@agh.edu.pl

P. Dunst, D. Lemański, J. Nawrocki, P. Nogaś
Astrogeodynamic Observatory
Space Research Centre PAS, AOS
Borowiec, Poland
nawrocki@cbk.poznan.pl

A. Czubla
Central Office of Measures, GUM
Warsaw, Poland
a.czubla@gum.gov.pl

A. Binczewski, W. Bogacki, P. Ostapowicz, M. Stroiński, K. Turza
Poznan Supercomputing and Networking Center, PSNC
Poznan, Poland
wojbor@man.poznan.pl

W. Adamowicz, J. Igalson, T. Pawszak, J. Pieczerak
Orange Polska, TPSA
Warsaw, Poland
Janusz.Piecezerak@orange.com

M. Zawada
Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Torun, Poland
zawada@fizyka.umk.pl

Abstract— The OPTIME project creates an ultra-precise time and frequency signals dissemination system based on telecommunication networks. End users obtain access to these signals without incurring huge costs for the purchase of their own atomic clocks, and receive the service related to laboratories generating international atomic time scales, to which any precise time must be referred. OPTIME dissemination system is based on three main elements: reference time and frequency laboratories, local time and frequency repositories and fiber optical network with specialized transmission equipment to transfers signals between laboratories, repositories and end users. This article describes OPTIME system with particular emphasis on a new 330 km long dissemination line between Space Research Centre PAS, Astrogeodynamic Observatory (AOS) at Borowiec and National Laboratory of Atomic, Molecular and Optical Physics (KL FAMO) at Torun.

Keywords— Atomic clock, fiber optical network, high precision dissemination of time and frequency reference signals, time and frequency transfer, local repositories.

I. PROJECT OPTIME

The OPTIME project started in December 2012. The main goal of the project is to design and ultra-precise time and frequency signals dissemination system and create a demonstrator of such system in wide area optical network.

The dissemination system is based on the three main elements:

- reference time and frequency laboratories – which provides time and frequency signals to the dissemination network,
- local time and frequency repositories – which are responsible for providing time and frequency signals during connection failure to reference time and frequency laboratories,
- fiber optical disseminating network – which is equipped in specialized transmission devices to transfers signals between reference laboratories, local repositories and end users systems.

In the first period of the project there was created 420 km long optical link which connected two UTC laboratories in Poland – Astrogeodynamic Observatory (AOS) in Borowiec which provides UTC(AOS) and the Central Office of Measures (GUM) in Warsaw, which provides UTC(PL).

During the project the dissemination system grew and was updated. The new version of transmission devices was developed. The new local repository with passive H-maser was designed, developed and installed in Poznan. In 2014, the new 330 km long dissemination line was built from Astrogeodynamic Observatory (AOS) at Borowiec to the National Laboratory of Atomic, Molecular and Optical Physics (KL FAMO) in Torun.

Now the National Distribution System for Atomic Clocks Time and Frequency Signals provides a 750 km dissemination network with two UTC laboratories and one local repository. The topology of the system is shown on Fig. 1.

The next step is to build third link between UTC(PL) laboratory and Orange Polska network synchronization center located in Warsaw-Anin. Link length is about 40 km.

The OPTIME project (no. PBS1/A3/13/2012) is co-funded by the National Center for Research and Development from Poland in the Applied Research Programme.
II. NEW 330 KM T&F DISSEMINATION LINE

The new 330 km time and frequency dissemination line became operational in December 2014. This line connects Astrogeodynamic Observatory (AOS) in Borowiec to the National Laboratory of Atomic, Molecular and Optical Physics (KL FAMO) in Torun.

The line contains 7 specialized optical Bidirectional Amplifiers, one Local Module was installed at the AOS in Borowiec and one Remote Module was installed at the KL FAMO in Torun. All these devices were designed and built by project partner AGH University of Science and Technology. The devices installed on the AOS – KL FAMO line are new versions of devices installed on Borowiec-Warsaw link. The new long-haul version of the ELTAB system is capable of compensating more than 1 µs of the fiber delay changes. Thus it may be used in very long-haul links up to about 1000 km, without any seasonal maintenance, needed for systems with smaller compensation range. The new solution is based on a hybrid setup exploiting a pair of continuously tuned electronic variable delay lines, and a set of switched optical delays. The optical delays are activated occasionally (a few times per year) in the way that the output 10 MHz and 1PPS signals are continuous and there are no observable phase jumps [1].

The choice of this line of optical fiber to create a dissemination network is not accidental. It links Borowiec Time and Frequency Laboratory which provides UTC(AOS) to The National Laboratory of Atomic, Molecular and Optical Physics (KL FAMO). The KL FAMO offers a system of two independent strontium optical lattice standards probed with a single shared ultra-narrow laser [2]. The two optical frequency standards (Sr1 and Sr2) are based on the $^3S_0 - ^3P_0$ transition in neutral strontium atoms (isotope $^{87}$Sr or $^{88}$Sr), recommended by the International Committee for Weights and Measures for practical realization of the meter and secondary representation of the second.

The two strontium optical lattice clocks operating at KL FAMO can be compared with UTC(AOS) realized by the active hydrogen maser at the Space Research Centre Astro-Geodynamic Borowiec Observatory. The absolute frequency of the $^3S_0 - ^3P_0$ clock transition in bosonic $^{88}$Sr measured in Sr2 in relation to UTC(AOS) is equal to 429 228 066 418 015 (14)syst(6)stat Hz.

**TABLE I.** LINE AOS – KL FAMO

<table>
<thead>
<tr>
<th>No.</th>
<th>Dissemination line between AOS – KL FAMO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type of device</td>
</tr>
<tr>
<td>1.</td>
<td>Local Module</td>
</tr>
</tbody>
</table>

First results of comparison of 10 MHz signal from H-Maser located at AOS Borowiec and Optical Clock located at KL FAMO Torun is shown on Fig. 2. Green points – Allan

![Fig. 2. First results of comparison of 10 MHz signal from H-Maser located at AOS Borowiec and Optical Clock located at KL FAMO Torun.](image-url)
deviation of comparisons of two optical clocks at KL FAMO. Red points – Allan deviation of measurements of optical comb connected at KL FAMO to 10 MHz from Borowiec, and one of the strontium clocks.

![Fig. 3. Stability of the feed-back coupling of the link AOS-KL FAMO in time domain](image)

The clocks at the KL FAMO use the 10 MHz reference frequency only. Therefore to estimate the quality of the link between the two laboratories, the 1 pps pulse sent to Torun and reflected to Borowiec is measured at AOS against the signal transmitted from AOS (overall distance of ~660 km). Precision of the single measurement is ~9 ps, this corresponds to the single shot precision of used WAT T4100U counter. For sampling times of 1000 s it goes to ~500 femtoseconds. Fig.3 shows stability of the feed-back coupling of the link AOS - KL FAMO in time domain.

### III. NEW LOCAL REPOSITORY

A second huge investment was a new local repository located in Poznan. The local repository is equipped with several specialized devices which are responsible for provision of time and frequency signals to the dissemination network when failure of connection to the reference time laboratories occurs. The local repository contains: atomic clock – model Vremya-CH VCH-1008, multichannel counter, FemtoStepper, and Time Transfer System TTS-4. It will be also equipped with a multiplexer, which is under construction by project partners PSNC. Currently, at the repository the calibration and testing processes are performed.

In Fig. 4, the comparisons of H-maser placed at PSNC Poznan local repository and UTC(AOS) are presented. The results were obtained using time transfer systems TTS-4, working simultaneously at the repository and AOS laboratory (For the computations of the results, Precise Point Positioning Method for GPS+GLONASS phase measurements was used). The distance between the two laboratories is in the range of 25 km. Short term precision of the measurements equals 10 – 11 ps for 30s and 60s averaging intervals.

![Fig. 4. Comparison of UTC(AOS) to H-maser in local repository in Poznan, GPS+GLONASS phase measurements](image)

Very high accuracy of the GNSS time transfer between AOS and PSNC, allowed for the preliminary estimation of the Vremya-CH VCH-1008 stability in reference to UTC (AOS). As a tool, the Modified Allan Deviation was used. Passive H-maser shows excellent short- and mid-term stability. For averaging time of 300 s it is in the range of a few parts of 10^-14. For one day, stability goes down to a few parts of 10^-15. (Fig. 5). The frequency drift of the VCH-1008 is the level of a few parts of 10^-16, but it is still under estimation. Better analysis of the behaviour of VCH-1008 clock will be possible when more precise optical link between AOS and PSNC will become operational (May 2015).

### IV. T&F DISSEMINATION LINE AOS – GUM

The optical 420 km link between Central Office of Measures (GUM) in Warsaw and Astrogeodynamical Observatory (AOS) in Borowiec near Poznan is fully operational and in permanent use for more than 3 years. It
allows for a continuous real-time comparisons of the UTC (PL) and UTC(AOS). This link is the base for OPTIME time and frequency dissemination network.

Fig. 6. The stability of comparisons UTC(AOS) and UTC(PL) over optical fiber

Fig. 6 presents the stability of comparisons of UTC(AOS) and UTC(PL). The precision of a single measurement (sampling time 5 s) is 30 ps. For averaging periods of 0.5 hour increases to 25 ps [3]. For longer intervals the noise of caesium clock located in GUM – UTC(PL) begins to dominate the comparisons.

REFERENCES


Design of the Optical Fiber Transmission Link in a Femtosecond-precision, Fiber-optic Timing Synchronization System

Huaiying Shang, Shengkang Zhang, Fan Shi, Hongbo Wang, Haifeng Wang, Hang Yi, Feng Nian, Keming Feng
Beijing Institute of Radio Metrology and Measurement
Science and Technology on Metrology and Calibration Laboratory
Beijing, CHINA
shanghuaiying2010@163.com

Abstract—In this paper, the details of the linear chirped femtosecond laser pulse used for timing synchronization when transferring in single mode fiber(SMF) and dispersion compensation fiber(DCF) are analyzed based on a numerical simulation. The broadening of the 416fs laser pulse used for timing synchronization by 9.7m DCF and then the compression by 33m SMF can be observed in the experiment. Based on the simulation and experiment, a design of the optical fiber transmission link which is made up of SMF and DCF in the femtosecond-precision fiber-optic timing synchronization system is introduced.

Keywords—Femtosecond laser; Timing synchronization; Dispersion compensation; Chirp

I. INTRODUCTION

Due to the limited precision of the existing satellite-based time synchronization techniques such as GPS common-view or Two-Way Satellite Time and Frequency Transfer[1], and the emergence of optical clock, frequency and time transfer via optical fiber which has advantages of high precision and anti-jamming becomes the trend of high precision frequency and time synchronization[2]. Frequency or time transfer via optical fiber has so far been investigated over regional distances with the use of three different methods for transmission: transmission of an optical carrier wave of a stable continuous wave laser[3], microwave signal transfer by means of an amplitude-modulated laser[4], or transmission of femtosecond laser pulses[5]. There is time information in the signal of the later two methods.

At present the timing synchronization systems based on optical fiber take the optical carrier wave amplitude-modulated by time code for transmission[6]. In this kind of system, active compensation loop is used for eliminating the time jitter introduced by the optical fiber link, which is measured by comparing the reference source and the transmitted signal in electrical field detected by a photoelectric detector from the optical signal. With the development of ultrashort-pulse mode-locked laser, time synchronization technology based on femtosecond laser pulses is becoming more and more valued. We can measure the time jitter between the transmitted pulses and the initial pulses either in electrical field or optical field. The first measurement scheme needs a band pass filter to obtain the harmonic of the pulse’s repetition frequency, whose phase jitter represents the time jitter[7]. The second scheme uses a balanced optical cross correlation modular composed of dichroic mirror and frequency doubling crystal for time jitter measurement. In order to obtain higher accuracy and sensitivity, we study on the femtosecond precision time synchronization technology based on femtosecond laser pulses, which measures time jitter for transmitted pulse using balanced optical cross correlation scheme[8].

The width of femtosecond laser pulses transmitted in optical fiber with dispersion will be broadened. The dispersion compensation of optical fiber transmission link in the femtosecond-precision fiber-optic timing synchronization system needs to be investigated. We should make the broadening amount as small as possible to reduce the influence on the balanced cross correlation detection progress of the timing synchronization system. In this paper, the details of the chirped femtosecond laser pulse used for timing synchronization when transferring in single mode fiber(SMF) and dispersion compensation fiber(DCF) are analyzed based on a numerical simulation. The broadening of the 416fs laser pulse used for timing synchronization by 9.7m DCF and then the compression by 33m SMF can be observed in the experiment. Based on the simulation and experiment, a design of the optical fiber transmission link which is made up of SMF and DCF in the femtosecond-precision fiber-optic timing synchronization system is introduced: Firstly, the absolute value of the femtosecond laser pulse’s initial chirp can be confirmed when the spectral width and pulse width of the laser source are known. Secondly, the sign of initial chirp can be confirmed with the broadening amount of the laser pulse transferred in a piece of DCF. At last, the length of DCF for dispersion compensation is calculated by the numerical simulation, after the final width of the pulse, the initial chirp of the laser source and the length of the SMF are figured out.
II. ANALYSIS OF CHIRPED FEMTOSECOND LASER PULSES TRANSMITTED IN FIBER

The optical signal can present many different effects such as dispersion, nonlinear and so on, when transmitted in the optical fiber, especially the femtosecond laser pulse whose width will be obviously broadened. The nonlinear Schrodinger equation which can describe the details of the femtosecond pulses transmitting in fiber is shown as follow:

\[
\frac{\partial A}{\partial z} + \frac{\alpha}{2} A + \frac{\beta_2}{2} \frac{\partial^2 A}{\partial T^2} + \frac{\beta_3}{6} \frac{\partial^3 A}{\partial T^3} = i \gamma |A|^2 A + \frac{i}{\omega_0} \frac{\partial}{\partial T} (|A|^2 - \delta_T A^* \delta_T A) \tag{1}
\]

Where \(A(z,t)\) is the slowly varying complex amplitude of the pulse, \(T=t-\beta_1 z\), \(\beta_1=1/v_g\), \(T\) is time in reference system moving along with pulse in a group speed \(v_g\). \(z\) is the transmitting distance in fiber of the pulse, \(\alpha\) the attenuation coefficient of fiber, \(\beta_2\) and \(\beta_3\) the two order dispersion coefficient and the third order dispersion coefficient. There are nonlinear dependencies on the right of equation.

The femtosecond laser pulse can be represented as linear chirped Gaussian pulse, expressed as:

\[
A(0,T)=\exp[-4\ln 2(1+i\zeta)T^2/2T_0^2] \tag{2}
\]

Where \(T_0\) is the full width at half maximum(FWHM) of the pulse, \(P_0\) the initial peak power, and \(C\) the chirp parameter.

After the analysis above, we can obtain the situation of the transmitted pulse by the numerical simulation established through the split-step Fourier method. Then the influence of various factors such as the initial chirp, dispersion, etc, to the width of transmitted pulse can be researched.

The initial chirp is respectively set to +1 and -1, and initial FWHM is set to 500fs. We choose SMF as the transmission medium in figure 1, DCF in figure 2. Broadening evolution of the chirped Gaussian pulse with transmission distance is shown in figure 1 and figure 2. The two curves in each figure have the same parameters except the sign of the initial chirp.

![Fig.1 Broadening evolution of the chirped Gaussian pulse with transmission distance(SMF)](image)

![Fig.2 Broadening evolution of the chirped Gaussian pulse with transmission distance(DCF)](image)

When the transmission medium is SMF, the pulse is broadened fast and the broadening amount keep raising along with the transmission distance increasing in the situation of \(C=1\). At the same time, we can firstly see an obvious compressing process if \(C=-1\), then the pulse becomes broadened and the broadening amount keep increasing with the same rate to \(C=1\). When the transmission medium is DCF, the appearance is opposed to SMF with the relevant of the sign of \(C\). All of these indicate that the initial chirp of pulse can compensate the dispersion, which is determined by the sign of the initial chirp and the type of fiber.

III. DESIGN OF THE OPTICAL FIBER TRANSMISSION LINK

When time synchronization is needed between the main station and the substation, the length of SMF is fixed in optical fiber network, and we should complete the deployment of femtosecond laser source, balanced optical cross correlation modular, active time delay compensation device and dispersion pre-compensation module in the main station. The femtosecond laser pulses carrying time information is transferred through the dispersion pre-compensation module and SMF in network from the main station to substation, and then transferred back to the main station by the same fiber link.

According to the preceding analysis, the length of DCF in the dispersion pre-compensation module can be calculated as follow: Firstly, the absolute value of the femtosecond laser pulse’s initial chirp can be confirmed when the spectral width and pulse width of the laser source are known. Secondly, the sign of initial chirp can be confirmed with the broadening amount of the laser pulse transferred in a piece of DCF. At last, the length of DCF for dispersion compensation is calculated by the numerical simulation, after the final width of the pulse, the initial chirp of the laser source and the length of the SMF are figured out.

A. Confirm the absolute value of initial chirp
The spectrum of the femtosecond laser source is measured by a spectrometer, whose center wavelength is 1556.1nm and half width in the 1/e power is 5.4nm, as is shown in figure 3. The auto-correlation trace of the initial laser pulse is measured by an autocorrelation instrument, whose FWHM is 416fs, as is shown in figure 4.

The half width in the 1/e power of spectrum $\Delta w$ and the FWHM of the initial laser pulse $T_0$ conform to the following equation:

$$\Delta w = 2(\ln 2)^{1/2}(1 + C^2)^{1/2} / T_0$$

Where $w=2\pi c/\lambda$, $c$ is the light speed, $\lambda$ is wavelength. The absolute value of the femtosecond laser pulse’s initial chirp can be confirmed as 0.32.

**B. Confirm the sign of initial chirp**

The initial chirp is respectively set to +0.32 and -0.32, and initial FWHM is set to 416fs, when choosing DCF as the transmission medium in the numerical simulation. In the transmission distance of 9.7m, if $C=0.32$, the FWHM of transmitted pulse is 5.71ps, at the same time, if $C=-0.32$, the FWHM is 5.98ps.

After the output of the laser is connected to the 9.7m DCF, the auto-correlation trace of the transmitted pulse can be measured by an autocorrelation instrument, as is shown in figure 5. The FWHM is 6.07ps, therefore the sign of initial chirp is minus and the chirp is -0.32.

**C. Calculate the length of DCF**

The initial chirp of the femtosecond laser pulses is -0.32, and the FWHM is 416fs. Assume the length of the SMF in the fiber network to be 33m, let the output of the femtosecond laser source transfers through a piece of DCF for dispersion pre-compensation, and then a 33m long SMF in the numerical simulation. Broadening evolution of the transmitted pulse with the length of DCF is shown in figure 6.

We can conclude that take the length of DCF about 9m can satisfy the dispersion compensation requirement of the femtosecond precision timing synchronization System.

In order to checkout our conclusion in experiment, the fiber link is consist of a 9.7m DCF and a 33m SMF. The auto-correlation trace of the transmitted pulse measured by an autocorrelation instrument is shown in figure 7, from which we can see the width of the femtosecond pulses is effectively controlled.

In addition, a faraday rotation mirror is needed in the fiber link at substation to decrease the influence of the polarization mode dispersion, when the length of the fiber link is long enough.
The femtosecond precision time synchronization system studied in this paper, which measures time jitter for transmitted pulse using balanced optical cross correlation scheme, has advantages of higher accuracy and sensitivity. We analyze the details of the chirped femtosecond laser pulse when transferring in single mode fiber(SMF) and dispersion compensation fiber(DCF) by means of a numerical simulation. Whether the initial chirp of pulse can compensate the dispersion or not is determined by the sign of the initial chirp and the type of fiber. On the basis of theoretical analysis and experiment, a design of the time synchronization system’s fiber link is proposed. The length of DCF for dispersion compensation is calculated by the numerical simulation, after the final width of the pulse, the initial chirp of the laser source and the length of the SMF are figured out.

ACKNOWLEDGMENT

This research is supported by the Science and Technology on Metrology and Calibration Laboratory in China. The authors would like to thank all members of the laboratory.

REFERENCES

The method of determination of GEO satellite precise clock bias during maneuvering

Mei-fang Wu
National Time Service Center of Chinese Academy of Sciences
Xi’an, China
Key Laboratory of Precision Navigation and Timing Technology, NTSC, CAS
Xi’an, China
University of Chinese Academy of Sciences
Beijing, China
wumeifang@ntsc.ac.cn

Pei Wei
National Time Service Center of Chinese Academy of Sciences
Xi’an, China
Key Laboratory of Precision Navigation and Timing Technology, NTSC, CAS
Xi’an, China

Xu-hai Yang
National Time Service Center of Chinese Academy of Sciences
Xi’an, China
Key Laboratory of Precision Navigation and Timing Technology, NTSC, CAS
Xi’an, China

Shou-gang Zhang
National Time Service Center of Chinese Academy of Sciences
Xi’an, China

Abstract—This paper aims to research and determine GEO satellite clock bias during maneuvering. By analyzing of GEO satellite clock bias data, quadratic polynomials, cubic spline and Lagrange are chose as interpolation methods. The result of the test in this paper shows that in most cases, cubic spline interpolation is the best one of the three interpolation methods. And the accuracy of cubic spline interpolation is at the level of 0.38ns~0.38ns which can meet the actual demand; besides the stability of cubic spline interpolation is obviously better than that of quadratic polynomials and Lagrange interpolations.

Keywords—GEO satellite; Maneuver; Satellite clock bias determination

I. INTRODUCTION

GNSS provides global continuous service, is a superb tool for scientific applications. The most important applications of GNSS include navigation, surveying and mapping, land resources monitoring, meteorology, seismology, security assistance and etc. [1]. So precision is not the only parameter to consider, availability and continuity are also worth taking on. For the China’s own BeiDou navigation system (hereinafter referred to as BDS), “Public service performance specification of BeiDou satellite navigation system (version 1)” explicitly stipulate relevant indicators such as availability and continuity in addition to the precision.

BDS of China includes GEO satellite as one kind of its constellation. But during the lifetime of GEO satellite, it is necessary to orbital maneuver within some time intervals to avoid radio frequency interference. The orbital maneuvers bring artificial and rough thrust, which result in dynamic modeling errors, and then affect precise orbit determination and satellite clock bias determination. So the GEO satellite precise clock product during maneuvering is not included in precise clock product at present, which influences availability and continuity of BDS and brings trouble in researching with GEO satellite clock product [2]. Therefore, it is important to research and determine the GEO satellite precise clock bias during maneuvering.

In essence, determination of GEO satellite precise clock bias during maneuvering is a problem about interpolation, which use GEO satellite clock bias data before and after satellites maneuver to determine GEO satellite precise clock bias during maneuvering by interpolation.

II. ANALYSIS OF INTERPOLATION METHODS

There are many mathematical methods of interpolation, such as quadratic polynomials, cubic spline and high-order polynomial (such as Lagrange, Chebyshev) etc. [2, 3, 4, 5]. In theory, all these methods can be used for the determination of satellite precise clock bias.

Satellite clock bias which can be described as quadratic polynomials, include clock offset, frequency deviation, frequency drift and random error of the clock. Assuming that quadratic polynomial coefficient is \(a_0, a_1, a_2\), we use Least Squares to calculate model parameters. As shown in “(1)”, satellite clock bias is defined as the difference between satellite clock time \(T\) and the time of navigation system \(t\). In “(1)”, \(t_0\) is the reference epoch, \(a_0\) is clock offset, \(a_1\) is clock speed and \(a_2\) is variable rate of clock speed. Therefore, quadratic polynomial can be used to interpolate satellite clock bias.

\[
T - t = a_0 + a_1(t - t_0) + a_2(t - t_0)^2 \quad (1)
\]

Meanwhile, the satellite clock error has irregular jump degeneration, but in the short term, the trend of clock error can be considered as approximate smooth, therefore, Lagrange polynomial interpolation model can also be used to interpolate satellite clock bias. In addition to clock offset, frequency deviation and frequency drift, satellite bias also includes the random error of the clock. Clock offset, frequency deviation and frequency drift can be calculate easily, while the random error of the clock is too complicated to calculate. So due to the random error of the clock, the interpolation precision of satellite clock bias by using high-order Lagrange polynomial is unsatisfactory. The relevant domestic scholars have proved the
above conclusion and point out that the best interpolation result (the standard deviation and the maximum error is minimum) is obtained when the order number of Lagrange polynomial is 2. And the precision of interpolation is more and lower when the order of Lagrange polynomial is more than 10[4]. That is, the higher the order numbers of Lagrange polynomial, the lower the interpolation precision of satellite clock bias. So we suggest that the order number of Lagrange polynomial should be no more than 10 in order to obtain a satisfied result when Lagrange polynomial is used to interpolate satellite clock bias.

From a mathematical point of view, quadratic polynomial has poor smoothness although easy to implement; high-order polynomial has the disadvantage of oscillation and Runge phenomenon; compared with them, cubic spline which is easy to implement is a section three function and every function node has a continuous two derivative, so it is widely used in interpolating satellite clock bias[3].

III. ANALYSIS OF EXAPLES

A. Define Test Plan

Firstly, we should construct test data during satellite maneuvering; then interpolate the test data to determine the satellite clock bias during maneuvering by interpolation models such as quadratic polynomials, cubic spline and Lagrange polynomial; finally we compare the satellite clock bias during maneuvering with IGS precise satellite clock bias and conclude.

B. Construct Test Data

When GEO satellite maneuvering, there are 8 hours of unavailable data in broadcast ephemeris and no precise satellite clock bias and orbit data in IGS. So we must use precise satellite clock bias data when GEO satellite dose not maneuver to construct GEO satellite data during maneuvering as test data.

The precise satellite clock bias data we used is supplied by GEO satellite C03 of BDS on 2014.7.13 when there is no satellite maneuvering. Fig.1 shows that the sampling interval of the data we used is 5minutes, and the amount of epoch is 288.

Because the time satellite maneuvering is no longer than 2 hours, so we construct test data by deleting 2hours of precise satellite clock bias data we used. Fig. 2 shows the test data which we construct has two parts. Part 1 is given data including epoch 1-132 and 155-288, the amount of epoch is 264. Part 2 is unknown data including epoch 133-156, the amount of epoch is 24.

C. The Result of Test

Because the precision of IGS precise GEO satellite clock bias is better than 1ns, it can be considered as true value to test the precision of the three models. We compare IGS precise satellite clock bias with satellite clock bias during maneuvering which we calculate then get the residuals and determine which model is the best one.

Because of its easy model and compact structure, Lagrange polynomial is a classical interpolation method [5]. Fig.3 shows the residuals determined by 2-order Lagrange polynomial. Low-order polynomial does not have obviously oscillation and Runge phenomenon, and the residuals are less than 0.2ns.

Fig. 3. The residuals of C03 determined by 2-order Lagrange polynomial

Fig. 4. The residuals of C03 determined by cubic spline polynomial

Fig. 5. The residuals of C03 determined by quadratic polynomial

Fig. 4 shows the residuals determined by cubic spline polynomial, and the residuals are less than 0.15ns. Quadratic polynomial has poor smoothness although easy to implement. Fig.5 shows the residuals determined by quadratic polynomial, and the residuals are less than 0.45ns.
<table>
<thead>
<tr>
<th>Satellite NO.</th>
<th>Models</th>
<th>Max</th>
<th>Mean</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>C01</td>
<td>Quadratic</td>
<td>0.4741</td>
<td>0.2131</td>
<td>0.4391</td>
</tr>
<tr>
<td></td>
<td>Cubic spline</td>
<td>0.5507</td>
<td>0.2613</td>
<td>0.3120</td>
</tr>
<tr>
<td></td>
<td>Lagrange</td>
<td>0.6743</td>
<td>0.3354</td>
<td>0.3889</td>
</tr>
<tr>
<td>C02</td>
<td>Quadratic</td>
<td>1.4215</td>
<td>0.7887</td>
<td>0.9123</td>
</tr>
<tr>
<td></td>
<td>Cubic spline</td>
<td>0.2020</td>
<td>0.1072</td>
<td>0.1235</td>
</tr>
<tr>
<td></td>
<td>Lagrange</td>
<td>0.4243</td>
<td>0.1486</td>
<td>0.1903</td>
</tr>
<tr>
<td>C03</td>
<td>Quadratic</td>
<td>0.4309</td>
<td>0.3438</td>
<td>0.3517</td>
</tr>
<tr>
<td></td>
<td>Cubic spline</td>
<td>0.1443</td>
<td>0.0721</td>
<td>0.0834</td>
</tr>
<tr>
<td></td>
<td>Lagrange</td>
<td>0.1823</td>
<td>0.0643</td>
<td>0.0735</td>
</tr>
<tr>
<td>C04</td>
<td>Quadratic</td>
<td>1.3796</td>
<td>1.0695</td>
<td>1.0940</td>
</tr>
<tr>
<td></td>
<td>Cubic spline</td>
<td>0.4524</td>
<td>0.2014</td>
<td>0.2335</td>
</tr>
<tr>
<td></td>
<td>Lagrange</td>
<td>0.7332</td>
<td>0.2661</td>
<td>0.3241</td>
</tr>
<tr>
<td>C05</td>
<td>Quadratic</td>
<td>0.9450</td>
<td>0.5265</td>
<td>0.6025</td>
</tr>
<tr>
<td></td>
<td>Cubic spline</td>
<td>0.3261</td>
<td>0.1554</td>
<td>0.1827</td>
</tr>
<tr>
<td></td>
<td>Lagrange</td>
<td>0.4202</td>
<td>0.1575</td>
<td>0.1912</td>
</tr>
</tbody>
</table>

The data of GEO satellite C01-C05 of BDS on 2014.7.13 is interpolated by three methods, and then the residuals are shown in Table 1. The result shows as follows.

1) The results of 5 GEO satellites show that the precision of quadratic polynomial is the worst, and the max residual is more than 1ns, so quadratic polynomial cannot be used to determine the GEO satellite clock bias during maneuvering.

2) The results of 5 GEO satellites show that the precisions of cubic spline and low-order Lagrange polynomial are at the level of sub-nanosecond; both of them can meet the practical needs.

3) The results of 5 GEO satellites show that the difference of the precisions of cubic spline and low-order Lagrange polynomial is tiny, which is almost at the level of 0.01ns.

4) The results of 4 GEO satellites including C01, C02, C04 and C05 show that the precisions of cubic spline is better than that of low-order Lagrange polynomial.

5) The results of 5 GEO satellites show that the stability of cubic spline polynomial is the best, and the max residual of cubic spline is less than that of the other two methods.

6) Above all, cubic spline and low-order Lagrange polynomial can be used to determine the GEO satellite clock bias during maneuvering, but cubic spline is better than low-order Lagrange polynomial.

IV. CONCLUSION

All told, in most cases, cubic spline polynomial is the best one of the three methods to determine the GEO satellite clock bias during maneuvering. In iGMAS (international GNSS Monitoring and Assessment Service), the precision of GEO satellite clock bias is at the level of 0.5ns, and the precision of cubic spline polynomial is at the level of 0.01ns-0.38ns, so cubic spline polynomial can meet the practical needs. And the stability of cubic spline polynomials is much better than that of quadratic polynomial and low-order Lagrange polynomial.

Also cubic spline polynomial which is easy to implement is a section three function and every function node has a continuous two derivative, so the result of it is continuous and smooth. Above all, cubic spline polynomial is the ideal method to determine the GEO satellite clock bias during maneuvering, which can improve the availability and continuity of BDS and make sense to the research with GEO satellite clock product.

ACKNOWLEDGMENT

The authors thank Professor Xu-hai Yang and Professor Xiao-hui Li for their persevering and meticulous guidance.

REFERENCES

Comparison of different carrier-envelope frequency stabilization methods for a high performance DPSSL frequency comb

Stefan Kundermann and Steve Lecomte
Centre Suisse d’Electronique et de Microtechnique (CSEM)
Jaquet-Droz 1, 2000 Neuchâtel, Switzerland
Stefan.kundermann@csem.ch

Abstract—Carrier envelope frequency ($f_{CEO}$) stabilization of frequency combs is traditionally achieved via power modulation of the pump of the comb oscillator. A further possibility is to shift the laser $f_{CEO}$ using an external acousto-optic frequency shifter (AOFS). In this case the optical frequency comb spectrum is shifted exactly by the RF modulation frequency of the AOFS. In this work different stabilization schemes in self-referenced frequency comb system architecture are compared. It is shown that AOFS frequency shifting represents a high performance alternative to the standard feedback control via pump power modulation.

Keywords—DPSSL; frequency comb

I. INTRODUCTION

Frequency combs have revolutionized frequency [1][2] and distance measurement metrology [3] since they allow phase coherent connection of optical and electronic frequencies [1]. They are a key component of optical atomic clocks [4] and can be employed for high precision spectroscopy [5]. Furthermore they can be used for generation of ultra-pure microwave signals with unprecedented low phase noise [6]. Drivers for the development of the frequency comb are of course the improvement of reliability and robustness to address a broader spectrum of technological applications as for example use for space applications [7]. On the other hand improvement of noise properties of the frequency comb, higher repetition rate [8], higher output power [9] or other laser stabilization methods using acousto-optic frequency shifters [10][11], SESAM modulation [12] or zero difference frequency generation [13] are subject to intense research and development activities. Lasers have shown cutting edge noise properties when the more traditional stabilization of $f_{CEO}$ via modulation of the pump laser power is used [14]. Potentially $f_{CEO}$ stabilization acting on the laser beam using an external acousto-optic frequency shifter (AOFS) is a promising solution to come to lower noise properties thanks to the AOFS's high modulation bandwidth[10][11]. In this paper high-performance $f_{CEO}$ stabilization of a diode-pumped solid-state laser (DPSSL) using the traditional pump laser power modulation and the external frequency shifting using an AOFS are presented and compared.

II. SCHEMATIC AND SETUP

A principle schematic representing the different stabilization schemes for $f_{CEO}$ is represented in Fig. 1. 190 fs transform limited pulses from a 100 MHz DPSSL laser oscillator emitting at 1560 nm are spectrally broadened in a highly nonlinear fiber (HNLF) to a bit more than one octave. The super continuum passes an AOFS and a part of it is deflected and frequency shifted by the AOFS driving frequency. F-2f interferometers at both outputs detect $f_{CEO}$ of the deflected and non-deflected beams. In Fig. 2 the optical setup downstream the AOFS is shown. The non-deflected beam f-2f interferometer (ND f-2f) is implemented by a standard Michelson f-2f interferometer architecture (red shaded part in Fig. 2). M5, which is a dichroic mirror, only reflects light around 1 μm carrying the $f_{CEO}$ beat note signal. All other parts of the supercontinuum are transmitted and thus rejected from detection in photodiode PD1. The deflected beam f-2f interferometer is more complex (see green shaded part in Fig. 2): low optical frequency red part (at 2 μm) and high optical frequency blue part (at 1 μm) of the spectrum are deflected at different angles. The red part at the higher deflection angle is reflected by a dichroic beam splitter BS2 and reflected collinearly by a mirror (M6) which is slightly misaligned down in direction of the optical table. This misalignment is introduced in order to split away the output of the Michelson interferometer with mirror M6.

![Fig. 1: Schematic of stabilization schemes.](image-url)
The blue part deflected at a smaller angle passes the dichroic beam splitter and is then reflected by two mirrors in such a way that it is collinearly superposed to the red part of the super continuum at the output of the interferometer. The output beam of the interferometer is focused on a non-linear crystal acting as second harmonic generator (SHG). The blue part and the frequency doubled red part are then focused on photo detector PD2, which detects the f CEO beat note of the diffracted beam. In this interferometer supercontinuum parts not carrying the fCEO beat note are rejected by the fact that the different wavelengths of the supercontinuum are angularly separated and only the useful parts pass the interferometer. The interferometers are used for the three different stabilization methods discussed later on.

III. REPETITION RATE AND FCEO STABILIZATION

A. Repetition rate lock

The electronic setup for the repetition rate lock is shown in Fig. 3. The repetition rate harmonic at 400MHz is detected by a fast photo diode. The signal is band pass filtered to isolate the 400 MHz harmonic only. It is frequency divided by a factor of 33. Signal phase is compared to the phase of a reference synthesizer and an error signal proportional to the relative phase is generated. The error signal is loop filtered, amplified and fed to the piezo transducer controlling laser cavity length.

B. fCEO lock a: Feedback on pump laser current

Regarding the overall configuration, the optical setup with the AOFS switched off and the ND f-2f has been used (Fig. 1 and Fig. 2). The electronic setup for the fCEO lock a is shown in Fig. 5. The fCEO harmonic peak at 390 MHz is detected and then band pass filtered, and frequency divided as in case for the repetition rate lock. The error signal representing the relative phase difference between reference synthesizer and fCEO is loop filtered and used for modulation of the pump laser diode current.

![Fig. 2: Optical setup AOFS with f-2f interferometers.](image)

![Fig. 3: Electronic setup for repetition rate lock.](image)

![Fig. 4: Phase noise and Allan Deviation of repetition rate lock.](image)

![Fig. 5: Electronic setup for fCEO lock a.](image)
Phase noise and Allan Deviation of \( f_{CEO} \) lock a are shown in Fig. 6. Phase noise diminishes to -82 dBc/Hz at an offset frequency of 0.5 Hz. A plateau at around -56 dBc/Hz follows between 40 Hz and 25 kHz. The peak at 25 kHz corresponds to the servo bump of the PLL. Integrated phase noise also measured with a noise equivalent bandwidth of 500 Hz amounts to 4.0 \( \times 10^{-11} \) at 1 s. Allan deviation has a \( \tau^{-1} \) slope until 200 s followed by a small shoulder resulting in a value of 6.0 \( \times 10^{-14} \) at 1000 s. In [12] \( 10^{-8} \) at 1 s was measured but noise equivalent bandwidth was not reported.

Fig. 7 shows the laser relative intensity noise (RIN) with and without \( f_{CEO} \) stabilization. The RIN of the free running laser exhibits technical noise to about 50 Hz followed by a plateau at -132 dBc/Hz up to a frequency of about 4 kHz. RIN then cuts off corresponding to the laser gain medium upper state lifetime [16]. The following plateau at -155 dBc/Hz corresponds to the shot noise limit of the measurement. When \( f_{CEO} \) is stabilized RIN diminishes by up to 20 dB within the lock bandwidth of 25 kHz. As already observed in [14] the \( f_{CEO} \) lock here efficiently also stabilizes laser power.

The feed forward \( f_{CEO} \) locking scheme b uses \( f_{CEO} \) detected from the ND f-2f for stabilization and the D f-2f for characterization of the lock phase noise as \( f_{CEO} \) for the diffracted beam is stabilized (Fig. 1). \( f_{CEO} \) is filtered by a band pass filter and mixed with the signal of a reference synthesizer at frequency \( f_S \) in such a way that \( f_{CEO} \) with its frequency fluctuations is shifted to the AOFS modulation frequency range. After this mixer the peak in the AOFS modulation frequency range is filtered by a band pass. The signal is amplified in order to drive the AOFS. The shift frequency \( f_S \) is directly present as \( f_{CEO} \) peak in the output signal of the D f-2f interferometer and could in principle be used for phase comparison to reference synthesizer 1 for the phase noise measurement. In order to reduce noise of the RF signal and phase detection, a different \( f_{CEO} \) peak is down shifted to \( f_S \) using a second reference synthesizer and a mixer was used. Both synthesizers were referenced to the same 10 MHz signal provided by an H-Maser.

Phase noise and Allan Deviation plots are shown in Fig. 9. When the AOM only is used to lock the CEO frequency, phase noise is dominated by technical noise up to a frequency of 10 Hz. Then it follows a plateau at -60 dBc/Hz, which cuts off at about 4 kHz. The shape of the plateau is similar to the laser RIN (see Fig. 7, red curve). A second feed forward lock has been performed using feedback control on the pump laser current at the same time. In this case, the phase noise is 10-15 dB lower in the range between 100 Hz and 5 kHz.

Astonishingly here the servo bump of the laser current feedback lock is also visible in the feed forward lock phase noise and its shape is rather similar to the shape of the RIN curve for the current feedback lock represented in Fig. 7, blue curve. The integrated phase noise of the two locks amounts to 160 mrad even if shape of phase noise curves is quite different for offset frequencies below 80 kHz. Allan deviation of the two feed forward locks are quite similar and frequency stability at 1 s is about 2 \( \times 10^{-10} \). \( \tau^{-1} \) behavior is observed from about 0.4 to 10 s indicating some drift with characteristic time scale of 10 s.

Fig. 6: Phase noise and Allan Deviation for \( f_{CEO} \) lock a.

Fig. 7: Laser RIN (free running and for \( f_{CEO} \) lock a).

Fig. 8: Electronic setup for \( f_{CEO} \) lock b.

C. \( f_{CEO} \) lock b: Feed forward on AOFS frequency

The feed forward \( f_{CEO} \) locking scheme b uses \( f_{CEO} \) detected from the ND f-2f for stabilization and the D f-2f for characterization of the lock phase noise as \( f_{CEO} \) for the diffracted beam is stabilized (Fig. 1). \( f_{CEO} \) is filtered by a band pass filter and mixed with the signal of a reference synthesizer at frequency \( f_S \) in such a way that \( f_{CEO} \) with its frequency fluctuations is shifted to the AOFS modulation frequency range. After this mixer the peak in the AOFS modulation frequency range is filtered by a band pass. The signal is amplified in order to drive the AOFS. The shift frequency \( f_S \) is directly present as \( f_{CEO} \) peak in the output signal of the D f-2f interferometer and could in principle be used for phase comparison to reference synthesizer 1 for the phase noise measurement. In order to reduce noise of the RF signal and phase detection, a different \( f_{CEO} \) peak is down shifted to \( f_S \) using a second reference synthesizer and a mixer was used. Both synthesizers were referenced to the same 10 MHz signal provided by an H-Maser.
A possibility for the transfer of the RIN to the phase noise is given by the fact that inside the AOFS there is considerable phase lag as a function of the AOFS modulation frequency. The reason for this shift is the distance between the piezo actuator and the optical beam. The relative phase between the piezo actuator and the optical beam. The different AOFS modulation frequencies with different phase lag between the piezo actuator and the optical beam. The relative phase difference imprinted onto the stabilized f CEO as a function of AOFS modulation frequency can be calculated using the formula shown in the center of Fig. 10, where D is the distance between piezo actuator and optical beam, c_{AC} the phase velocity of the acoustic wave in the crystal and f_1 and f_2 two different AOFS modulation frequencies. For example when the AOFS modulation frequency is changed (by varying the pump laser current and thus changing free running f_{CEO}) by 200 kHz and D=6mm, a phase change of 3 rad can be expected. The right panel indicates a measurement of the relative phase between the reference synthesizer with f_S and stabilized f_{CEO} of the diffracted beam. Indeed when the free running f_{CEO} is changed by 200 kHz relative phase changes by approximately 3 rad. Furthermore in the graph depicting relative phase a phase drift can be observed, which is due to the thermal drifts of the ND f-2f and the D f-2f interferometers.

The third f_{CEO} lock discussed here is the feedback lock of f_{CEO} via modulation of the AOFS frequency. Here the D f-2f interferometer is used as detection element for the PLL. f_{CEO} is band pass filtered, divided and phase compared to a reference synthesizer generating an error signal. The error signal is loop filtered and added to a DC voltage. The sum voltage drives a VCO generating the driving frequency of the AOFS. The output voltage of the summing amplifier is adjusted in such a way that if the correction signal coming from the loop filter is zero, the AOFS modulation frequency is 79 MHz. The third f_{CEO} lock discussed here is the feedback lock of f_{CEO} via modulation of the AOFS frequency. Here the D f-2f interferometer is used as detection element for the PLL. f_{CEO} is band pass filtered, divided and phase compared to a reference synthesizer generating an error signal. The error signal is loop filtered and added to a DC voltage. The sum voltage drives a VCO generating the driving frequency of the AOFS. The output voltage of the summing amplifier is adjusted in such a way that if the correction signal coming from the loop filter is zero, the AOFS modulation frequency is 79 MHz.

The phase drifts exhibit a servo bump at about 100 kHz demonstrating the high lock bandwidth when using the AOFS. Allan Deviation follows a 3.0·10^{-12} s^{-1} from 0.001 to 1000 s for the AOFS lock only (see Fig. 12, blue curve) and 4.5·10^{-13} s^{-1} from 0.001 to 20 s for the AOFS lock with current lock (red curve).
IV. CONCLUSIONS AND OUTLOOK

Performance of the three $f_{CEO}$ locks is compared in the following table:

<table>
<thead>
<tr>
<th>Property</th>
<th>$f_{CEO}$ lock a: Current feedback stabilization $f_{CEO}$</th>
<th>$f_{CEO}$ lock b: AOFs feed forward $f_{CEO}$ stabilization</th>
<th>$f_{CEO}$ lock c: AOFs feedback $f_{CEO}$ stabilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPN @390 MHz</td>
<td>384 mrad</td>
<td>160 mrad AOFs lock only</td>
<td>118 mrad AOFs lock only</td>
</tr>
<tr>
<td>ADEV at 390 MHz, @1s</td>
<td>$4.0 \times 10^{-11}$</td>
<td>$1.1 \times 10^{-10}$ AOFs lock only</td>
<td>$3.0 \times 10^{-11}$ AOFs lock only</td>
</tr>
<tr>
<td>Long term locking stability</td>
<td>Pump laser current for $f_{CEO}$ has a large actuation range</td>
<td>$f_{CEO}$ may drift out of AOFs frequency range resulting in unlocking of laser (in case of AOFs lock only)</td>
<td>Complex f-2f setup in diffracted beam of AOFs; could be made less complex with prism compensation and chirp pre compensation using chirped mirrors [10]</td>
</tr>
<tr>
<td>Complexity of optical setup</td>
<td>One simple potentially collinear f-2f interferometer</td>
<td>Colinear f-2f interferometer. For lock diagnostics complex second f-2f interferometer needed lock on diffracted beam.</td>
<td>Complex f-2f setup in diffracted beam of AOFs; could be made less complex with prism compensation and chirp pre compensation using chirped mirrors [10].</td>
</tr>
</tbody>
</table>

As can be noticed from the table, AOFs locks have better noise properties. This is specially due to the fact that AOFs modulation bandwidth is much higher than current modulation bandwidth, which is limited by the upper laser state lifetime of the DPSSL gain medium. However it needs to be pointed out that with in depth study even lower phase noise with standard pump power modulation feedback lock can be achieved [14]. The feed forward lock (b) using the AOFs suffers from phase lag in AOFs and is only attractive when using feedback lock via pump power in parallel. AOFs feedback lock exhibits best lock performance and does not suffer from phase lag. It requires however complex optical setup with an advanced Michelson interferometer which is demanding to align. Furthermore it implies that the useful optical beam at the output of the D f-2f interferometer for stabilization is limiting the laser power available for the user of the stabilized laser system. RIN is not improved when lock is performed with the AOFS only. This is evident, since the AOFS in the present configuration does not directly act on the laser power but only on the $f_{CEO}$. A laser power stabilization using the AOFS might be implemented as well by modulation of the AOFS RF signal power. It can be concluded, that the traditional $f_{CEO}$ lock via pump power modulation still presents a valid and powerful approach. In case an AOFs is used on top phase noise may be improved as the AOFS acts here as a second cleaning step for $f_{CEO}$. Here lock b is preferable as it only requires one collinear f-2f interferometer for the lock, which makes the system simpler, more robust and less sensitive to drift.

ACKNOWLEDGMENT

The authors acknowledge funding by the Swiss National Science Foundation (SNF) and the Canton of Neuchâtel.

REFERENCES

Development of an erbium-fiber-laser-based optical frequency comb at NTSC

Zhang Yanyan1, Yan Lulu1, Fan Songtao1,2, Zhang Long1, Zhao Wenyu1,2, Guo Wenge1, Zhang Shougang1 and Jiang Haifeng1*
1 Key Lab. of Time and Freq. Primary Standards, National Time Service Center, Xi’an China
2 Graduate University of Chinese Academy of Sciences Beijing, China
3 School of Science, Xi’an Shiyou University, Xi’an, 710065, China
Email: zhangyanyan@ntsc.ac.cn; * haifeng.jiang@ntsc.ac.cn

Abstract—We report the research progress of an erbium-fiber-based optical frequency comb with repetition rate 232 MHz. Its repetition rate is stabilized to a continuous wave laser via an intra-cavity electro-optic modulator and a piezo-transducer, yielding an in-loop frequency instability about 2.1×10^-16 @ 1s. The carrier envelope offset (CEO) frequency with a signal-to-noise ratio of 45 dB for 300 kHz resolution spectrum is detected by using a common path f-f interferometer. CEO frequency is locked to a RF reference frequency by controlling the pump current. The frequency instability induced by in-loop CEO frequency is about 2.9×10^-16@ 1s. The frequency count in use is a Π-type counter from K&K.

Keywords—Optical frequency comb, Fiber laser, Frequency stabilization, Frequency instability

I. INTRODUCTION

Many basic and applied research fields, including attosecond science[1], spectroscopy[2], astronomical spectrograph[3], low-noise microwave generation[4], distance measurement[5], time & frequency transfer[6], optical clock development and application[7] etc., benefit from the optical frequency combs. Among kinds of optical frequency combs[8-12], Er: fiber frequency combs are of great interest, because of their compactness, reliability, low power consumption and directly covering telecommunication wavelength. Now, the performance of the Er: fiber-based frequency combs is as good as that of the best Ti: sapphire-based combs[13].

In this paper, we present the research progress of a home-made erbium-fiber-laser-based optical frequency comb at NTSC. First, We introduce our home-made Er: fiber femtosecond (fs) laser. Then, we give the results of the phase locking of both carrier-envelope-offset frequency (f_{CEO}) and repetition rate (f_{r}) of the fs laser. Finally, we show the in-loop frequency instability of f_{r} and f_{CEO}.

II. EXPERIMENTAL SETUP AND MEASUREMENT RESULTS

The schematic setup of the frequency-stabilized optical comb is shown in Fig.1. The source is a home-made Er: fiber ring laser employing nonlinear polarization rotation mechanism for mode locking. The laser cavity is composed by 43.5 cm of 110 dB/m Erbium-doped gain fiber, 34.5 cm SMF-28 fiber and free space part including two collimators, four wave plates, a polarization beam splitter and an EOM. The net dispersion of the oscillator cavity is about -2000 fs² at 1550 nm. A 2-cm PZT mounted on a stage allows the repetition rate to be adjusted more than 3 kHz. The EOM is an 8-mm long lithium niobate crystal with a 10 MHz response bandwidth.

![Fig 1. Schematic diagram of the Er:fiber-laser-based frequency comb.](image)

Fig 1. Schematic diagram of the Er:fiber-laser-based frequency comb. The thick solid lines represent free-space paths, the thin solid lines represent fiber paths and the dotted lines represent electrical paths.

**PLL**

This work was supported by National Natural Science Foundation of China (Grant No. 91336101 and 61127901) and West Light Foundation of The Chinese Academy of Sciences (2013ZD02)
the laser is approximately 120 nm centered at ~1570 nm. Fig. 2 (b) shows the RF spectrum of the laser’s power, which is detected by using a 2 GHz bandwidth InGaAs photodetector (EOT-3000A). We can clearly see that the laser’s repletion rate is about 232 MHz.

The output of the ring laser is distributed into three parts: Part 1 is the comb’s output for utilization; Part 2 is to produce optical beatnote between the optical comb and a continuous wave (CW) fiber laser (NKT koheras) at 1555.5 nm; Part 3 is to generate $f_{ceo}$ signal.

We combine both CW laser and comb’s signals into a 2 GHz bandwidth InGaAs PD in free space. Plenty of optical beatnotes are produced at PD output. We pick up one at 274.4 MHz and 80-times frequency-divide it to be a 3.43 MHz signal. After that, we lock phase of the 3.43 MHz signal to a RF reference frequency by controlling both the EOM and the PZT with home-made loop filters.

Part 3 laser is pre-chirped, amplified and width-compressed with an EDFA, yielding ~60 fs pulses with an average power ~300 mW. We splice a highly nonlinear fiber (HNLF) to output of the nonlinear EDFA. An octave spanning supercontinuum (SC) is then generated in HNLF. The SC, shown in Fig. 3(a), covers a span from 1000 nm to 2100 nm for 25 dB bandwidth, and most of the energy distributes in two solitons at 1 μm and 2 μm.

Fig 3. (a) The octave-spanning supercontinuum produced with HNLF; (b) Measured CEO beat SNR in a 300 kHz resolution bandwidth.

The SC light is sent into a common pass $f$-2$f$ interferometer which produces $f_{ceo}$ signal. The signal-noise-ratio (S/N) of $f_{ceo}$ is about 45 dB with 300 kHz resolution, as shown in fig. 3(b). High S/N enables the system generating a clean feedback signal to control the pumping power, for stabilizing $f_{ceo}$ on a RF reference frequency. The coefficient of $f_{ceo}$ versus pump current is 0.2 MHz/mA. Note that a 20-times frequency-divider is used to enlarge the phase-locking tolerance.

Figure 4 shows spectra of controlled $f_{ceo}$ and $f_r$. The servo bandwidths, indicated by the servo gain bumps, are 75 kHz and 150 kHz for $f_{ceo}$ and $f_r$ respectively. The bump shown up at about 35 kHz for both spectra is supposed to be induced by cross-talking between two controlled loops.

Fig 4. RF spectrum of (a) the carrier envelope offset frequency; (b) the optical beatnote between the comb mode and the CW laser.

Frequency stability of the optical frequency combs is the most important parameters. We should observe an out-of-loop beat between two identical comb systems to evaluate the performance of the combs. However, the second independent system is not ready now. Hence, we measure only in-loop frequency stabilities.

Fig 5. In-loop frequency stability of $f_r$ (solid red round) and $f_{ceo}$ (solid black square).

As shown in figure 5, the relative frequency stability of $f_r$ (solid red round) is $\sim2.1\times10^{-16}$ at 1 s, and integrates down to $\sim6\times10^{-20}$ at 10^5 s. The relative frequency stability of in-loop $f_{ceo}$ (solid black square) is $\sim2.9\times10^{-15}/\tau$ for short terms. Such stabilities are good enough to support our frequency comb system measuring transition frequency os the best optical atomic clocks with frequency stability of $\sim3.2\times10^{-16}/\sqrt{\tau}$[14]. Note that the frequency count in use is a Π-type counter from K&K; these results are more than one order of magnitude
worse that those obtained with A-type counter from Agilent, for more high frequency noise contribution observed\(^\text{[15]}\).

**III. CONCLUSION**

In summary, we demonstrate a home-made Erbium-fiber-laser-based 232 MHz optical frequency comb. We phase lock both the comb’s CEO frequency and repetition rate over long time, thanks to the high bandwidth frequency control and high signal-to-noise ratio of frequency signals detection. The relative in-loop frequency stabilities are in the order of $10^{-16}$ at 1s integration time and roll down to $10^{-19}$ level at $10^4$ s. This work paves the way for Sr optical clock frequency measurement and application.

**REFERENCES**


High spectral purity laser characterization with a self-heterodyne frequency discriminator

O. Llopis¹, Z. Abdallah¹,3, V. Auroux¹,4, A. Fernandez¹,2
¹ CNRS, LAAS, Univ. de Toulouse, Toulouse, France
² Univ. de Toulouse, UPS, Toulouse, France
³ CNES, Toulouse, France ; ⁴ OSAT, Toulouse, France
E-mail: llopis@laas.fr

Abstract—The performance of a delay line discriminator dedicated to the characterization of high spectral purity lasers is investigated. The system noise floor is obtained using a symmetrical delay. The vibrations are one of the main problems with this approach at the low frequency offsets. However, the system is able to characterize commercially available narrow linewidth lasers, such as external cavity lasers and fiber lasers.

Keywords—laser; frequency noise; phase noise; metrology; frequency discriminator

I. INTRODUCTION

The today availability of small size narrow linewidth lasers, such as semiconductor lasers optically coupled to high quality factor optical resonators (WGM resonators, fiber resonators...), is an opportunity to develop new embedded applications for which the laser phase noise level is mandatory. This is the case, as an example, of high precision interferometers and of microwave to terahertz signal generation using optics. Before being used in these systems, this type of laser has to be characterized in terms of phase or frequency noise. However, phase noise measurement in the optical domain is not as developed as it is for radio frequency (RF) sources, although some commercial systems are already available (ex : OEwaves). The reason is that it is much more difficult to design a high quality optical frequency synthesizer than a low phase noise RF synthesizer. Therefore, the easier way to measure this type of sources is still to set up the frequency discriminator technique.

In this approach, the laser signal is split in two paths, one path is delayed with a delay smaller than the coherence length of the laser (contrarily to linewidth measurement case) and the signals are combined on a photodiode [1-4]. In the self-heterodyne case, an acousto-optic modulator is added in one of the paths in order to shift the signal frequency around a few tens of MHz. In this case, the noise is analyzed using an RF phase noise test bench at this frequency [3-5], which prevents the detection of some parasitic contributions of low frequency noise, such as the laser 1/f amplitude noise (but also increase the system cost and complexity).

The main difficulty in this approach is to evaluate the measurement noise floor. In microwave frequency discriminators, it is measured using a bypass of the delay line. However, in the optical case, part of the noise may come from nonlinear effects in the fiber [6], or from vibration sensitivity of the fiber spool [7,8] or of the whole system, and the noise floor has to be measured with the fiber maintained in the system.

In this paper, measurements of two types of high spectral purity lasers using a self-heterodyne system are presented. Then, the noise floor of the system is evaluated using a set of two identical fiber spools. Finally, we focus on the choice of the delay and on the importance of the reduction of vibration sensitivity.

II. MEASUREMENT SYSTEM

The measurement system is depicted in Figures 1 and 3. The laser signal is split in two paths. On one path, an acousto-optic modulator shifts the laser frequency of 80 MHz. On the other path, the signal is delayed using a fiber spool, and then recombined to the first path. A polarization controller is added after the fiber spool in order to improve the output signal level. The beat frequency is recovered using a photodiode, amplified and, finally, feeds an RF phase noise measurement bench (in our case, an Agilent E5052B).

Fig 1: Self heterodyne frequency noise measurement system

Using this approach, the measured RF phase noise at 80 MHz is proportional to the laser frequency noise, at least at offset frequencies close to the carrier [3,4]. Far from the
carrier, the output spectrum is attenuated with a classical $\sin x/x$ shape. This is clearly visible in Figure 2, in which the phase noise of the 80 MHz output is depicted in case of an external cavity semiconductor laser (RIO Inc.) feeding the bench, and using a 2 km fiber spool in the bench. At 100 kHz, the measurement sensitivity locally drops down to zero because of the first zero of the $\sin x/x$ function.

Because of this response, we have limited our measurements on the $1 \text{ Hz} - 100 \text{ kHz}$ offset range, and we have corrected the measured data using the reverse function of the $\sin x/x$ response of the bench and the time delay of the spools we are using. More precisely, the laser frequency noise can be computed from the measured 80 MHz RF phase noise [3] using equation 1.

$$S_{\Delta f \text{ in } Hz/\sqrt{Hz}} = \frac{f_m}{\sqrt{2} \sin(\pi f_m \tau)} 10^{L_{\text{RF}}(f_m)/20}$$ (1)

were $f_m$ is the offset frequency, $\tau$ is the time delay between the two paths and $L_{\text{RF}}(f_m)$ is the measured single sideband phase noise, in dBc/Hz, of the RF signal.

Another important result of equation 1 is that it gives the proportionality coefficient between the RF phase noise and the laser close to carrier frequency noise, and that this coefficient is itself proportional to the delay $\tau$. Thus, even if a large delay limits the measurement bandwidth because of the $\sin x/x$ effect, it enhances the measurement sensitivity and, as we will see, reduces the noise floor.

Equation 1 is used to compute the laser frequency noise data. However, when a noise floor is measured instead of a true laser frequency fluctuation, this noise floor is not affected by the $\sin x/x$ response, because it is either a noise superimposed on the system output, or the demodulation of the laser amplitude noise. Therefore, to convert the measured noise floor into frequency noise, we use a constant coefficient which is the limit of the above coefficient at low frequencies.

$$S_{\Delta f \text{ noise floor}} = \frac{1}{\sqrt{2} \pi \tau} 10^{L_{\text{RF}}(f_m)/20}$$ (2)

This frequency discriminator has been used to measure two different high spectral purity lasers: a semiconductor laser stabilized on an external resonator, available in a small butterfly package (RIO), and a laboratory high quality fiber laser (Koheras). Both lasers are used at a relatively low output power: 7 mW (60 mA) for the RIO laser and 16 mW for the Koheras laser. The power dependence of the spectral purity is relatively weak for this type of lasers, and such an output power is largely sufficient to drive our bench (however, a small improvement of about 2 to 3 dB is observed on the RIO laser phase noise when biased at 120 mA instead of 60 mA).

Figure 4 depicts the result of the measurements performed on these two lasers, under these conditions. The measurement is performed in a Faraday’s shielded room using the E5052 system and the frequency noise is computed using equation 1.
In Figure 5, the same measurements are plotted in terms of single sideband phase noise, which shows the difference in dB between these two lasers, and also give an idea of the laser spectral shape, taking into account that when the phase noise is sufficiently low compared to 0 dB, it represents the lateral noise wings of the spectrum (we do not want to compute the linewidth, because this parameter is too much dependent on the integration time and has, in our opinion, no rigorous meaning for this type of high spectral purity optical sources).

In a preceding paper [4], the phase noise of the Koheras laser had been a few dB overestimated in the 10 Hz-100 Hz range, because of a vibration noise sensitivity of the fiber spool we were using. Indeed, like in any type of noise measurement, it is essential to determine if the measured data are not too close from the measurement bench noise floor. The next paragraph is dedicated to the evaluation and the optimisation of this noise floor.

III. MEASUREMENT NOISE FLOOR

An obvious technique to measure a frequency discriminator noise floor is to cancel the delay \( \tau \) by removing the fiber spool. In this case, the system sensitivity to the laser frequency noise is null, and only remains the other noise contributions. However, such an evaluation of the noise floor is quite optimistic. Indeed, it also cancels any noise contribution coming from the fiber spool. There are mainly three types of low frequency noise that can be generated in a fiber spool: noise related to optical nonlinear effects [6], such as Stimulated Brillouin Scattering (SBS); diffusion noise on impurities, such as Rayleigh scattering noise; noise due to a vibration sensitivity of the spool [7,8].

Therefore, in order to get a more precise evaluation of the noise floor, two identical fiber spools have been used to cancel the delay \( \tau \), one in each path of the interferometer. In this case, the above described effects can still be observed, and are included in the real noise floor.

The plastic rod spool used in our first experiments [4] has been replaced by a packaged spool (Fig. 3). The laser is free of constraint in this package, but it is however mechanically maintained thanks to the relative small size of the package and to an inside plastic protection. Two set of spools have been used, one of 1 km (\( \tau = 5.1 \mu s \)) and one of 2 km (\( \tau = 10.05 \mu s \)).

Figure 6 depicts the results of the noise floor measurements in the 2 km fiber spool case, with the battery biased RIO laser feeding the bench (7 mW optical input power). A clear contribution of vibration noise is observed between 10 Hz and 2 kHz when the system is set up on a normal table (green curve). When the system is put on a low vibration table (but still inside the Faraday shielded room), most of the vibration contribution is reduced and the noise floor measured with the two spools technique (red curve) is close to the noise floor measured with the spool removed from the bench (two harms with short delay; black curve).

These observations allow us to clarify the importance of the vibration noise contribution to the noise floor, and to demonstrate the absence of non-linear optical noise contribution. Indeed, in both cases (Koheras or RIO lasers), we are feeding the fiber spool with optical power levels much below the SBS threshold, which is in the range of 30 mW for a 2 km fiber spool [9].

The same experiments have been performed with the 1 km spools. On Figure 7, it is clear that the vibration problem is still observed with the 1 km spool. Finally, with this configuration, the noise floor is increased of about 6 dB, which is directly related to the reduction of the delay \( \tau \) of a factor 2, and thus of the same reduction in sensitivity.

Finally, the 2 km spool appears to be a good compromise between a good sensitivity, a high SBS threshold and a large offset frequency bandwidth, and this fiber length is selected for further experiments.

The last question is: what is limiting the noise floor when the fiber spool is removed. At high frequency offsets, it is
clear that the signal to noise ratio may have an influence, and this signal to noise ratio is mainly determined by the laser RIN at the IF frequency (80 MHz). However, close to the carrier, we also found an influence of the oscillator which is driving the acousto-optic modulator. This oscillator features a -43.5 dBc/Hz phase noise level at 1 Hz, with a 30 dB/dec slope in this region. It determines the optical frequency discriminator noise floor between 1 Hz and 5 Hz offset, as shown in Figure 8. However, between 5 Hz and 2 kHz, the noise floor seems still to be determined by vibration sensitivity. Finally, between 2 kHz and 100 kHz, the noise floor changes with the laser type and power, which means that it should be mostly related to the laser RIN.

The next step is now to improve the experimental set up to overcome these limitations. The problem of the acousto-optic modulator driver is easy to solve: this source will be replaced by a high spectral purity oscillator (a high quality OCXO) followed by a high power amplifier, and this should lead to a 30 dB improvement of the 1 Hz phase noise (case of a vibration free experiment). However, the noise between 10 Hz and 2 kHz will be more difficult to improve: it is directly related to the interferometer sensitivity to vibrations. A new set up will be designed, featuring higher compactness, and on which mechanical isolation will be easier to provide.

IV. CONCLUSION

The measurement of two different high spectral purity lasers has been performed with a delay line frequency discriminator. The problem of the measurement noise floor has been carefully studied and it has been found that it is mainly limited by the mechanical vibration sensitivity of the fiber spool, and also of the whole interferometer set-up. The system sensitivity and dynamic range is however sufficient to investigate on the two examples of lasers presented in this paper. However, for ultra-high spectral purity lasers characterization, the set-up must still be improved.

REFERENCES

Abstract—We have implemented a long-term frequency stabilization system for external cavity diode laser (ECDL) based on mode boundary detection method. In this system, the saturated absorption spectroscopy was used. The current and the grating of the ECDL were controlled by a computer-based feedback control system. By checking any mode boundaries in the spectrum, the control system determined how to adjust current to avoid mode hopping. This procedure was executed periodically to ensure the long-term stabilization of ECDL in the absence of mode hops. This laser diode system without antireflection-coating had over operated in the condition of long-term mode hopping free stabilization for almost 400 hours. Last year, we have added template matching technics to the system. The ECDL stabilization time has been extended to almost 1200 hours without any manual intervention further. This is a significant improvement of ECDL frequency stabilization system. This technique is very useful in some applications such as laser atomic cooling and atom fountain etc.

Keywords—ECDL; long-term; frequency stabilization; mode boundary detection;

I. INTRODUCTION

Frequency stabilization for external cavity diode laser (ECDL) plays an important role in quantum optics, cold atom physics and optical clock [1], [2]. It is one of the key technologies in modern precision measurements. Frequency stabilization for the diode laser has been widely implemented. However, for ECDL it requires the precise control of temperature, current and grating, as external optical feedbacks which exist in ECDL. The long-term frequency stabilization for ECDL especially with non-antireflection-coated (non-AR-coated) diode is more difficult to achieve due to mode hopping.

Theoretically, mode hopping is formed by the asynchronous drift of internal mode, external mode and grating profile of ECDL [3], [4], due to the drift of temperature and other influences such as aging of diode.

Currently, several methods have been used to suppress mode hopping. Antireflection (AR) coating of the laser diode’s output facet [5] and simultaneous adjustment of the diode current and grating in an appropriate ratio, which is called feed forward technique, can broaden continuous tuning range with mode-hop-free [6], [7]. Improving the stability of temperature controller can minimize the impact due to environment changing. But these methods mentioned above are passive methods, which merely mitigate the disturbance caused by mode hopping and postpone the occurrence of mode hopping. They cannot suppress mode hopping radically. Some long-term influences caused by aging of diode and mechanical structure gradual deformation cannot be eliminated by these methods. This makes it difficult to stabilize ECDL frequency for a long time. Active methods such as current feedback by monitoring the frequency or amplitude of noise [8] can stabilize ECDL for 2 to 12 days, but its means of monitoring current is complicated and restrict conditions for photo detector are required. With high-performance control circuits, the continuous stabilizing time can be extended to 100 hours [9]. Moreover, an automatic system to control the operation of ECDL can relock the frequency when it loses locking [10]. This system requires the spectrum and its differential signal to be simple shaped, otherwise the automatic function cannot work properly.

In this paper, we present a novel technique for long-term frequency stabilization, which is based on the mode boundary periodical detection and template matching. A computer-based feedback control system for ECDL is designed and implemented. We used a diode with non-AR-coated. The system had stabilized the ECDL frequency for 1200 hours without any without any manual intervention.

II. MODE BOUNDARY

Mode hopping can be easily discerned from the spectrum. The spectrum is a continuous waveform when the laser diode works within a single mode in the grating scanning range, even at the peak of the saturated absorption spectrum. Nevertheless there would be steps in the spectrum if the diode operates within two or more modes. The steps indicate that the diode mode hopping occurs in the grating scan range and there are two modes with different frequency beside the step. We refer to this step as mode boundary, which is shown in Fig. 1. Part a.

The ECDL frequency is usually stabilized and locked to the peak of the saturated absorption spectrum by Lock-in or doppler-free Dichroic Atomic Vapor Laser Locking (DAVLL) method through grating feedback circuit. The relative position between mode boundary and the frequency locked peak would drift with the long-term influences of temperature fluctuation,
aging of diode and mechanical structure gradual deformation. This drift cannot be restrained by the single grating feedback method at all. When the mode boundary comes into contact with the frequency locked peak, the frequency error signal which is essential to the feedback circuit will be disturbed, and the locked frequency will be lost.

Fig. 1. Mode boundaries in spectrum (Part a) and its differential signal (Part b).

The mode hop often occurs when the non-antireflection-coated diode is used, because the mode-hop-free spectrum range of this type of diode is narrow, about 2-3 GHz or less. It means that the distance between two mode boundaries is very short. The locked frequency of ECDL is often to lose in half or one hour when the mode boundary comes into contact with the locked peak at little drift of the temperature or other parameters, even if the locked peak is at the very middle of two mode boundaries when the feedback system started.

III. SYSTEM DESIGN

A. Mode boundary detection and elimination

Checking the differential signal of the spectrum is an effective way to detect the mode boundary by a computer. The differential signal of a step is infinite impulse in math, whereas it is an impulse with high value much higher than that of the peak of saturated absorption spectrum by Lock-in or numerical calculating method. Therefore, it is the criteria of judgment of occurrence of mode boundaries in the spectrum to check whether there are any impulses with high value higher than preset value. Once the high impulse appears, the mode boundary should be detected.

Fig. 1 shows the spectrum and its corresponding differential signal. Part a is the saturated absorption spectrum of 87Rb and 85Rb with mode boundaries in it, as part b is the differential signal calculated by a computer program. From the waveform, the impulse corresponding to mode boundaries in differential signal is much higher than that of peak of saturated absorption. So calculating differential signal by a program is precise enough to detect mode boundaries.

The detailed standard of detection is shown in Fig. 2. The main curve is the differential signal and point B is the locked point. In the range of hundreds of MHz at two side of the locked point, two reference levels are set, shown as two broken lines in the figure. According to the shape of differential signal, the reference levels are set above and below differential signal, to form an incompact outline of differential signal. The situation that any signal oversteps the reference levels is judged as mode boundaries have been detected. This is easily accomplished in a program. Besides, it is not necessary to set the reference levels too close to the differential signal to avoid misjudgment, for the misjudgment is occasionally caused by unexpected noise pulses and signal distortion due to different working point of laser diode’s current.

After mode boundary has been detected, the regulating measures to suppress mode hopping should be taken to prevent mode boundary from moving to the frequency locked peak further. Both current and temperature adjustment are effective methods, while the current adjustment is more sensitive and quick. Current tiny adjustment can withdraw mode boundary rapidly and precisely.

In addition to this, we have added an ordinary technics of template matching to the system. Under this technics, the stabilization system can recognize the spectrum if it is the appointed saturated absorption and some adjustment will be carried out if necessary.

Fig. 2. The reference levels of mode boundary detection.

B. Periodical detection procedures and coordination program

In the method above, an indispensable step is to obtain the spectrum, otherwise it is unknown whether mode boundary is getting close to locked peak. But during the normal frequency stabilizing operation of ECDL, there are no any methods but scanning grating at least once to obtain the spectrum. The frequency has been changed during the grating scanning, but this arbitrary change of frequency is not allowed in the experiments. For example, in our experiment for atomic fountain, laser frequency must be kept stabilized strictly
without any scanning or unlocking during a whole procedure, which lasts for 5 seconds including trapping, cooling, launching and detection of the atoms. Apparently, the grating scanning is forbidden, but we designed a periodical detection procedures and coordination program and had solved this problem successfully.

A communication line was set between computer A for controlling the atomic fountain experiment and the computer B for stabilizing frequency of ECDL to coordinate their sequence signal. During the operation, computer B performs the normal PID frequency stabilizing program to provide frequency stabilized laser beam for atomic fountain, when computer A carries out the atomic fountain procedures. After one of more whole procedures of atomic fountain (we set 2 whole procedures lasting for 10 seconds), computer A sends a signal to computer B that the atomic fountain procedures have been finished and the grating scanning is allowed now, and computer A turns into a waiting step immediately. Once computer B has received the signal, it stops the normal PID and turns into grating scanning immediately to obtain the spectrum, which is centered at the current grating offset and cover a few hundreds of MHz range. Computer B calculates the differential signal from the spectrum to detect mode boundary, and calculates its relative position to frequency locked peak. Then possible current adjustment is gotten ready. Finally, normal PID restarts and the frequency is locked to the former peak of saturated absorption signal and the prepared current adjustment is performed to eliminate the mode boundary. At the same time, computer B sends a signal to computer A that the detection procedures have been finished and the atomic fountain could be continued. The whole detection procedure lasts for 0.1 seconds in our setting, which is mainly determined by the speed of grating scanning. In this way, these two computers work automatically, circularly and synchronously through the coordination program. The whole detection procedure lasts for 0.1 seconds. Thus, there is little possibility that the frequency loses locking instantaneously.

This method requires that the procedures of the experiment such as atomic fountain, are interruptible, otherwise it is impossible to detect mode boundary. For instance, pause could be inserted between every procedure for atomic fountain in the experiment. This requirement constrains its application to certain extent. But it can be met in most of the experiments in quantum optics and cold atom physics. It is also acceptable to modify the procedure to insert MBPD step, in order to improve the frequency long-term stability, even if the quondam procedure is consecutive.

C. Experiment setup

The frequency stabilization system based on MBPD was designed. The system is composed of ECDL and its driving modules, optical system with saturated absorption technique, lock-in module and computers. The schematic of the setups is shown in Fig. 4.

The ECDL is a commercial laser with Littrow configuration sold by UniQuanta Co., Ltd. with a non-AR-coated diode in it. It works in 780 nm, and the line width is about 600 kHz in freely running tested by spectrum analyzer through frequency beat with another ECDL. A small part of the output laser beam was injected into the photo detector through a saturated absorption setup to generate the spectrum signal, and the majority of the beam is supplied to the atomic fountain experiment.

Current module and grating’s piezoelectric-transducer (PZT) driving module are controlled by computer B. Temperature controller was developed by ourselves. Lock-in amplifier

![Fig. 3. Programs’ flow chart](image-url)
(Femto LIA-MV-200-H) generates sine wave which is added to PZT module for modulation, and we set the frequency to 10 kHz. It also receives the signal from the photo detector and outputs the error signal by phase sensitive technique for feedback control system.

![Image: The schematic of experimental setup](image)

The feedback control system is implemented by a common computer with the plugged PCI card of the model PCI-6229 produced by NI, Inc. The signal from photo detector and Lock-in amplifier are converted to digital value by A/D converter on PCI card and are transmitted to computer B. The digital calculating results are converted to analog signal by D/A converter to control the PZT and current module. Normal PID stabilizing program, MBPD program and coordination program run on the computer B. These control programs are designed and implemented on Labview platform. Computer A carries out the atomic fountain procedures.

**IV. PERFORMANCE**

This system can resist many disturbances, which are caused by the temperature fluctuation and other environmental factors, from mode boundary successfully, and eliminate the mode boundary rapidly and effectively. With this excellent trait, the system has good performance in long-term stability with mode-hop-free. In the experiment, we installed the non-AR-coated diode in ECDL and started the control system with locked frequency appointed at the saturated absorption peak of hyperfine transition CO 2-23 of 87Rb. Under the automatic control of the system, the frequency had been stabilized for 1200 hours. During the operation, mode boundary got close to frequency locking peak for hundreds of times, but the current adjustment eliminated it successfully. The stabilized time of ECDL with non-AR-coated diode under this system was improved almost 1000 times than that of the old normal feedback circuit.

There are two ways to evaluate the system stability, Allan variance and error signal. It is not mathematically rigorous to use Allan variance in this system, because periodical scan is inserted between normal PID stabilization in MBPD. Error signal is not as precise as Allan variance, but it can indicate any frequency fluctuation and excursion relative to locked point in time domain, and it also can be used for estimating line width. So here we use error signal to evaluate system stability.

Fig. 5 shows the time traces for 1200 hours of mode-hop-free operation with frequency locked. Part a is the error signal of frequency locked acquired from Lock-in Amplifier.

![Image: Mode-hop-free operation with frequency locked lasting 1200 hours](image)

According to the slope calculated from the middle of the scanning error signal (point B in Fig 2.) and the level width of locking error signal, we estimate the laser line width is about 4 MHz.

Part b in Fig. 5 is the PZT voltage. It indicates the operation of normal PID frequency stabilization. The partially correlation of the these parameters are shown, but the correlation is not perfect due to the other environmental influences, and this is the reason why the simultaneous adjustment of current and grating in a simple ratio (feed forward technique) is unable to suppress mode hopping radically. Not only the virtue of this system is that the continuous time mode-hop-free operation is extended considerably, but also the rigorous demand for stability of temperature controller has been lowered extremely.

**V. DISCUSSION**

The system based on the MBPD and coordination program is proved to be effective and the goal for long-term frequency stabilization has been achieved successfully. The key of MBPD method is to prevent the disturbance before it occurs. The previous periodical detection and measures for suppressing mode hopping are taken. From this perspective, feed-forward approach has also been embodied.

MBPD method can be used in system in which the error signal is generated by other technique such as DAVLL etc., not
just confined to modulation combined with lock-in technique in
this experiment.

In some applications, high stability of laser power is
required. Thus, temperature adjustment can be taken
simultaneously with current adjustment. Both the frequency and
power of ECDL may be stabilized in long-term at the same
time, for the temperature adjustment has the same effect with
current adjustment.

REFERENCES

1998.
et al., “A compact grating-stabilized diode laser system for atomic
3675-3685, June 2005.
[7] C. Petridis, I. D. Lindsay, D. J. M. Stothard, and M. Ebrahimzadeh,
“Mode-hop-free tuning over 80 GHz of an extended cavity diode laser
without antireflection coating,” Rev. Sci. Instrum. vol. 72, pp. 3811-
3815, October 2001.
[8] Sheng-wei Chiow, Quan Long, Christoph Vo, Holger Müller, and
Steven Chu, “Extended-cavity diode lasers with tracked resonances,”
[9] Tong Zhou, Xianghui Qi, Qing Wang, Wei Xiong, Jun Duan et al.,
“Frequency-stabilized diode laser at 780 nm with a continuously locked
system to control the operation of an extended cavity diode laser,”Rev.
Faraday Anomalous Dispersion Optical Filter at 461nm Utilizing a Strontium Hollow Cathode Lamp

Duo Pan, Xiaobo Xue, Xiang Peng, Jingbiao Chen, and Hong Guo
State Key Laboratory of Advanced Optical Communication, Systems and Networks
School of Electronics Engineering and Computer Science
Peking University
Beijing 100871, China
Email: jbchen@pku.edu.cn, hongguo@pku.edu.cn

Bin Luo
State Key Laboratory of Information Photonics and Optical Communications
Beijing University of Posts and Telecommunications
Beijing 100876, China
Email: luobin@bupt.edu.cn

Abstract—We realize a Faraday anomalous dispersion optical filter (FADOF) operating on the $^{88}\text{Sr} (5s^2)^1S_0-(5s5p)^1P_1$ transition, utilizing a strontium hollow cathode lamp (HCL). At the magnetic field strength of 1460 G and the HCL discharge current of 22 mA, a single transmission peak with a maximum transmission of 28.8% is obtained. The dependence of transmission on magnetic field and HCL discharge current is qualitatively studied. This demonstration will be used to build a Faraday laser, which will be applied in the Sr optical clock.

Keywords—Faraday effect; filters; strontium; hollow cathode lamp; transmission.

I. Introduction

Faraday anomalous dispersion optical filter (FADOF) [1,2] has advantages of ultra-narrow bandwidth [3], high transmission, and high noise rejection [4,5], which makes it an excellent frequency selection component widely used in free-space laser communication [6-8] and lidar remote sensing system [9-11]. In addition, the FADOF can also be used for laser frequency stabilization [12, 13] and directly made into an optical clock [14]. Electrodeless lamps and hollow cathode lamps are also utilized in FADOFs. Although they may not be suitable for the filtering of weak light because of the fluorescence background, the abundant atomic transitions they provide bring them potential applications for frequency stabilization [15, 16].

In this paper, a Faraday anomalous dispersion optical filter at $^{88}\text{Sr} (5s^2)^1S_0-(5s5p)^1P_1$ transition (Fig. 1) is demonstrated utilizing a Sr hollow-cathode lamp (HCL). A single transmission peak with a maximum transmission of 28.8% is obtained, at the magnetic field strength of 1460 G and the HCL discharge current of 22 mA. The dependence of transmission on magnetic field and HCL discharge current is qualitatively studied.

In the current available Sr optical clocks, 461 nm external cavity diode lasers (ECDLs) are commonly used to produce probing beams. But these lasers are very sensitive to the outside vibrations, and must be set to exact temperature and current values to keep an accurate output frequency. The FADOF demonstrated in this paper will be used to build a Faraday laser, which is available in the Sr optical clock. The output frequency of the Faraday laser will be limited at the $^{88}\text{Sr} (5s^2)^1S_0-(5s5p)^1P_1$ transition, making it immune to fluctuations of current and temperature.

Fig. 1. (Color online) Energy diagram of $^{88}\text{Sr} (5s^2)^1S_0-(5s5p)^1P_1$ transition.

II. Experimental setup

The experimental setup of the FADOF is displayed schematically in Fig. 2. The configuration in the dashed box is a FADOF using $^{88}\text{Sr}$ hollow-cathode lamp, with H1 and H2 being a pair of permanent magnets, and P1 and P2 being a pair of crossed polarizer, and the extinction ratio is $1 \times 10^{-5}$. A 461 nm diode laser provides the probing beam, and the transmission spectrum is detected by a photodiode (PD).

Fig. 2. (Color online) Experimental setup of a FADOF utilizing Sr hollow-cathode lamp. 461 nm laser: the probe laser, P1 and P2: crossed polarizer, H1 and H2: permanent magnets, PD: photodiode.
The Sr HCL used as the atomic source can easily adjust vapor pressure by regulating the discharge current. It contains inside a ring-shaped anode and a cathode having a length of 20 mm and a bore diameter of 3 mm. Compared with the widely used Sr atomic ovens [17,18], it has significantly reduced the size and cost of experimental systems.

III. Results and discussion

First a 461 nm probing laser is used to measure the optical density (OD) of the HCL without magnetic field. The spatial distribution of the Sr atoms in the HCL is not uniform, and it contains more atoms near the wall of the cathode [19, 20]. In our experiment, the probing laser beam interacts with atoms at 0.75 mm from the center, with the spot size to be 0.5 mm × 0.7 mm. The absorption spectrum of the $^{88}\text{Sr} \left(5s^2\right)^1S_0 - \left(5s5p\right)^1P_1$ transition at different discharge current is shown in Fig. 3. The fluorescence background of the HCL (for example, 36 mV at the discharge current of 20 mA) has been removed from the detected intensity. With the formula $\text{OD} = \ln(P/P_0)$, the optical density is calculated to be 0.50, 1.59, 2.49, and 3.85 under the discharge current of 5 mA, 10 mA, 15 mA, and 20 mA, respectively. The change of the voltage at far detuning is due to the power change of probe laser when sweeping the frequency.

![Absorption Spectrum](image)

Fig. 3. (Color online) Doppler broadened absorption spectrum of the Sr $^1S_0 - ^1P_1$ transition under different discharge currents, measured at 0.75 mm from the HCL center.

Then we measured the transmission spectrum under different HCL discharge current and magnetic field. Fig. 4 depicts the transmission as a function of the discharge current, with fixed magnetic field and probe power of 1460 G and 9.3 mW. Fig. 4 (a) shows the transmission peaks under different discharge currents, and Fig. 4 (b) shows the transmittance versus the discharge current. When the current is relatively low, atomic density is the main limitation of the transmittance, and the transmittance rapidly increase with the current. However, since the increasing of discharge current causes more absorption of probe laser, the growth rate of transmittance slows down and the transmittance gradually tends to a stable value.

![Transmission Spectrum](image)

Fig. 4. (Color online) The transmission as a function of the discharge current, with fixed magnetic field and probe power of 1460 G and 9.3 mW. (a) Transmission peaks under different discharge currents. (b) The transmittance versus the discharge current.

This current, the transmittance can be rapidly improved by increasing the magnetic field strength. When the magnetic field is small, a twin-peak transmission spectrum is obtained. With the magnetic field increasing, the transmission spectrum gradually changes into single-peak. A maximum transmittance of 28.82% is obtained at the magnetic field strength of 1460 G and the HCL discharge current of 22 mA. It seems that if increasing the magnetic field, the transmittance can be further improved.

The relevant theory is being calculated. The FADOF of alkali metal already have a quite maturational theory model, but due to the difference of level structure, the Sr FADOF theory should be more similar to that of casium [21].

IV. Conclusion

In this paper, we have demonstrated a Faraday anomalous dispersion optical filter at Sr (5s$^2$)$^1S_0 - (5s5p)^1P_1$ (461 nm) utilizing a Sr hollow-cathode lamp. A single transmission peak with a maximum transmission of 28.8% is gained, at the magnetic field strength of 1460 G and the HCL discharge current of 22 mA. The dependence of transmission on magnetic field and HCL discharge current is qualitatively studied.
Transmission peaks under different magnetic fields.

The transmittance versus the magnetic field strength.

Fig. 5. (Color online) The dependence of the transmission on the magnetic field strength, with fixed HCL discharge current and probe power of 22 mA and 9.3 mW. (a) Transmission peaks under different magnetic fields. (b) The transmittance versus the magnetic field strength.

When combined with the all-optical locking technique [13], the FADOF can be used as a frequency selector, with the transmission signal providing feedback to the laser diode. So we can realize a compact frequency-fixed laser system, in which the laser would work immediately at the $^{88}$Sr $(5s^2)^1S_0 - (5s5p)^1P_1$ resonance line when turned on. For future work, we will build a Faraday laser system [12, 14] working at 461 nm resonance line used for $^{88}$Sr optical clock. The frequency of this laser will be immune to the fluctuations of injection current and laser diode temperature.

REFERENCES


A Cavityless Laser Using Cesium Cell with 459 nm Laser Pumping

Xiaobo Xue, Duo Pan, and Jingbiao Chen*
State Key Laboratory of Advanced Optical Communication System and Network, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China.
*E-mail: jbchen@pku.edu.cn

Abstract—In this paper, we present a laser scheme without cavity in a 10 cm cesium cell as a potential active optical frequency standard. Since no cavity is applied, the stimulated light is emitted from the atoms directly and the center frequency of output laser light is solely determined by atoms without cavity pulling. The threshold characteristic of the emitted light is measured. This scheme can be used to investigate the relation between stimulated emission and amplified spontaneous emission. It is expected to be further extended to generate other narrow linewidth frequency standard signals.

Keywords—amplified spontaneous emission, cavityless lasers, frequency standards.

I. INTRODUCTION

The concept of active optical clock [1], [2] is proposed not long ago and is expected to provide laser light source with potential much narrower linewidth than that of conventional cavity stabilized lasers for optical clocks, since many cavity noises, including the Brownian thermal noise [3], [4], can be reduced by orders of magnitude. Some practical schemes for building active optical clocks have been proposed and experimentally investigated. A previous experiment based on Cs four-level system in an atomic cell [5], [6] is done and the linewidth of the emitted laser light is 331 Hz. The essential difference in principle from original frequency standards is the application of stimulated emission in “bad-cavity” with a relative low finesse, the linewidth of which is much wider than the gain linewidth of atoms.

To investigate the effect of the “bad-cavity” a laser scheme without cavity using a cesium cell is proposed and realized in this paper. We name this a “cavityless laser” temporarily. This scheme can be used to investigate the relation between the stimulated emission and the amplified spontaneous emission [7], [8], [9]. It is expected to be further extended to generate other much narrower linewidth frequency standard signals.

II. SCHEME OF A CAVITYLESS LASER

The Cs four-level system is chosen because some research has been done earlier and the pumping lasers are prepared. The energy levels are shown in Fig. 1. A beam of laser operating at 459 nm will pump the atoms from ground state to 7P1/2 state. The atomic system will radiate 1470 nm fluorescence through spontaneous radiation. A mirror (M1 in Fig. 2) with 459 nm high reflection coating is applied to amplify the spontaneous radiation and generate a beam of 1470 nm cavityless laser. The 1470 nm cavityless laser light will beat with another 1470 nm laser light that generated from an active optical frequency standard at bad cavity regime.

III. EXPERIMENTAL SETUP

An experiment is built to investigate the threshold characteristic and the linewidth of 1470 nm laser. The experimental setup is shown in Fig. 2. The 459 nm laser is an external cavity diode laser in littrow configuration used as a pumping laser. The Cs cell is 10 cm in length and heated by a heater from 150°C to nearly 180°C. M1 is a coated concave mirror with 459 nm laser light anti-reflection and 1470 nm laser light high-reflection coating. The radius of curvature is 8000 mm. M2 is a 1470 nm filter that isolates the 459 nm laser light and transmits the 1470 nm laser light. PD1 and PD2 are commercial photodetectors. An iris diaphragm is applied to modify the size of the light spot. The light spot is set to be 1.4 mm in horizontal and 0.7 mm in vertical when the photodetector detects the biggest signal.

Fig. 1. (Color online) The relevant energy levels of the Cs cavityless laser.
IV. Experimental Results

The 1470 nm cavityless laser light is observed. The resonance spectrum, the threshold characteristic and the effect of amplified spontaneous radiation are measured.

A. Resonance Spectrum

The 459 nm pumping laser is swept in frequency within a range of about 12 GHz, which can cover the 9.2 GHz hyperfine splitting of the ground state. The mode hopping free range is larger than 16 GHz ensuring that no mode jump exists within the sweeping. The 1470 nm cavityless laser light is observed when the frequency of the pumping laser light resonates with Cs 6S_{1/2} (F=4) to 7P_{1/2} (F’=3) transition. The 1470 nm resonance spectrum is shown in Fig. 3. Besides the transition from 6S_{1/2} (F=4) to 7P_{1/2}, a very weak 1470 nm fluorescence can also be observed when the 459 nm laser frequency is resonant with 6S_{1/2} (F=3) to 7P_{1/2}.

B. Threshold Characteristic

The threshold characteristic of the output laser is measured. Fig. 4 shows a typical threshold curve at 170°C. The threshold power of the pumping laser is about 0.9 mW. Stabilizing the pumping laser power at 2.4 mW, which is the maximum output power after the iris diaphragm, the maximum 1470 nm laser output power is measured under different Cs cell temperature, as shown in Fig. 5. From the figure one can see that the maximum power occurs between 166°C and 170°C.

V. Discussions

A. Linewidth of 1470 nm Laser

To investigate the linewidth of the 1470 nm laser, we make the 1470 nm laser light beat with another 1470 nm laser light that generated from an active optical frequency standard [5](see Fig.6). The beat signal hasn’t been observed for now. Some possible reasons may be that the pumping lasers have too much noise. Furthermore, the power of the 1470 nm laser for beat is too small, which is about several µW.

B. Influence from the Size of the Light Spot

In this paper, the light spot is 1.4 mm in horizontal and 0.7 mm in vertical. This is because at a small size, the 1470
nm output laser is relatively easier to observe. While from the experimental data earlier, when there is no iris diaphragm, the 1470 nm output laser can still be observed. The threshold characteristic was also measured when pumped with a larger spot size (3.2 mm in horizontal and 1.3 mm in vertical), as shown in Fig. 7. The reason why the adjustment is harder with a bigger laser spot size is not clear for now. We will do more work to investigate this issue.

C. Generation of the 1470 nm Cavityless Laser Light

It seems that the generation of the 1470 nm cavityless laser results from the effect of amplified spontaneous emission, but the laser light has a very good collimation, that the intensity of the 1470 nm signal doesn’t decrease at a longer distance. Meanwhile the intensity decrease quickly when moving the detector in the plane that perpendicular to the direction of pumping laser. From this phenomenon one can see that the 1470 nm laser is more likely to be generated by simulated emission. To prove this idea, the ratio of 1470 nm cavityless laser power to 1470 nm fluorescence power is calculated from the experimental data. Here the 1470 nm laser power is represented by the intensity of the highest peak in the resonance spectrum (see Fig. 3), and the 1470 nm fluorescence power is represented by the intensity of the highest peak in the resonance spectrum. The result is shown in Fig. 8. With 1470 nm laser power increasing, more atoms will transit from $7S_{1/2}$ to $6P_{1/2}$ state through stimulated emission process which is emitted by 1470 nm photons from the reflection of M1 with increased spatial coherence. This also needs to be further investigated. Moreover, to measure the frequency of the 1470 nm laser light, a 1470 nm laser will be applied to distinguish the transition.

VI. CONCLUSIONS AND FUTURE WORK

In conclusion, a cavityless laser operating at 1470 nm is built to study its possibility as potential active frequency standard. The maximum output power is about 20 µW at about 170°C with pumping power 2.4 mW when the frequency of the pumping power is resonate with Cs $6S_{1/2}$ ($F=4$) to $7P_{1/2}$ ($F'=3$) transition. This system can be used as a tool to investigate the relation between active optical frequency standards working in “bad-cavity” condition and amplified spontaneous emission. It is expected to be further extended to generate other narrow linewidth frequency standard signals. For future work, the origin and linewidth of the 1470 nm laser need to be investigated. A 1470 nm external cavity diode laser will be used to precisely measure the frequency of the 1470 nm laser that emitted from the cavityless laser. Comparing with previous mirrorless lasing [8], our interest is in the potential spectrum purity and accuracy in the view of active optical frequency standard. This cavityless laser will be used as a tool to investigate the difference and evolution of linewidth and lasing spectra between lasers with bad cavity and lasers without cavity.

ACKNOWLEDGMENT

We thank Junhui Li for useful discussions in analyzing the experimental results.

REFERENCES


Active Optical Frequency Standard Based on Narrow Bandwidth Faraday Atomic Filter

Xiaogang Zhang and Jingbiao Chen
State Key Laboratory of Advanced Optical Communication System and Network,
School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China
Email: ganggunther@gmail.com, jbchen@pku.edu.cn
Wei Zhuang
National Institute of Metrology,
Beijing 100013, China
Email: zhuangwei@nim.ac.cn

Abstract—Active optical frequency standard based on narrow bandwidth Faraday atomic filter is proposed. The output laser power achieved to 42 μW. The laser frequency is determined by narrow bandwidth Faraday atomic filter, corresponding to the Cs 6S_{1/2} F=4 to 6P_{3/2} F'=4, 5 crossover transition line. The transmission of the Faraday atomic filter is 23% and the bandwidth is 52.6MHz. Considering the loss in the cavity, the linewidth of the cavity mode is 345.1MHz. The bad cavity coefficient is 6.6. Working in the bad cavity scheme, the output laser signal of active Faraday optical frequency standard is insensitive to the cavity vibration and thermal noise. Besides, we redesign an integrated structure to decrease the influence of the environmental vibration. With more steady structure, hundreds of millihertz level of the active Faraday optical frequency linewidth is expected.

Keywords—Active optical frequency standard; Faraday atomic filter; Bad cavity scheme;

I. INTRODUCTION

Proposed in 2005\textsuperscript{1, 2}, active optical clock based on the stimulated emission in a bad cavity has become a very promising optical clock. Compared to the conventional passive optical clock\textsuperscript{3-6}, the advantage of the active optical clock is clear. Representing state of the art, the optical lattice clock based on neutral atoms has shown the stability and accuracy at 10\textsuperscript{-18} level\textsuperscript{3-6}. But the limitation is still the Brownian thermo-mechanical noise of high-fineness Fabry-Perot cavity for frequency stabilization of the oscillation laser\textsuperscript{7, 8}. With a bad cavity in active optical clock, the laser frequency is determined by the atomic transition profile, which is much narrower than the linewidth of the cavity mode. And this results that the frequency of the active optical clock is insensitive to the Brownian thermo-mechanical noise of optical cavity. In addition, the principle of the stimulated emission will make the linewidth of the laser frequency surpass the limitation of quantum-limited linewidth. Millihertz level of the linewidth is expected\textsuperscript{9-16}.

With these prior advantages, the active optical clock also suffers some problems. First, an appropriate atom ensemble which has narrow atomic transition spectroscopy is hard to find to realize the stimulated emission in a bad cavity. Second, the low laser output power is difficult to detect. For now, several experimental schemes have been realized\textsuperscript{17, 18}. In this paper, we demonstrated the active optical frequency standard based on narrow bandwidth Faraday atomic filter, which is called active Faraday optical frequency standard.

Faraday magneto-optical effect was discovered by Michael Faraday in 1845\textsuperscript{19}. Through 170 years development, it has been extended to laser frequency stabilization\textsuperscript{20-25}. Due to its high transmission, high noise rejection and narrow bandwidth, Faraday atomic filter is further utilized as a frequency reference in active optical frequency standard. As shown in Figure 1, the atom ensemble which dips in a homogeneous magnetic field can be vapor cell atoms, atomic beam, neutral cold atoms and trapped ion with narrow transition lines. On both sides of the atomic ensemble, there are two crossing Glan-Taylor prisms to realize the Faraday atomic filter. The gain medium could be chose with Ti:Sapphire, dye and semiconductor. Two mirrors constitute a bad cavity with one mirror having a low reflectivity. The departure of frequency reference and gain medium overcomes the limitation of low laser output power in active Faraday optical frequency standard.

In this system, active Faraday optical frequency standard is based on Cs 852 nm narrow bandwidth Faraday atomic filter in an extended bad-cavity of a semiconductor diode. The narrow bandwidth Faraday atomic filter is realized by using a velocity selection spectroscopy with pumping laser and probing laser in opposite propagation direction. The stimulated emission is demonstrated by using the Faraday atomic filter frequency

Fig. 1: Active optical frequency standard based on narrow bandwidth Faraday atomic filter
reference and the semiconductor gain medium in a bad cavity. The output laser frequency is determined by the Cs $^6S_{1/2} F=4$ to $^6P_{3/2} F'=4, 5$ crossover transition line. The stable active Faraday optical frequency standard has been realized. The preliminary experimental results have been published. Then, we redesign the system in an integrated construction to reduce the influence of the mechanical and thermal vibrations. The whole experimental system with two separated active Faraday optical frequency standard has been built. The transmission of the Faraday atomic filter has achieved to 23% and the bandwidth has achieved 52.6MHz. The length of the extended bad cavity is 35cm and the linewidth of the cavity mode is 345.1MHz with a bad cavity coefficient 6.6. With bad cavity feedback, the active optical frequency standard has realized the $42\mu W$ stable laser output.

II. FARADAY ATOMIC FILTER

The experimental setup of active Faraday optical frequency standard is shown in Figure 1. A specific experimental setup has been shown in reference 18. Based on the system in reference 18, a new experimental structure is designed. The system is integrated for much more steady structure in Figure 2. The 852nm laser diode (Eagleyard EYP-RWE-0860-06010-1500-SOT02-0000) is anti-reflective coating diode and the output is collimated with an aspheric lens. The temperature of the diode is controlled. An 80% reflectivity mirror and the diode constitute an extended bad cavity. The length of the extended bad cavity is 35cm. Between the two Glan-Taylor prisms, Cs vapor cell surrounded homogeneous magnetic field coil is set in magnetic shield box. The length of the Cs vapor cell is 10cm and the diameter is 2.54cm. The cat eye structure is for collimated active optical frequency signal output. The pumping laser and the probing laser are outside the integrated box. The pumping laser is locked to the atomic transition line and power amplification by a light amplifier. The probing laser is used to adjust the narrow bandwidth Faraday atomic filter and is eliminated while active Faraday optical frequency standard works.

With the integrated structure, narrow bandwidth Faraday atomic filter is realized. The basic Cs energy level is shown in Figure 3. The Zeeman energy levels are split in homogeneous magnetic field. The $\sigma^+$ polarized pumping laser pumps the ground state atoms to the $^6P_{3/2} F'=5$ energy level. A $\pi$ polarized probing laser can be separated $\sigma^+$ and $\sigma^-$ polarized probing laser. When $\pi$ polarization probing laser goes through these particular atoms, the transition intensity will be unbalanced between $\sigma^+$ and $\sigma^-$ polarized probing laser. And it will lead to the polarization rotation of the probing laser. With two crossed Glan-Taylor prisms, the Faraday atomic filter signal is observed.

In this paper, a 10 cm long Cs vapor cell is used. The surface transmission of Cs vapor cell is 70%. The homogeneous magnetic field is generated by the magnetic coil surrounded the Cs vapor cell. The inhomogeneity of the magnetic field is tested and only a little inhomogeneity occurs on the edge of Cs vapor cell. The extinction ratio of two crossed Glan-Taylor prisms serving as an frequency analyzer is $1 \times 10^{-5}$. The transmission of a single Glan-Taylor prism is 85%.

![Fig. 2: Integrated active optical frequency standard based on narrow bandwidth Faraday atomic filter](image)

![Fig. 3: Cs energy level](image)

![Fig. 4: Narrow bandwidth Faraday atomic filter](image)
lock the pumping laser to the Cs $^6S_{1/2} F=4$ to $^6P_{3/2} F' = 4, 5$ crossover transition line.

Corresponding to the Cs $^6S_{1/2}$ to $^6P_{3/2}$ saturation spectroscopy, narrow bandwidth Faraday atomic filter signal is calibrated on the Cs $^6S_{1/2} F=4$ to $^6P_{3/2} F' = 4, 5$ crossover transition line. The bandwidth is 52.6 MHz and the transmission reaches to 23%. Figure 6 shows the transmission changing of the Faraday atomic filter with the changing of the homogeneous field and the pumping laser power. The transmission is maximized with 3G magnetic field. The transmission increases with the increasing pumping laser power. Due to the limitation of the pumping laser power, we choose a 10mW pumping laser power and the transmission is 15%. We also increase the temperature of the Cs vapor cell by using a heater band, but the transmission decreases with the increasing of the Cs vapor cell’s temperature. The reason is the absorption of the long Cs vapor cell increases with the increasing of the Cs vapor cell’s temperature. The temperature of the Cs vapor cell maintains 23°C.

In this system, we build two separated narrow bandwidth Faraday atomic filter with the same pumping laser, probing laser, different Cs vapor cell and different experimental structure. One is the reconstructed separated structure which is the same as the reference 18 and the other is the new designed integrated structure. The above experiment results come from the reconstructed separated structure. Similar results in the new designed integrated structure have been achieved. In the new designed integrated structure, the max transmission is 13% and the bandwidth is 42.5MHz with 10mW pumping laser, 190 μW probing laser and 4G optimized homogeneous magnetic field.

III. ACTIVE FARADAY OPTICAL FREQUENCY STANDARD

An anti-reflective coating laser diode with collimated aspheric lens and an 80% reflectivity mirror constitute the extended bad cavity. The length of the extended bad cavity is 35cm. As frequency discrimination, narrow bandwidth Faraday atomic filter is set in the extended bad cavity. Combining with the loss in the cavity (Cs vapor cell transmission 70%, single Glan-Taylor prism transmission 85%, reflectivity mirror 80% and Faraday atomic filter transmission 15%), the linewidth of the cavity mode is expressed as

$$\Gamma_c = -\frac{c\ln\sqrt{R}}{2\pi L}$$

Where c is the light speed, L is the cavity length and R is the reflectivity when the light propagates a round trip. So the linewidth of the cavity mode is 345.1MHz. Corresponding to the 52.6MHz bandwidth Faraday atomic filter, the bad cavity coefficient is 6.6, which means the active Faraday optical frequency standard works in the bad cavity regime.

Adjusting the feedback mirror to realize the optical feedback, the active optical frequency standard signal output is shown in Figure 7. When the pumping laser swept to the Cs $^6S_{1/2} F=4$ to $^6P_{3/2} F' = 4, 5$ crossover transition line, the stimulated emission of active optical frequency standard is realized due to the transmission of narrow bandwidth Faraday atomic filter. The output power of active Faraday optical frequency standard is 42μW. The output laser frequency is determined by the atomic transition line. Because the gain medium is the diode laser, the output laser power is
not limited by the finite atoms. It makes the observation of the active optical frequency standard signal much easier.

In reference 18, the linewidth of Faraday optical frequency standard has been measured with two separated active Faraday optical frequency standard. In order to decrease the influence of the environmental vibration, we redesign an integrated system. For now, the new compact system has not realized the active Faraday optical frequency signal output because of the relative low transmission of the narrow bandwidth Faraday atomic filter and the linewidth can’t be measured for better results.

IV. CONCLUSION

In conclusion, we realized a new kind of active optical frequency standard based on narrow bandwidth Faraday atomic filter. With an extended bad cavity, the output frequency standard is determined by narrow bandwidth Faraday atomic filter, corresponding to the Cs crossover transition line 6^2S_{1/2} F=4 to 6^2P_{3/2} F'=4, 5. The output laser power can reach to 42 μW, which is easy to detect. In this paper, the realized Faraday atomic filter is with 23% transmission and 52.6MHz bandwidth. Considering the loss in the extended cavity, the linewidth of the cavity is 345.1MHz. The bad cavity coefficient is 6.6, which means the active Faraday optical frequency standard works in the bad cavity regime. The influence of the extended cavity is decreased.

With the proof-of-principle in active Faraday atomic filter, much narrower atomic ensembles and larger bad cavity coefficient will be explored. Possessing the great advantages of stimulated emission and cavity vibration immunity, millihertz linewidth of active optical frequency standard is expected.

REFERENCES


All-fiber Implementation of Modulation Transfer Spectroscopy for $^4$He Atoms

Wei Gong, Xiang Peng*, Wenhao Li, Teng Wu, Haidong Wang, Jingbiao Chen and Hong Guo†
State Key Laboratory of Advanced Optical Communication Systems and Networks,
School of Electronics Engineering and Computer Science, and Center for Quantum Information Technology,
Peking University, Beijing 100871, China.
Email: *xiangpeng@pku.edu.cn, †hongguo@pku.edu.cn

Abstract—The demonstration of modulation transfer spectroscopy (MTS) for $^4$He atoms with an all-fiber experimental scheme is presented, which exhibits more flexibility than the free-space one. MTS signals at $^4$He transition lines around 1083 nm are obtained and optimized. The application in laser frequency stabilization is also discussed.

Keywords—modulation transfer spectroscopy, metastable helium, frequency stabilization.

I. INTRODUCTION

Modulation transfer spectroscopy (MTS) is an optical heterodyne spectroscopy which combines both radio-frequency external modulation and phase-sensitive detection techniques. It is achieved by heterodyning the optical carrier with the modulated sidebands [1], which is efficient and leads to smaller technical noise [2]. In MTS, the modulation of the probe beam is transferred from the phase modulated pump beam under the four-wave mixing process which occurs only when the sub-Doppler resonance condition is satisfied [3], [4], [5], [6]. Therefore, the MTS signal has relatively high signal-to-noise ratio and has been demonstrated with various elements in precise metrology applications since 1980s [6], [7], [8], [9]. In our previous work [10], MTS of $^4$He atoms is observed with a free-space experimental scheme and furthermore, it is applied for locking a fiber laser to the transition lines of metastable $^4$He atoms around 1083 nm.

Since the metastable $^4$He atom has very simple structure, it is increasingly used in fields like laser cooling, optical pumping and optical magnetometry [11], where frequency stabilization of the laser source around 1083 nm is necessary. Compared with the laser frequency stabilization techniques with Saturated Absorption Spectroscopy (SAS) [12], Polarization Spectroscopy (PS) [13] and Frequency Modulation Spectroscopy (FMS), MTS exhibits superior locking performance [10].

In this paper, we present the demonstration of MTS for $^4$He atoms with an all-fiber experimental scheme, which exhibits more flexibility than the free-space one. MTS signals at $^4$He transition lines around 1083 nm are obtained and optimized. The application in laser frequency stabilization is discussed and the locking performance is evaluated.

This work is supported by the National Science Fund for Distinguished Young Scholars of China (Grant No. 61225003), the National Natural Science Foundation of China (Grant No. 61101081), and the National Hi-Tech Research and Development (863) Program.

II. EXPERIMENTAL SCHEMES

Based on the principle and free-space realization of MTS in [10], we design the all-fiber MTS experimental scheme for $^4$He atoms, which is illustrated in Fig. 1. All the optical elements in this scheme are fiber-connected and polarization maintaining ones.

The linear polarized laser beam from the 1083 nm fiber laser (with optical power of about 30 mW) enters a 50:50 fiber coupler and splits into two—the probe and the pump beams. A fiber LiNbO$_3$ electro-optic modulator (EOM) is used to modulate the phase of the pump beam, which is driven by a radio-frequency oscillator. The oscillator also produces synchronous signal as the reference for the following phase-sensitive detection. The optical circulator is used to determine the propagating directions of both the pump and the probe beams. The probe and the modulated pump beams are aligned collinearly within a $^4$He atomic vapor cell which is fixed on a U-bench fiber-to-fiber collimator. $^4$He atoms in the vapor cell are excited to the metastable state by a radio-frequency cell (0.3 Torr at room temperature) is about 500 mm long, which is just as long as the U-Bench so that relatively more atoms participate in the interaction process. With the four-wave-mixing effect in the vapor cell, the modulation of the pump beam can be transferred to the originally unmodulated probe beam.

Finally the probe beam with sidebands is heterodyne detected with an InGaAs photodetector (PD). By utilizing standard phase-sensitive detection with a lock-in amplifier, the dispersive-like MTS signal can be generated. MTS signal could be optimized with beam power adjustment through the...
two in-line optical power attenuators. After optimization, the dispersive-like MTS signal from the lock-in amplifier could be further used as the feedback error signal to control the laser frequency.

It should be noted that the polarization of the pump and probe beams would influence the MTS signal [14], therefore the MTS signal could be improved by manipulating the polarization of the two beams. However, all the optical elements are polarization maintaining ones, which means the polarization of the propagating beams are settled in the all-fiber experimental scheme. The only way to alter the polarization of the beams is to insert wave plates inside the U-Bench collimator at the cost of the vapor cell length, which might deteriorate the MTS signal in return. So polarization manipulation of the beams is temporarily not taken into consideration in our experiment.

III. RESULTS AND DISCUSSIONS

With the experimental scheme in Fig. 1, MTS signals around 1083 nm at D\textsubscript{0}, D\textsubscript{1}, D\textsubscript{2} transition lines and the cross line between D\textsubscript{1} and D\textsubscript{2} are obtained respectively, which is displayed in Fig. 2. It is obvious that the MTS signals exhibit sharp slopes around resonance and sit on zero baseline, which is particularly suitable for laser frequency locking.

Actually, the frequency of the fiber laser can be locked to each of the transition lines around 1083 nm. For simplicity, the optimization of the MTS signal at D\textsubscript{1} line is chosen as an example to clarify. As mentioned before, the two optical power attenuators on the output arms of the fiber coupler in Fig. 1 are used to adjust the optical power of the counterpropagating pump and probe beams respectively. According to the experience from our previous work [10], pump beam with over ten times higher optical power than the probe beam will lead to a MTS signal with relative large amplitude and sharp slope near resonance. To ensure the signal-to-noise ratio of the detection, the probe beam power should not be too small as well. Aside from the beams power, the modulation frequency and the demodulation phase also influence the amplitude and the slope of the dispersive-like MTS signal. After numerous experiments, the relatively optimal signal is achieved with the pump beam power of 10 mW and probe beam 2 mW, which is shown in Fig. 3. The optimized signal has a peak-to-peak amplitude of 2.1 V and linewidth of 25 MHz.

Following, we use the optimized MTS signal from Fig. 3 as the feedback error signal to control the frequency of the 1083 nm fiber laser. Combined with the servo circuit, the laser frequency could be locked to the D\textsubscript{1} transition line of \textsuperscript{4}He atoms easily. To evaluate the frequency stabilization performance, we calibrated the error signal to the laser frequency fluctuation with a Fabry-Pérot interferometer [15]. Fluctuation signals under both the free running and the frequency locked cases are recorded for 1000 s duration with a data acquisition card (sampling rate 100 Hz). Fig. 5 shows the frequency locking performance clearly: for the open-loop free running case, the center frequency of the fiber laser drifts with a peak-to-peak value up to 5 MHz, and for the close-loop stabilized case, the frequency drift disappears and the fluctuation is suppressed to 10 kHz order of magnitude (root-mean-square).

The experimental results show that the all-fiber MTS scheme for \textsuperscript{4}He atoms is practicable and compared with the free-space scheme in our previous work [10], the frequency...
locking performance of the 1083 nm fiber laser to the transition lines of metastable helium atoms is equally outstanding. What is more, the all-fiber scheme exhibits obvious flexibility than the free-space one and is more promising in miniaturization for future use. In addition, as mentioned above, the scheme could be further improved with introducing reasonable beam polarization manipulation mechanism and methods.

IV. CONCLUSION

In this paper, we demonstrate the all-fiber implementation of MTS for $^4$He atoms. With replacing key elements and devices into corresponding fiber ones and introducing necessary mechanism, the MTS signal of the $^4$He atoms around 1083 nm is obtained. We optimize the dispersive-like MTS signal by experimentally studying the influence of the beam power, modulation frequency and demodulation phase. With using the optimized MTS signal as the feedback error signal, we lock the fiber laser frequency to the transition lines of $^4$He atoms. The frequency locking performance of the all-fiber MTS scheme shows it is as useful as the free-space scheme in laser frequency stabilization application. And with its flexibility and room for improvement, the all-fiber MTS scheme is much more promising for the future use.

ACKNOWLEDGMENT

We thank Mr. Zaisheng Lin for his help in data acquisition and processing.

REFERENCES

Large Waist Cavity for Ultra-narrow Transition Spectroscopy

Stefan A. Schäffer, Sigrid S. Adsersen, Bjarke T. R. Christensen, Jan W. Thomsen
Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, 2100 Copenhagen, Denmark
Email: schafer@nbi.dk

Abstract—Cavity enhanced spectroscopy on ultra-narrow transitions in atoms and molecules is a promising tool for frequency stabilization of clock lasers. However, the small beam waist of a typical cavity can effectively reduce the number of atoms in the system or introduce transit time broadening in the less controlled molecular systems. We present a cavity design with an internal telescope optimized to increase the waist from about 0.5 mm to 5 mm while still maintaining a convenient range of tuning parameters. Its usefulness in three separate spectroscopic systems is clarified.

I. INTRODUCTION

State-of-the-art optical atomic and molecular clocks employ spectroscopy on ultra-narrow atomic transitions. This exerts high demands on the spectral quality and stability of the interrogation laser. In order to obtain a high spectral purity, extensive setups are necessary to sufficiently stabilize the reference cavity of the laser. Rather than using an empty cavity as a frequency discriminator, spectroscopy directly on atomic or molecular transitions can provide a sufficiently sensitive phase response to use as a feedback signal for locking the laser frequency. Using a cavity to enhance the light-matter interaction, one can greatly improve the locking signal of the spectroscopic system [1]–[4], thus allowing for a large signal-to-noise ratio without the high demands on cooling and shielding of a bare reference cavity [5]. Such a system, however, will typically have a small intracavity beam waist resulting in a restriction on the number of atoms or molecules and a spectral broadening of the resonance feature due to transit time broadening in molecular samples.

Here we present a cavity designed with an internal telescope in order to increase the cavity waist radius of the beam while maintaining a robust construction. We have investigated the optical limitations of a variety of different focal configurations, and present a characterization of a cavity design with a large TEM₀₀ mode (see Fig. 1). The system finesse of the cavity is in the few hundreds and limited by the quality of the optical coating as well as spherical aberrations.

In a typical large-waist cavity, end mirrors with a radius of curvature (ROC) of 9 m provides a waist radius of the order of 0.5 mm. By using an intracavity telescope a beam diameter of the order of centimeters can be achieved in the optical domain. We are currently developing a system for use in a molecular clock setup for iodine (I₂) or acetylene (ethyne) as well as an atomic ⁸⁸Sr beamline. For these experiments waist radii of 3.5 to 6 mm will be implemented. This is an increase in cavity waist size by an order of magnitude compared to two-mirror cavities while still maintaining a cavity that is well stabilized within an experimentally feasible range of tuning parameters.

II. TARGET SYSTEMS

When working with an atomic or molecular sample in a cavity the coupling of one atom or molecule to another through the intracavity field can be characterized by the collective cooperativity (C) of the system [3]. As the atom number N increases, the collective cooperativity scales linearly: \( C = \frac{4N\gamma^2}{\pi\kappa} \). Here \( g \) is the single atom-cavity coupling constant, \( \kappa \) is the cavity decay rate, \( \Gamma \) the transition linewidth of the molecule or atom, and \( C \) is unitless. Frequency stability of the interrogation laser in current atomic clocks is limited due to Brownian motion in the reference cavity mirrors [6]. Using a system in the bad cavity regime of \( \Gamma \ll \kappa \), the cavity mirrors will no longer be the limiting factor [1]. Instead the narrow linewidth \( \Gamma \) of the atoms is exploited. The cavity finesse serves to increase the effective interaction length as well as the coupling between single atoms or molecules. In this regime a high cooperativity \( C \) is obtained by having a large number of atoms or molecules rather than a high finesse. Even deep in the bad cavity regime, where the finesse is in the hundreds, this may provide laser frequency linewidths of few mHz [1], [3].

One of several planned experiments to take advantage of a large cavity mode to increase cooperativity is a ⁸⁸Sr beamline. In such a setup a jet of slow atoms traveling perpendicular to the probing mode allows for continuous direct interrogation of the ¹S₀ – ³P₁ clock transition at \( \lambda = 689 \) nm and \( \Gamma = 2\pi \times 7.6 \) kHz. The velocity of the atoms moving through the cavity mode is limited in order to reduce transit time broadening. Lower atom-velocities, however, imposes higher demands on the transverse velocity and thus results in loss of atoms. The large waist in an optical cavity such as the one proposed here, allows probing of faster atoms, effectively increasing the number of atoms interrogated in the system. A higher atom number boosts our collective cooperativity and thus facilitates a better error signal, reducing the local oscillator instability [1]. The finesse of the cavity (\( \propto \frac{1}{\alpha} \)) influences collective cooperativity in the same way as \( N \), and though we remain in the bad cavity regime it should not be neglected.

Other approaches will focus on molecular spectroscopy on gas cells, and here the sample is typically at or close to room temperature. This means that a large waist diameter greatly
reduces the transit time broadening of the hot molecules. This broadening occurs as a fast atom or molecule passes through the interrogation mode. The time spent inside the beam sets a lower Fourier limit to the spectral linewidth. The broadening is given by [7]:

\[
\Delta \nu_{tt} = \frac{1}{2\pi} \sqrt{\frac{2 \ln 2 \frac{\pi k_B T}{2M}}{w(z)}},
\]

where \( T \) is the sample temperature, \( M \) is the molecular or atomic weight of the sample, and \( k_B \) the Boltzmann constant. Here the proportionality to the inverse of the position-dependent waist radius \( w(z) \) is of special interest to us. We interrogate the \(^{127}\text{I}_2\) P(13) 43-0;\( \alpha_2 \) transition at \( \lambda = 514.6 \) nm in iodine with a transition width of \( \Gamma = 2\pi \times 284 \) kHz [8], and the \(^{13}\text{C}_2\text{H}_2\) P(16);\( \nu_1 + \nu_2 \) transition at \( \lambda = 1542.4 \) nm in acetylene which has an expected transition width of a few Hz (estimated from similar transition in [9]). These simple systems are very promising for compact clock systems [10]–[12] relevant e.g. for space or telecom applications. The use of a cavity, rather than simply counter-propagating beams, will effectively increase the interrogation length by a factor of the order of the finesse. When combined with the great reduction in transit time broadening due to a large-waist cavity, this system could exceed current state-of-the-art molecular clock systems in terms of frequency stability [12].

III. SYSTEM SETUP

In order to obtain a large waist inside the cavity in which the sample is placed, measures must be taken to widen the beam beyond typical waist radii. As seen in Fig. 1, the cavity consists of four different optical elements. An in-coupling mirror (\( M_1 \)) with a focal length of \( f_{M_1} = 500 \) mm followed by a telescope mirror (\( M_2 \)) with \( f_{M_2} = 12 \) mm, a telescope lens (\( M_3 \)) with \( f_{M_3} = 200 \) mm, and a flat out-coupling mirror (\( M_4 \)). The sample cell containing the sample is placed between \( M_3 \) and \( M_4 \). The cell should have Brewster cut windows in order to minimize reflection loss at the surfaces. While we only show a folded cavity in Fig. 1, a linear version of the same setup was also investigated. Here the mirror \( M_2 \) was replaced by a lens with the same focal properties.

Typically, symmetrical intracavity telescopes are employed to obtain a sharp focal point where intensity in some nonlinear material or gain-medium is high. In this setup, however, an asymmetric telescope is essential as we wish to enlarge the beam waist. As we are not interested in the small intr telescope waist a convex \( f = -12 \) mm could be used at \( M_2 \) rather than the concave mirror shown in Fig. 1. This slightly increases the compactness of the setup, but could be more sensitive to alignment errors. The focal properties of the end mirrors \( M_1 \) and \( M_4 \) are available off-the-shelf optics, chosen for convenience. Using ROC = 9 m \( f = 4.5 \) m at both positions will increase the waist radius by about 7% for distances cited below and could reduce alignment instabilities at \( M_4 \).

IV. CHARACTERIZATION OF CAVITY PARAMETERS

When designing an optical cavity, demands on size, accessibility and mechanical stability sets a number of limitations on optical elements and feasible parameters. If the cavity dimensions are too large, instabilities due to vibrations become important, and semi-compact applications are no longer possible. Correspondingly the size of the mirrors and lenses set limits on the beam size, as a significant reduction in the finesse occurs when the Gaussian tail is lost outside optical components.

In a cavity of the type presented here with an internal beam expander the mode is divided into three sections. In the third section \( d_3 \), see Fig. 1) the mode is wide, and the wavefront curvature is insignificant \( R_{min} = 3 \times 10^7 \) m. The wavefront curvature introduces broadening effects according to [7]:

\[
\Delta \nu_{wf} = \Delta \nu_{tt} \left( \frac{\pi w(z)^2}{R(z)\lambda} \right)^2,
\]

hence a large curvature radius \( R(z) \) of the wavefront is desirable to minimize such broadening. The section \( d_3 \) is also the section where spectroscopy is performed, and it can be trimmed to fit the experiment at hand as it is conveniently insensitive to changes in length. The telescope section \( d_2 \) is very sensitive to changes in length, and is thus entirely determined by the focal lengths of \( M_2 \) and \( M_4 \). In Fig. 2 the waist size in \( d_3 \) can be seen as a function of distances \( d_1 \) and \( d_2 \) and it is seen that the distance \( d_1 \) can be used to adjust the spectroscopic waist described by:

\[
w^2 = \frac{\lambda}{\pi} \frac{2B}{\sqrt{4 - (A + D)^2}},
\]

Here the stability criterion is given as \(-1 < S < 1 \) and \( S = \frac{\lambda + 2B}{2} \). A, B, and D are given by the transfer matrix of the cavity and \( \lambda \) is the wavelength of the light. All measurements presented here were performed at \( \lambda = 532 \) nm.

The stable region of the cavity maps out a valley in the parameter \( d_2 \). The width of this valley diverges as \( d_1 \rightarrow 0 \).
we have chosen the waist size of the system as a function of inter-mirror/lens distances. Here we have chosen \( f_M = 500 \text{ mm}, \ f_M = 12 \text{ mm}, \ f_M = 200 \text{ mm}, \) and a flat mirror at \( M_4 \). As the beam is almost collimated over the length of the sample cell, waist size and stability is insensitive to the value of \( d_2 \). The waist radius is shown as a function of distances \( d \) and \( d_2 \), and a white circle marks the parameters for the measurements presented in Fig. 4. Eq. (3) was used to express the waist size, and the wavelength used was \( \lambda = 532 \text{ nm} \). The hatched area marks unstable cavity parameters where the waist value becomes imaginary.

The waist size diverges as the cavity becomes unstable, and has imaginary values in the hatched, unstable region of Fig. 2. To obtain a cavity with high stability we consider only values of \( d_2 \) in the center of the stability interval, providing a minimal waist size but also a minimal change in waist size \( \frac{\Delta w}{d_2} \) with respect to the most sensitive parameter. Increasing the distance \( d_1 \) will also decrease the width of the stable region in \( d_2 \). In Fig. 2 a white circle marks the position in parameter-space for the cavity distances presented here. The telescope focal lengths were chosen to have a relation of approximately 1:16 as larger magnification becomes impractical in terms of both focal lengths and physical sizes when off-the-shelf optical elements are considered. As can be seen from the figure, much larger mode volumes are possible even with these focal lengths, but at the cost of increasing the cavity length. The bent cavity allows for a large parameter space in \( d_1 \), depending on sample cell construction, without significantly increasing the physical footprint. A three-element linear cavity with no mirror \( M_1 \) was also investigated, but found to yield unsatisfactory waist sizes.

A. Errors and deviations

Deviations due to physical rather than ideal optical elements put restrictions on our parameter space and adds additional conditions to cavity stability. The most important errors in the system described here are spherical aberrations from mirror and lens geometries, as well as astigmatism introduced when elements are tilted and the cylindrical symmetry of the setup is broken. Whereas errors due to astigmatism can be directly calculated using the transfer-matrix of Gaussian beams separately for the two orthogonal dimensions in the beam cross-section, spherical aberrations due to lens errors cannot be as easily implemented. Ideally all mirrors and lenses should be designed aberration-free for the setup at hand, but such a design is timeconsuming and costly, and by simple considerations the largest errors can be avoided. Spherical lenses and mirrors are typically cheaper than hyperbolic lenses and parabolic mirrors thus making them preferable if no significant losses occur. In order to reduce the physical footprint of the cavity as well as removing excess surfaces which can reduce the finesse of the cavity, it is of interest to fold it by replacing telescopic lenses with mirrors.

1) Spherical aberrations: In the ray picture of light the spherical aberration of a lens can be described by the dependence of the focal length \( f_{sph}(r) \) on the radial distance \( r \), of the incoming ray from the lens center. This dependence results in a smearing of the focal point which can be given by a non-zero spot diameter [13]:

\[
\Delta x \propto \frac{r^3}{f_p},
\]

where \( \Delta x \) is the spot diameter, \( r \) is the distance of a parallel incident ray to the lens center and \( f_p \) is the paraxial focal length of the lens. Though this is a description in the ray picture, its simplicity gives a useful guideline for the resonator considered here. The beam shape of the resonator can be directly calculated for ideal lenses using the transfer matrix approach, and a qualitative overview can be seen in Fig. 1. From this we see, as expected, that the beam waist will be much larger at \( M_3 \) and \( M_4 \) than at \( M_1 \) and \( M_2 \). Since we place ourselves in the center of the stable \( d_2 \) region the beam radius will be close to 16 times larger at \( M_3 \) compared to \( M_2 \) due to the equal but inverse relationship between the focal lengths. From (4) it follows that \( M_3 \) will be more sensitive to spherical aberrations than \( M_2 \). This means that efforts should be concentrated on reducing spherical aberrations here, e.g. by using a hyperbolic rather than a spherical lens. In the measurements presented below, a plano-convex rather than a bi-convex lens was used to minimize loss due to spherical aberrations.

2) Astigmatism: Astigmatism can be calculated directly assuming ideal lenses and splitting the transfer matrices into two orthogonal planes. In Fig. 3 the resonator stability parameter has been plotted as a function of the angle of incidence (AOI) of \( M_2 \) and \( M_3 \) respectively. In both cases all other AOIs were assumed to be zero. Once more the large beam diameter at \( M_3 \) results in a much more sensitive parameter. Already at an AOI of \( 4^\circ \) the cavity becomes unstable, whereas an AOI of \( 16^\circ \) is permissible in the case of \( M_2 \). This means that bending the cavity at \( M_3 \) is not experimentally feasible, due to the waist size at \( d_3 \). However, a bend at \( M_2 \) is possible and an AOI as low as \( 6^\circ \) to \( 8^\circ \) should be permissible without significantly increasing losses from other aberrations. The bend will result in a slightly elliptical beam at \( d_3 \), which should not affect performance. It also reduces the number of scattering surfaces and results in a more compact cavity.
Fig. 3. The stability of the cavity is sensitive to the angles of the mirrors and lenses. In order to reduce the number of surfaces as well as increase the compactness of the setup, a bent cavity is used rather than a linear one. As the stability is very sensitive to the angle of M3, there is a bend only at the position of M2. Here the stability parameter is shown as a function of angle of incidence (AOI) for the two lenses in both the sagittal (S) and tangential (T) planes.

B. Finesse

The cavity finesse determines the effective interaction length of the photons with the sample cell. The average number of times a photon is reflected inside the cavity is thus given as $N = \frac{2}{\pi F}$, where $F$ is the cavity finesse. All loss inside the cavity due to either out-coupling or absorption will reduce the finesse. For a surface reflectance of $R > 0.5$ the finesse can be approximated as:

$$F \approx \frac{\pi \sqrt{R^n}}{1 - R^n},$$

(5)

where $R^n$ is the product of the intensity reflectance of all $n$ surfaces assuming identical coatings. When lenses are used in the cavity rather than mirrors the transmission coefficient $T$ of each lens surface enters in the above equation on an equal footing to $R$, however, a disadvantage of using lenses is the increase in the number of surfaces encountered by the photons. A lens also adds absorption in the lens material which is expressed as $1 - A = e^{-\alpha z}$, where $\alpha$ is the absorption coefficient of the material and $z$ the thickness of the lens. Once more $1 - A$ enters in Eq. (5) as would $R$. Some of the best commercially available Anti-Reflective (AR) and reflective coatings have guaranteed transmission and reflection coefficients of $T \geq 0.998$. The absorption in a $z = 6.2$ mm BK7 lens at 532 nm is $1 - A = 0.999$. In Table I an overview of the achievable finesse with five (the folded cavity presented here) and six (a linear version using a lens at M2) surfaces in the cavity is given.

V. MEASUREMENT OF PARAMETERS

Measurements of the cavity finesse using a series of different lenses in a linear cavity as well as a bent cavity yielded finesses comparable to theoretical values. By using achromatic and hyperbolic lenses to correct for spherical aberrations at M2 measurements in agreement with Eq. (4) were obtained. Optimal results with available mirrors and lenses were obtained for the bent cavity and a plano-convex lens at M3. Here we used end mirrors M1 and M4 with $R = 0.998$ and $R = 0.995$ respectively, a single surface (mirror) at M2 with $R = 0.999$, and two surfaces (plano-convex lens) at M3 with $T = 0.995$. For lenses with no aberrations, but including material absorption, this gives a theoretical finesse of $F = 157$, whereas the measured value was $F_{\text{meas}} = 139$, presumably due to spherical aberrations from M3. Using a hyperbolic lens at M3 should reduce the loss and boost the finesse to approach the theoretical value.

The waist size was measured by fitting a Gaussian curve to the pixel intensity of a photograph as shown in Fig. 4. Here the known mirror size is used for scaling. The measured waist radius is $w_0 = 3.44 \pm 0.06$ mm which should be compared to the theoretical value of $w_0 = 3.38$ mm.

VI. CONCLUSION

In the case of iodine or acetylene, a large waist cavity will not boost the effective number of molecules, as available velocities are tuned only by the temperature of the cell. Instead the large waist will directly reduce some of the limiting broadening effects in the system: Transit time broadening, see Eq.
and wavefront curvature broadening, see Eq. (2). A waist size of \( w = 3.4 \) mm with an iodine gas at room temperature results in a transit time broadening of \( \Delta \nu_{tt} = 6.8 \) kHz, which is more than an order of magnitude smaller than the natural linewidth \( \Gamma \). With a two-mirror cavity the broadening would be of the order of \( \Gamma \). Using a finesse of \( \mathcal{F} = 139 \) and a sample cell of \( l = 20 \) cm we have an effective saturated absorption cell length of \( l^2_{\mathcal{F}} = 8.9 \) m compared to \( 1.8 \) m in [12]. In acetylene the waist radius increases to \( w = 5.8 \) mm due to the increase in wavelength and the broadening will be reduced to \( \Delta \nu_{tt} = 12 \) kHz at room temperature. This is an improvement of an order of magnitude compared to the \( \Delta \nu_{tt} = 135 \) kHz achieved by [10], but still much larger than the expected natural linewidth \( \Gamma \) of the transition. The total linewidth of the system will thus still be limited by the transition time broadening. The wavefront curvature broadening is practically removed for all optical wavelengths (\(<\) mHz) allowing a larger interrogation volume without the increased broadening of traditional cavities.

In a \(^{88}\)Sr sample the gas would typically be generated by an oven and cooled using a Zeeman slower. If a beamline is used the slowest atoms are lost due to gravitational effects as well as broadening of the beam inside the slower, and increasing the velocity of the atoms thus reduces loss. With the longer wavelength of the transition in \(^{88}\)Sr (689 nm) the waist becomes \( w = 3.8 \) mm and a speed of \( v = 74 \) m/s gives a broadening equal to the transition width \( \Delta \nu_{tt} = \frac{\Gamma}{2\pi} \). With a 0.5 mm waist radius this velocity would be restricted to \( v = 10 \) m/s. The higher mean atomic velocity reduces loss and thus effectively boosts the number of atoms \( N \).

A cavity as the one presented here can thus boost the effective atom number or significantly reduce the wavefront curvature and transit time broadening of atomic and molecular spectroscopic systems. Both effects are of interest to spectroscopy on narrow linewidths and the technique could thus be implemented in order to increase the sensitivity of such systems, and transgress current technical limitations on optical frequency stabilization.

ACKNOWLEDGMENT

The authors would like to acknowledge support from the Danish Research Counsil and ESA Contract No. 4000108303/13/NL/PA-NPI272-2012.

REFERENCES

Resonant Infrared Detector Based on a Piezoelectric Fishnet Metasurface

Yu Hui, Zhenyun Qian, and Matteo Rinaldi
Department of Electrical and Computer Engineering
Northeastern University
Boston, MA, USA
yhui@ece.neu.edu; rinaldi@ece.neu.edu

Abstract—This paper reports on the first demonstration of a high resolution (noise equivalent power NEP of 1.9 nW/Hz$^{1/2}$ at 200 Hz bandwidth) and fast (thermal time constant of 5.3 ms) infrared (IR) detector based on a nanoelectromechanical system (NEMS) resonant piezoelectric fishnet-like metasurface (PFM). For the first time, an ultrathin (650 nm) piezoelectric fishnet-like metasurface is employed to form the vibrating body of a nanomechanical resonator with a unique combination of optical, thermal and electromechanical properties. Efficient sensing and actuation (electromechanical coupling coefficient, $k_t^2$~1.4%) of a high frequency (172 MHz) and high quality factor ($Q$~1254) bulk acoustic mode of vibration in the free-standing ultrathin structure is achieved thanks to the superior piezoelectric transduction properties of the proposed metasurface. Strong absorption (60%) of short wavelength infrared (SWIR) radiation in the ultra-low volume resonant device is obtained thanks to the properly engineered optical properties of the fishnet-like metasurface which provide a Fabry-Perot like resonance at ~4 μm to the structure.

Keywords—Infrared detector; plasmonic metasurface; aluminum nitride; piezoelectric resonator; NEMS.

I. INTRODUCTION

A micro/nanoelectromechanical system (MEMS/NEMS) resonant infrared (IR) detector is a particular class of thermal detectors that relies on a transduction scheme based on the change in vibration frequency of a microelectromechanical resonator with a temperature dependent mechanical resonance frequency. The advantage of MEMS resonant IR detectors over other technologies is the unique combination of high sensitivity and low noise performance [1]. Specifically, piezoelectric MEMS resonant IR detectors have recently been demonstrated using quartz [2], gallium nitride [3] and aluminum nitride resonators [4], showing promising performance. Nevertheless, a fundamental challenge associated to the implementation of high performance MEMS resonant IR detectors is the integration of efficient IR absorbing materials that are also compatible with conventional transduction and microfabrication techniques.

In this paper, a stepping stone towards the implementation of high performance MEMS resonant IR detectors is set by demonstrating an ultrathin piezoelectric fishnet-like metasurface (PFM). For the first time, an ultrathin (650 nm) PFM is employed to form the vibrating body of a nanomechanical resonator with a unique combination of optical, thermal and electromechanical properties: (1) Efficient sensing and actuation (electromechanical coupling coefficient, $k_t^2$~1.4%) of a high frequency (172 MHz) and high quality factor ($Q$~1254) bulk acoustic mode of vibration in the free-standing ultrathin structure is achieved thanks to the intrinsic piezoelectric transduction properties of the proposed metasurface (based on a thin-film Aluminum Nitride, AlN). (2) Strong absorption (60%) of short wavelength infrared (SWIR) radiation in the ultra-low volume resonant device (without the need of any additional IR absorbing material) is obtained thanks to the properly engineered optical properties of the fishnet-like metasurface which provides a Fabry-Perot like resonance at ~4 μm to the structure. (3) Maximum thermal isolation of the resonant metasurface from the heat sink (thermal resistance ~$2.3\times10^5$ K/W, one order of magnitude higher than previous demonstrations), hence maximum device responsivity (~530 Hz/μW), are attained by using nanoscale metallic anchors (with minimum cross-section) to support the freestanding vibrating body of the structure. Thanks to such unique features, the fabricated PFM IR detector shows a 7× enhanced responsivity and unchanged noise performance compared to a conventional AlN nano-plate resonant thermal detector. The demonstrated properties of the proposed PFM address the most important and fundamental challenges associated with the development of performing resonant IR detectors, marking a milestone towards the implementation of a new class of high performance, miniaturized and low power IR spectroscopy and multi-spectral imaging systems.

II. EXPERIMENTAL RESULTS

The core of the proposed NEMS resonant IR detector is an AlN piezoelectric nano plate (500 nm thick) sandwiched between a 100 nm thick platinum (Pt) bottom inter-digital transducer (IDT) and a top electrically floating layer (50 nm thick gold, Au) patterned with the goal of confining the electric field induced by the bottom IDT across the piezoelectric nano plate, and, at the same time, enabling absorption of IR radiation in the ultrathin piezoelectric nano plate thanks to suitably tailored plasmonic resonances (Figure 1). The proper patterning of plasmonic nanostructures in the top metal layer of the device significantly enhances field concentration which provide a Fabry-Perot like resonance at ~4 μm, enabling strong...
absorption of SWIR radiation over the ultrathin structure (inset in Figure 3). The resonant body of the device is supported by two metallic anchors (with minimum cross-section) to maximize the thermal isolation of the device from the substrate (heat sink), thus maximizing the thermal resistance of the IR detector (Figure 1).

The electromechanical performance of the fabricated AlN PFM IR detector was characterized by measuring its admittance amplitude versus frequency using an Agilent E5071C network analyzer. The measured resonance frequency was found to be 172 MHz (Figure 2). High electromechanical performance (mechanical quality factor, $Q$, of 2254 and electromechanical coupling coefficient, $k^2$, of 1.42%), comparable to the one of conventional ultrathin film AlN MEMS resonators [1], was achieved, indicating that the use of an ultrathin piezoelectric fishnet metasurface to form the resonant body of the device does not deteriorate the electromechanical performance of the resonator (unlike the integration of conventional IR absorbers [3]). The temperature coefficient of frequency (TCF) of the device was measured using a temperature controlled RF probe-station and found to be $\sim 3.84 \text{ kHz/K}$ ($\sim 22.3 \text{ ppm/K}$), which is comparable to the typical values recorded for conventional 500 nm thick AlN contour-mode resonators [1].

The thermal properties of the AlN PFM IR detector were estimated by 3-dimentionsal finite element method (FEM) simulation using COMSOL multiphysics. The device thermal resistance was simulated by applying a thermal power of 1 $\mu$W to the metasurface and monitoring the temperature change. A temperature increase of 230 mK was recorded, resulting in a thermal resistance of $2.3 \times 10^5$ K/W. Such high thermal resistance is at least one order of magnitude higher than previous demonstrations [2], thanks to the maximized thermal isolation of the resonant metasurface from the heat sink by using nanoscale metallic anchors (with minimum cross-section) to support the freestanding vibrating body of the structure. The device thermal time constant was simulated by monitoring the transient response of the device temperature (inset in Figure 2), upon an applied thermal power of 1 $\mu$W. A thermal time constant of 5.3 ms was recorded. The achieved fast device response was due to the ultra-low volume of the resonant structure, thus low thermal mass attributed to the nano plate AlN piezoelectric layer.

The reflection spectrum, $R$, of the fabricated AlN PFM was measured using a Bruker V70 Fourier transform infrared (FTIR) spectrometer and Hyperion 1000 microscope. The spectral absorptance, $A$, was estimated as $A=1-R$, assuming no transmission through the resonant structure (which is a reasonable assumption given the large metal coverage of both the top Au fishnet structures and bottom Pt IDT). The measured result shown in the inset of Figure 3 shows that a strong absorption of SWIR (3-5 $\mu$m) radiation higher than 60% was readily achieved. The device responsivity was calculated by multiplying the measured TCF, absorption $A$, by the simulated $R_{th}$ and it was found to be $\sim 530 \text{ Hz/} \mu\text{W}$.

Figure 2. Measured admittance amplitude versus frequency and modified Butterworth van dye (MBVD) fitting of the NEMS resonator showing high mechanical quality factor ($Q=2254$ and electromechanical coupling coefficient ($k^2=1.42\%$). The inset shows the 3D FEM simulated thermal resistance, time constant and temperature profile.

The reflection spectrum, $R$, of the fabricated AlN PFM was measured using a Bruker V70 Fourier transform infrared (FTIR) spectrometer and Hyperion 1000 microscope. The spectral absorptance, $A$, was estimated as $A=1-R$, assuming no transmission through the resonant structure (which is a reasonable assumption given the large metal coverage of both the top Au fishnet structures and bottom Pt IDT). The measured result shown in the inset of Figure 3 shows that a strong absorption of SWIR (3-5 $\mu$m) radiation higher than 60% was readily achieved. The device responsivity was calculated by multiplying the measured TCF, absorption $A$, by the simulated $R_{th}$ and it was found to be $\sim 530 \text{ Hz/} \mu\text{W}$.

Figure 3. Measured responses of the fabricated IR detector and a reference device (conventional AlN MEMS resonator) upon exposure to a 5 $\mu$m quantum cascaded laser (QCL) IR source, showing a 7x enhancement in the device responsivity. The inset shows the measured absorption of the fishnet-like metasurface.
The responses of the fabricated AlN PFM IR detector and a reference device (a conventional AlN nano plate resonator with same TCF and $R_0$) to IR radiation were characterized using a 5 μm quantum cascaded laser (QCL) as an illumination source. The measured frequency responses in Figure 3 show that a $\sim 7 \times$ enhanced responsivity was recorded for the PFM IR detector, thanks to its properly engineered absorption properties. Based on the measured noise spectral density of $\sim 1 \text{Hz/Hz}^{1/2}$, the NEP of the detector was calculated by dividing the noise spectral density by the responsivity, and found to be $\sim 1.9 \text{nW/Hz}^{1/2}$.

III. CONCLUSIONS

This paper introduces a transformative approach to the development of high resolution and fast uncooled resonant IR detectors in which an ultrathin piezoelectric fishnet-like metasurface is used to form the vibrating body of a NEMS resonator with a unique combination of optical and electromechanical properties not found in nature. High electromechanical performance (quality factor, $Q \sim 2254$ and electromechanical coupling coefficient, $k_t^2 \sim 1.42\%$) and strong absorption of SWIR radiation ($\sim 60\%$ absorption for a spectral wavelength of 3-5 μm) in an ultra-low volume piezoelectric resonant device are simultaneously achieved, resulting in the first prototype of a high resolution (NEP $\sim 1.9 \text{nW/Hz}^{1/2}$) and fast (thermal time constant $\tau \sim 5.3 \text{ ms}$) piezoelectric NEMS resonant IR detector. The demonstrated resonant PFM detector technology marks a milestone towards the implementation of a new class of high performance, miniaturized and low power IR spectroscopy and multi-spectral imaging systems.

ACKNOWLEDGMENT

The authors wish to thank the staff of the George J. Kostas Nanotechnology and Manufacturing Facility, Northeastern University for their support in device fabrication. This project was supported by the DARPA Young Faculty Award N66001-12-1-4221 and the NSF CAREER Award ECCS-1350114.

REFERENCES


**NSPUDT using c-axis tilted ScAlN Thin Film**

Abhay Kochhar¹, Yasuo Yamamoto², Akihiko Teshigahara², Ken-ya Hashimoto³, Shuji Tanaka⁴ and Masayoshi Esashi¹

¹ WPI-Advanced Institute for Materials Research (WPI-AIMR), Tohoku University, Sendai (JAPAN),
² Research Laboratories, DENSO CORPORATION, Nisshin (JAPAN),
³ Chiba University, Chiba (JAPAN),
⁴ Graduate School of Engineering, Tohoku University, Sendai (JAPAN).

E-mail: abhay.kochhar@ieee.org

Abstract — This paper reports finding of directionality in the 44% Scandium doped Aluminum Nitride thin film based surface acoustic wave devices. Some previously reported SAW devices using bulk substrates showed higher power of acoustic signals in either forward or backward direction depending on their crystal orientations and are called natural single-phase unidirectional transducers (NSPUDT). As these reports were based on bulk substrates, for the first time, we report NSPUDT using c-axis tilted Scandium doped Aluminum Nitride thin film SAW devices on sapphire. In addition, we also examined the c-axis tilt dependency to improve transducers return loss. It is worth to put in notice here that our observance of directionality is specifically in Sezawa wave. Hence, the comparison for both acoustic waves i.e. Rayleigh and Sezawa is also reported.

Keywords—NSPUDT, SAW devices, Sc-doped Aluminum Nitride thin film, c-axis tilted thin film.

I. INTRODUCTION

Surface acoustic wave (SAW) devices are low cost, simple fabrication and produces outstanding repeatability. Its application range is very wide from high-end filters to specific sensors. Previously reported SAW devices [1, 2] using bulk substrates showed higher power of acoustic signals in either forward or backward direction depending on their crystal orientations and are called natural single-phase unidirectional transducers (NSPUDT). These SAW transducers were reported previously for oriented quartz and LiNbO₃ substrates [1], ST-25 X quartz, Y-51.25Z LiTaO₃, 50Y-25X La₃Ga₅SiO₁₄ substrates [2], etc. These substrates observed 15-20 dB of directivity. The basic reason behind the directionality is in the shift in the transduction and reflection coefficient as mentioned by [3, 4]. There are different reported models that can simulate and evaluate the directionality orientations in the substrates [2], [5, 6]. The advantages of NSPUDT [7] are i) less spurious resonance, ii) reduction in triple transit echo, iii) impedance matching easiness, iv) avoiding 3 dB radiation loss in a bidirectional IDT, etc.

The use of bulk substrates has some disadvantages. It increases complications when integrating with sensor or actuator components such as CMOS/MEMS devices. To overcome the integration limitation, monolithically transferred thin film fabrication process [8] will become a great option for the use of thin piezoelectric film based surface acoustic wave devices. Apart from the integrated sensor application such as for high temperature application, the piezoelectric bulk substrates always have a temperature limitation. This could be neglected in the case of thin piezoelectric films as the thermal limit such as in aluminum nitride can go up to 1000 °C. As natural directionality in the SAW devices enables the improvisation in acoustic signals, hence for both integrated sensor and high temperature applications, the thin film based directional SAW devices can be utilized.

Sezawa wave is one of the preferred acoustic modes to be operated at higher frequency due to the lithography limitations. That is why it becomes very important to evaluate the natural directionality in the Sezawa wave SAW-based IDT devices. Recently, doped AlN [9-12] that has higher piezoelectricity are receiving considerable focus for producing increased k² Q product as compared to non-doped AlN in the film based bulk and surface acoustic wave devices and technology development. It was mentioned in [13], that Sezawa wave showed higher coupling factor and higher phase velocity for Scandium doped AlN (ScAlN) film.

In this paper, we report the directionality similar to bulk substrates (due to crystalline orientation) in c-axis tilted ScAlN thin film based SAW devices. Experimental results were performed for c-axis tilt between 0° & 4.5°. The c-axis tilt dependency on directivity and improvement in return loss is reported. In addition, it is found out that the directionality in the thin film produces greater influence on the transducer’s return loss for Sezawa wave as compared to the Rayleigh wave.

II. DEVICE AND EXPERIMENTAL DETAILS

The brief schematic of thin film SAW IDT transducer on c-plane sapphire is shown in Fig. 1. x is the propagation direction. z is the thickness direction. The SAW propagation direction is changed in x-y plane. For this study, the c-axis tilt is in the wave propagation direction (i.e. x). S1, G1 and S2, G2 are source and ground connections for port 1 and port 2 respectively. 44% Sc-doped AlN film on 3” sapphire wafer is utilized. The Sapphire wafer is c-axis oriented and its orientation flat is in the x-axis. In our thin film deposition system, it is possible to obtain varied tilt angle with respect to the thickness direction (z). The c-axis tilting angle depends on the average incident angle of sputtered particle. The distribution of tilting angle is naturally formed due to the...
layout of substrate and target. The c-axis tilt is varied from $0^\circ$ to $4.5^\circ$.

### A. SAW device details

To confirm the directionality due to c-axis tilt in the thin film, previously reported IDT structure [1] as shown in Fig. 2 is utilized. Two device structures are compared, NS: “normal IDT (port 1) – split IDT (port 2)" and SN: “split IDT (port 1) – normal IDT (port 2)". Both configurations have spacing of 50 $\mu$m between port 1 and port 2. In normal IDT, fingers are spaced at $\lambda/4$ having finger width of 0.5 $\mu$m. Number of finger pairs are 40 and the aperture is 40 $\lambda$. Split IDT consists of split-finger electrodes with the width of the finger as 0.5 $\mu$m where each pair of fingers is spaced at 0.5 $\mu$m. The number of finger pairs is limited to 10 and the aperture is 40 $\lambda$. Split IDT is utilized to eliminate internal reflections and thus ensure that it is bidirectional on any crystal orientation [1]. By comparing the NS and SN configurations, directionality in the forward and backward direction is measured respectively. For observing the behavior of transducer’s insertion and return loss, the configuration shown in Fig. 3, NN: “normal IDT (port 1) – normal IDT (port 2)”, is utilized.

![Fig. 1. Two-port thin film ScAlN (1 $\mu$m) based SAW device on sapphire.](image1)

![Fig. 2. SAW IDT configurations utilized to check directionality.](image2)

![Fig. 3. SAW IDT configuration utilized as transversal filter for the observation of return loss behavior due to directionality.](image3)

### III. EXPERIMENTAL RESULTS

#### A. c-axis tilt dependency on directivity and input admittance (Sezawa)

The directivity is calculated by comparing the insertion loss of forward and backward direction. From NS & SN type, the directivity is 1, 7 & 11.5 dB for $0^\circ$, $3^\circ$ & $4.5^\circ$ c-axis tilted ScAlN film respectively as shown in Fig. 4 (a) – (c). This is due to the shift in transduction and reflection coefficient as described in [3, 4]. In addition to the above, the input admittance’s maximum shifts to the lower frequency for larger c-axis tilt. This is basically due to the phase shift between the transduction and reflection centers. The input admittance change due to the c-axis tilt is shown in Fig. 4 (d).

#### B. c-axis tilt dependency on the return loss (Sezawa)

The transversal structure for return loss analysis is shown in Fig. 5 (a). The return loss dependency on c-axis tilt angle is plotted in Fig. 5 (b). The improvement in the return loss from -15.5 dB (S11) as shown in Fig. 5 (c) for c-axis tilt angle of $0^\circ$ to -6.6 dB (S22) as shown in Fig. 5 (f) for c-axis tilt angle of $4.5^\circ$ is observed. It should be noted as shown in Fig. 5 (b), that from c-axis tilt angle of $0^\circ$ to $3^\circ$ the S11 and S22 are opposing.

---

**Experiment details:**

Sc-doped AlN thin films of 2 $\mu$m are prepared by co-sputtering system utilizing dual target comprising $\text{Sc}_{0.6}\text{Al}_{0.4}$ alloy target and Al target in the nitrogen reactive environment. By changing the power of these two targets, 44% Sc-doped AlN thin film is deposited on 3" c-sapphire wafer. The c-axis tilt in (002) of ScAlN film is confirmed by Bruker’s D8 DISCOVER 2D XRD measurement system using pole figure analysis. Electron beam (EB) sensitive ZEP502A resist is spun over the sample at 3000 rpm for 60 sec and baked at 180° C for 3 min. Later, ESpacer 300Z is spun at 1500 rpm for 60 sec to avoid any charging phenomenon during EB writing. IDT patterns are exposed by electron beam (EB) lithography equipment (Elionix ELS-G125S). Fine patterns for up to 50 nm line and space can be fabricated at an EB acceleration voltage of 130 eV. However, in this case for split finger electrodes, 500 nm line and space are patterned with the EB dose of 260 $\mu$C/cm². For development and rinsing, ZED-N50 and ZMD-B solutions are used respectively. Later, metal electrode Au (70 nm) with Cr (20 nm) adhesion layer is deposited by electron beam (EB) evaporation equipment with the respective power requirement for gold and chromium metals. Finally, lift-off process is utilized to remove resist and the leftover metal over resist by stripping the resist using 1165 remover. In addition, IPA rinse and jet spray is utilized to remove any metal leftovers (between the line and space) remained in the lift-off process. For characterization of two port SAW devices, on wafer RF prober system with Cascade’s GS probes and Agilent’s E5071C Vector Network Analyzer (VNA) is utilized. VNA’s time domain analysis with the gating option is utilized to isolate the reflection areas. A touchstone format “.s2p” files are later analyzed in the MATLAB.

---

**TRANSLATION:**

Sc-doped AlN thin films of 2 $\mu$m are prepared by co-sputtering system utilizing dual target comprising $\text{Sc}_{0.6}\text{Al}_{0.4}$ alloy target and Al target in the nitrogen reactive environment. By changing the power of these two targets, 44% Sc-doped AlN thin film is deposited on 3" c-sapphire wafer. The c-axis tilt in (002) of ScAlN film is confirmed by Bruker’s D8 DISCOVER 2D XRD measurement system using pole figure analysis. Electron beam (EB) sensitive ZEP502A resist is spun over the sample at 3000 rpm for 60 sec and baked at 180° C for 3 min. Later, ESpacer 300Z is spun at 1500 rpm for 60 sec to avoid any charging phenomenon during EB writing. IDT patterns are exposed by electron beam (EB) lithography equipment (Elionix ELS-G125S). Fine patterns for up to 50 nm line and space can be fabricated at an EB acceleration voltage of 130 eV. However, in this case for split finger electrodes, 500 nm line and space are patterned with the EB dose of 260 $\mu$C/cm². For development and rinsing, ZED-N50 and ZMD-B solutions are used respectively. Later, metal electrode Au (70 nm) with Cr (20 nm) adhesion layer is deposited by electron beam (EB) evaporation equipment with the respective power requirement for gold and chromium metals. Finally, lift-off process is utilized to remove resist and the leftover metal over resist by stripping the resist using 1165 remover. In addition, IPA rinse and jet spray is utilized to remove any metal leftovers (between the line and space) remained in the lift-off process. For characterization of two port SAW devices, on wafer RF prober system with Cascade’s GS probes and Agilent’s E5071C Vector Network Analyzer (VNA) is utilized. VNA’s time domain analysis with the gating option is utilized to isolate the reflection areas. A touchstone format “.s2p” files are later analyzed in the MATLAB.
each other and hence the difference is increasing. However, an increase in both $S_{11}$ and $S_{22}$ by 4 dB was observed for c-axis tilt angle of 4.5° as compared to the 3°.

C. Sezawa and Rayleigh wave comparison

The comparison of $|S_{11}|$-$|S_{22}|$ for Rayleigh and Sezawa wave for 2° c-axis tilt in the ScAlN thin film is reported in Fig. 6. The propagation direction of SAW device is changed in the x-y plane. As it can be evaluated from the figure, the change in Rayleigh wave is not as large as observed in Sezawa wave. This is primarily due to the higher coupling factor of the Sezawa wave. Fig. 7 shows the comparison for 90° propagation at 2° c-axis tilted ScAlN film. The difference in $S_{11}$ & $S_{22}$ for Sezawa wave is 13.3 dB as compared to just marginal 1.25 dB for Rayleigh wave.

![Fig. 4. (a) – (c) Insertion loss for the forward and backward signals for 0°, 3° & 4.5° c-axis tilted ScAlN film utilizing split IDT to analyze the directionality; (d) Input Conductance and Susceptance vs Frequency for c-axis tilts 0°, 3° & 4.5°.](image)

![Fig. 5. (a) Transversal filter structure for return loss analysis, (b) c-axis tilt dependency on return loss, and (c)-(f) $S_{11}$ and $S_{22}$ for c-axis tilts 0°, 1.5°, 3° and 4.5°, respectively.](image)
The directionality in ScAlN thin film due to the c-axis tilt orientation in the surface acoustic wave (SAW) devices is reported. The directionality is confirmed by comparing forward and backward direction insertion loss for Normal IDT – Split IDT and Split IDT – Normal IDT two-port structures. Sezawa wave is monitored from 0° to 4.5° c-axis (z) tilted Sc-AlN film on sapphire. The directivity is found to be 1, 7 & 11.5 dB for 0°, 3° & 4.5° c-axis tilted ScAlN thin film respectively. For analyzing return loss improvement, the transversal filter structure is utilized. In Sezawa wave, the return loss for tilts from 0° to 4.5° showed shift (improvement) of as large as 18 dB for 90° propagation at 2° c-axis tilted ScAlN film. Finally, the Rayleigh and Sezawa wave of a SAW device with 2° c-axis tilted ScAlN thin film is analyzed for both directionality and the improvement in transducers loss (by changing the propagation wave directions from 0° to 150° in the x-y plane). Rayleigh wave did not show much improvement as large as Sezawa wave. The originality in finding the directionality for the thin film based SAW devices is to allow improvement in transducers loss i) to be utilized at higher temperature (if needed with proper compensation using positive temperature coefficient of SiO2 film) and ii) for the integrated sensor applications. Such improvement will have great impact on the applications utilizing thin film based SAW devices.

ACKNOWLEDGMENT

Authors would like to acknowledge Prof. M. Kadota of Tohoku University, Mr. T. Iwaki and Mr. H. Wado of Research Laboratories, DENSO CORPORATION for the rigorous technical discussions. This work was supported by World Premier International Research Center Initiative (WPI), MEXT, Japan.

REFERENCES

Abstract— This paper describes the achieved results of a new space Rubidium Atomic Frequency Standard (named Robust-RAFS) through the description of the internal coefficients influence reduction and the positive consequences in term of clock frequency stability and predictability improvements. Performances achievements during uninterrupted operation of several months demonstrate a monotonic behavior, a stability of $1\times10^{-14}$ @10^5 sec. and a drift per day of few $10^{-14}$.

Keywords—Space; atomic clock, Rubidium, frequency stability, RAFS, GNSS

I. INTRODUCTION

Accurate and ultra-stable atomic clocks have been recognized as the critical equipment for the precision Global Navigation Satellite Systems (GNSS). In parallel to Rb clocks produced for industrial applications, SpectraTime (SpT) is a space clocks manufacturer of Rubidium Atomic Frequency Standard (RAFS) for various navigation systems (European, Chinese and Indian) and other programs. In the last ten years, the 56 SpT RAFS units in-orbit have cumulated 170 years of flight heritage. In addition, almost 130 SpT RAFS flight units have been manufactured and characterized. Those compact clocks (3 kg only) provide good performances and reach the expected reliability figure. The foreseen stability up to two hours is in specification and, analysis of the telemetries demonstrates a nominal behavior over the time. Nevertheless, from the long term records of the frequency stability of those units, it appears that a weakness exists when considering the predictability of their behavior over time period interval longer than few hours. As long as the GNSS infrastructure was designed to compensate the frequency variations within an interval of one or two hours, this issue was not too critical.

According to the latest GNSS requirements, such RAFS behavior becomes critical. Predictability of the frequency is expected over much longer period of operation. In order to reach such new requirements, an update of the RAFS physics package was initiated through ESA EGEP funding in 2012.

II. SUMMARY OF THE WORK PERFORMED

The execution of the activities was divided in two Phases.

- Phase 1 was the Design Assessment and Definition. This first phase was dedicated to the review and assessment of the current RAFS design and performance. This first phase also addressed the definition of the final EM Robust-RAFS. The phase 1 was concluded with a Design Review held with ESA.
- Phase 2 was the Manufacture, Test and Validation. This second phase was dedicated to the manufacture, assembly, test and validation of the Robust-RAFS in accordance with the plan agreed at the Design Review with ESA. The project was terminated by a Final Review where all the development results were commented and conclusions agreed with ESA.

III. IMPROVEMENT APPROACH

In order to improve the robustness of the existing RAFS design, the reported occurrence and the amplitude of frequency anomalies were strongly reduced. The origin of the frequency anomalies could be internal or external to the clock. In this development, mainly the internal causes were considered. The SpT strategy was to reduce the influence of the identified causes through their intrinsic coefficients. Essentially, the light shift and the power shift coefficients were concerned. The light shift being the dominant factor of the two.

IV. PHYSIC PACKAGE CONFIGURATIONS AND ASSOCIATED ELECTRICAL DESIGN

Two main configurations were investigated:

- Separated Rb85 filter. This configuration is working with a Rb87 lamp, a Rb85 isotopic filter and a Rb87 microwave frequency discriminator cell. Each of those elements is thermally regulated separately and their operational temperatures are optimized.
- Integrated Rb85 filter. In this configuration, the isotopic filter is included within the microwave frequency discriminator cell. In such case, the Rb isotopes are mixed. The temperature is identical for the two functions.
V. SUMMARY OF THE DIFFERENT CONFIGURATION TESTED

The prototypes were mounted with separated isotopic filter or integrated filter, operated with plasma in different modes and various optical filters. In total twelve configurations were first evaluated. From which, after this initial performance evaluation, six were selected and further studied.

In each of these six configurations, an important work of optimization of the different temperatures (lamp, filter, cell), gas pressure and geometry was performed. In addition, the volume of the micro-wave cell was increased to have a better homogeneity within the active area where the atoms are interrogated.

At the end of this process, one configuration was selected, integrated within an Engineering Model and intensively tested. The integrated filter solution was selected. It provides the best value in term of reliability due to the reduction of active elements without penalizing the frequency stability performances. It also simplifies the adjustment procedure.

VI. GENERAL MECHANICAL AND THERMAL DESIGN

The RAFA unit is composed of two main parts. The clock itself named “RAFA core” and the Electronic Power Conditioning named “EPC” which includes the DC/DC converter and the electrical interface to the satellite. The “RAFA core” includes the microwave cell and the lamp.
VII. RAFS Core

The Core module is thermally coupled with a cold base-plate through calibrated thermal resistance in order to allow its thermal regulation by the use of heaters. One dedicated cover fixed to the intermediate thermally regulated base-plate will provide magnetic and radiation shield, and a second cover fixed to the cold base-plate will reinforce the shielding function.

Fig 4. RAFS core details with PCB’s integrated covers removed

VIII. DC/DC Converter Design (EPC)

The EPC module is accommodated in a separate box beside the RAFS core. The base plate is common between the EPC and the RAFS core. Figure below shows the complete EPC housing and the connector locations. The EPC is covered by a steel housing for radiation shield. The secondary voltages U1 and U2 are connected via feed through EMC filters. This will minimize the EMC influence of the EPC to the RAFS core in the frequency range above 10 MHz.

IX. Achieved Results

An engineering model (EM) was realized including the selected improved microwave cell, lamp and other constituents. A new EPC design was also implemented.

As expected on the base of the prototypes preliminary results, the critical coefficients of the EM were reduced by significant factors:

- light shift coefficient; reduction factor of 10 versus average current design results
- power shift coefficient; reduction factor of 5.

In addition, the atomic line bandwidth was reduced from 1 KHz to 300 Hz leading to a better clock frequency stability both in short and long term.

<table>
<thead>
<tr>
<th></th>
<th>Current RAFS</th>
<th>Robust-RAFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light shift ($10^{-12}$/%)</td>
<td>100 to 200</td>
<td>+10</td>
</tr>
<tr>
<td>Pressure shift ($10^{-12}$/mbar/°C)</td>
<td>+50</td>
<td>+1.0</td>
</tr>
<tr>
<td>Power shift ($10^{-12}$/%)</td>
<td>+15</td>
<td>+3</td>
</tr>
</tbody>
</table>

Table 1. Coefficients comparison between current RAFS and Robust-RAFS

Fig 5. EPC Module Housing

Fig 6. Illustration of the light shift coefficient improvement. On top, current design. A light intensity discontinuity (blue trace) induces a frequency discontinuity (red trace). Below Robust-RAFS EM measurement, the light discontinuity has a negligible effect on the frequency.
It must be noted that the new design does not introduce new processes in order to benefit of the previous qualification heritage.

Encouraged by those promising design improvement results, an intensive test campaign was conducted on the EM. After the usual characterizations in terms of electrical interfaces, phase noise and EMC, a long term measurement was performed with a total of 9 months of measurement in three periods of three months. The interruptions were introduced to verify the retrace of the clock ($< 5\times 10^{-11}$). During the nine months of measurement, several light intensity discontinuities occurred with negligible impact on the frequency stability performances.

A second important result is the better predictability of the frequency over long observation period. The improvements of internal coefficients also minimize their impact on the frequency drift evolution over the time. The comparison between the GIOVE-A results published by ESA and latest Robust RAHS ones illustrates this better predictability.

Finally, the good overall Robust-RAFS behavior is well demonstrated by the Hadamard Deviation result. The stability is around $1\times 10^{-14}$ at one day. A linear drift of $-3.4\times 10^{-14}$/day is measured after few weeks of stabilization. A dynamic Allan Deviation plot is provided below for the demonstration of the overall frequency stability performance of the Robust RAFS.

Fig 7. On the left, Frequency evolution over the time of GIOVE-A RAHS as per ESA document SP-1320 (end of year 2008). On the right, three months continued measurement of the Robust-RAFS EM. The frequency drift predictability is strongly improved.

Fig 8. Robust-RAFS EM frequency stability expressed through Hadamard Deviation.
Fig 9. Robust-RAFS EM frequency stability expressed through Dynamic Allan Deviation (drift removed)

Fig 10. Robust RAFS Engineering Model
REFERENCES


Use of two traveling GPS receivers for a relative calibration campaign among European laboratories

P. Uhrich, G.D. Rovera, B. Chupin
LNE-SYRTE
Observatoire de Paris, LNE, CNRS, UPMC
Paris, France
Pierre.Uhrich@obspm.fr

J. Galindo, H. Esteban
Real Instituto y Observatorio de la Armada
San Fernando, Spain

K. Jaldehag, C. Rieck
SP Technical Research Institute of Sweden
Borås, Sweden

A. Bauch, T. Polewka
Physikalisch-Technische Bundesanstalt
38116 Braunschweig, Germany

G. Cerretto, G. Fantino
Istituto Nazionale di Ricerca Metrologica (INRIM)
Torino, Italy

R. Piriz
GMV
Tres Cantos, Spain

Abstract—We report about a GPS receiver relative calibration campaign, which took place between five European National Metrology Institutes or Designated Institutes: LNE-SYRTE in Observatoire de Paris (Paris, France), where the reference receiver of the campaign was located, ROA (San Fernando, Spain), SP (Borås, Sweden), PTB (Braunschweig, Germany) and INRIM (Torino, Italy). We used as traveling equipment two main units, both connected to a single antenna, and we kept track of the offset between both traveling units in all the visited sites. An external validation of the resulting hardware delays is provided against the time scale differences derived from the UTC – UTC(k) data published by BIPM in its monthly Circular T. Thanks to a very good stability of the traveling equipment, we obtained expanded uncertainty estimates within 2.0 ns (k = 2) for the hardware delays.

Keywords—Time transfer, GPS receiver, relative calibration.

I. INTRODUCTION

One well-known limiting factor of GPS receiver relative calibration of hardware delays is the stability during the measurement campaign of the traveling equipment [1, 2, 3]. One way to estimate such stability is to compute the difference of the traveling receiver hardware delays between the start and the end of the campaign, which is called the deviation from closure, against the reference receiver of the campaign. But unexpected events on the traveling receiver delays at a given location might remain unnoticed, introducing that way a bias in the calibration results for this site. The use of two receivers as traveling equipment might help to detect abnormal behavior during the campaign or to study effects not clearly understood yet [4, 5]. In this paper, we report about a GPS receiver relative calibration campaign, which took place during Autumn 2014 between five European National Measurement Institutes or Designated Institutes listed in visited order: LNE-SYRTE in Observatoire de Paris (OP, Paris, France), where the reference receiver of the campaign was located, Real Instituto y Observatorio de la Armada (ROA, San Fernando, Spain), SP Technical Research Institute of Sweden (SP, Borås, Sweden), Physikalisch-Technische Bundesanstalt (PTB, Braunschweig, Germany) and Istituto Nazionale de la Ricerca Metrologica (INRIM, Torino, Italy). We used as traveling equipment two main units, both connected to a single antenna.

After description of the organization of the campaign and characterization of the traveling equipment, we show an example of the collected data. We kept track of the offset between both traveling receivers in all the visited sites, providing that way an analysis of their stability during the campaign. The external validation of the resulting hardware delays is obtained either against the time scale differences derived from the UTC – UTC(k) data published monthly by the Bureau International des Poids et Mesures (BIPM) in Circular T, or against another time transfer technique: Two-Way Satellite Time and Frequency Transfer (TWSTFT). We finally provide an example of uncertainty budget computations, leading to the k = 2 expanded estimates for all receivers.

II. ORGANIZATION OF THE CAMPAIGN

A. Laboratories and receivers involved

This relative calibration campaign took place from August to December 2014. The manufacturers and types of the GPS receivers involved are as follows.

- OP: OPM1 (ASHTECH Z12-T), OPM7 and OPM8 (both Septentrio POLARX4).
- ROA: RO_5 (DICOM GTR50) and RO_6 (Septentrio PolaRx3aTR).
- SP: SP01 and SP02 (both JAVAD EGGD).
- PTB: PT07 (DICOM GTR50), PT10 (DICOM GTR51) and PTBB (ASHTECH Z12-T).
- INRIM: GTRB and GTRI (both DICOM GTR50) and IENG (ASHTECH Z12-T)*.

B. Characterization of OP traveling equipment

The reference receiver of the campaign, OPMT, had been relatively calibrated twice by BIPM against an absolutely calibrated receiver, and stayed continuously monitored since, in particular against OPM2 in a common-clock and common antenna set-up. OPM7 and OPM8 were the traveling main units. We used two short cables and one power splitter combiner to connect both main units to the same antenna cable and Choke-Ring antenna. A Time Interval Counter (TIC) was also part of the traveling equipment, aiming at minimizing the remaining bias for cable delay measurements on site. Table I provides for both OPM7 and OPM8 the P-code delays on both L-band GPS carriers, hence P1- and P2-code delays, measured against OPMT at the start and at the end of the campaign. The deviation from closure stays between 0.3 and 0.7 ns, which shows an apparent sub-ns stability of the traveling equipment. The mean values were used for the receiver hardware internal delay computations in all the visited sites, OPMT being assumed as perfectly calibrated.

<table>
<thead>
<tr>
<th>TABLE I.</th>
<th>OPM7 AND OPM8 DELAYS AGAINST OPMT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P-code</td>
</tr>
<tr>
<td>OPMT</td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td>P2</td>
</tr>
<tr>
<td></td>
<td>P1</td>
</tr>
<tr>
<td></td>
<td>P2</td>
</tr>
</tbody>
</table>

III. Collected data

For all the receivers implemented in common-clock set-up, GPS data were collected in the geodetic RINEX 2.1 format [6], 30 s sampled. We compute the antenna coordinates by using the Precise Point Positioning (PPP) software developed by National Resources Canada (NRCan) [7], and we build the P1- and P2-code differences between a traveling receiver and a local one in each site, Fig. 1 shows an example of such data between PTBB and OPM8 when implemented in PTB, each mean value providing the required PTBB delays. For all the involved receivers, we have computed the mean values of the delays obtained by using OMP7 and OPM8 as the results of the calibration [8]. In addition, we also compute the Allan Time Standard Deviation (TDEV) of such data, as can be seen in Fig. 2 based on the example given in Fig. 1. For GPS receivers in common-clock set-up, a Flicker Phase Noise Modulation takes over from the White Noise Modulation at about a 1 d averaging period [9]. Therefore, as a minimum, at least three days of common data have to be continuously collected on site, in order to obtain a significant TDEV at 1 d, which is providing the resulting noise of the calibration.

Fig. 1. P-code differences between PTBB and OPM8 from October 15 to 29, 2014, built from RINEX files, in blue for P1 and in lilac for P2.

Fig. 2. TDEV of P-Code differences between PTBB and OPM8 from MJD 56945 to 56959, in blue for P1 and in lilac for P2.

IV. Traveling equipment stability

We have computed Common-Views (CV) between the two traveling receivers in common-clock and common antenna set-up when implemented in each one of the visited sites. We used for CV computations CGGTTs data files [10,11] generated from RINEX files according to the TAIP3 processing [12] developed by P. Defraigne (Royal Observatory of Belgium) and available on the BIPM website, together with an averaging software developed in OP [13]. The TAIP3 CV between OPM7 and OPM8 have been computed by using as internal delay parameters the mean values of the delays against OPMT between the start and the end of the campaign, provided in Table I above. Fig. 3 shows the TAIP3 CV between OPM7 and OPM8 in all the visited sites, the mean values of such direct comparisons being given in Table II. We see that during the campaign both traveling receivers stayed close to each other within 0.60 ns peak to peak of the mean values, the noise staying consistently well within 100 ps in all sites. This offset seems mostly due to a small apparent change between the site, together with an averaging software developed in OP [13]. The TAIP3 CV between OPM7 and OPM8 have been computed by using as internal delay parameters the mean values of the delays against OPMT between the start and the end of the campaign, provided in Table I above. Fig. 3 shows the TAIP3 CV between OPM7 and OPM8 in all the visited sites, the mean values of such direct comparisons being given in Table II. We see that during the campaign both traveling receivers stayed close to each other within 0.60 ns peak to peak of the mean values, the noise staying consistently well within 100 ps in all sites. This offset seems mostly due to a small apparent change between the

* Product names and model numbers of the equipment are included for reference only. No endorsement or critique is implied.
receivers when located in PTB with respect to the other locations. We nevertheless obtain here again a sub-ns stability for the traveling equipment during the campaign.

Fig. 3. Averaged TAIP3 CV between OPM7 and OPM8 in all the visited sites during the relative calibration campaign: OP at the start (green), ROA (blue), SP (lilac), PTB (cyan), INRIM (yellow), and back to OP at the end (black).

TABLE II. AVERAGED TAIP3 CV MEAN VALUES BETWEEN OPM7 AND OPM8 IN COMMON-CLOCK AND COMMON-ANTENNA SET-UP FROM THE DATA PLOTTED IN FIG. 3.

<table>
<thead>
<tr>
<th>OPM8 – OPM7</th>
<th>TAIP3 CV mean value (ns)</th>
<th>Standard deviation (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP (campaign start)</td>
<td>0.13</td>
<td>0.08</td>
</tr>
<tr>
<td>ROA</td>
<td>-0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>SP</td>
<td>-0.16</td>
<td>0.08</td>
</tr>
<tr>
<td>PTB</td>
<td>0.44</td>
<td>0.07</td>
</tr>
<tr>
<td>INRIM</td>
<td>-0.15</td>
<td>0.08</td>
</tr>
<tr>
<td>OP (campaign end)</td>
<td>-0.05</td>
<td>0.08</td>
</tr>
</tbody>
</table>

V. VALIDATION OF THE RESULTS

A. Consistency of calibrated delays

The TAIP3 CV between the traveling equipment and the local equipment have been similarly computed, using as parameters for the local receivers the newly calibrated delays. What we expect here are mean values close to 0 ns together with a given standard deviation, both values small enough to consider a proper consistency of the computations. The mean values and related noise are given in Table III. We see that the largest deviation from 0 is 0.32 ns, which clearly stays below the noise level as all standard deviations are within 0.70 ns. The figures obtained for either OPM7 or OPM8 when implemented in PTB also seem to show that the offset observed in the Section above is mostly coming from a deviation of OPM7. But the origin of this deviation remains unknown. This result nevertheless provides confidence in the data collected and in the computations achieved.

B. External validation against BIPM Circular T

This external validation is based on the comparison of resulting TAIP3 CV time transfer after calibrated delays implementation against BIPM Circular T [14]. For each visited laboratory k, we compute the link to UTC(OP) by using the 5 d sampled UTC – UTC(OP) and UTC – UTC(k) data. Note that, for all the links considered here, BIPM uses preferably TWSTFT data for Circular T, together with some long-term smoothing. Hence some of the discrepancies appearing here below are issued from the calibration deviations between both techniques. In Circular T, the combined uncertainty on UTC – UTC(OP) is given as u = 2.1 ns in September 2014 and u = 1.3 ns from October to December 2014. TAIP3 CV were computed not only for the local receivers but also by using the OPM7 and OPM8 traveling equipment on site.

Fig. 4 shows this external validation for ROA. We see on the mid-point of the interval a sub-ns consistency between OP-ROA TAIP3 CV and the point issued from BIPM Circular T 321. This confirms the good quality of the relative calibration results for ROA receivers. Note however that the combined uncertainty on the UTC – UTC(ROA) differences published at that time in Circular T is given as u = 5.0 ns. Fig.5 shows a similar plot for SP. Here too, the consistency between the TAIP3 CV and BIPM Circular T 321 is sub-ns for the two points available from Circular T, where, for UTC – UTC(SP), the combined uncertainty is given as u = 1.2 ns. Fig. 6 shows the related plot for INRIM. Again, the consistency between the TAIP3 CV and BIPM Circular T 323 is sub-ns for all three points available from Circular T, where, for UTC – UTC(IT), the combined uncertainty is given as 1.4 ns.

TABLE III. AVERAGED TAIP3 CV MEAN VALUES BETWEEN GPS RECEIVERS IN COMMON-CLOCK SET-UP AFTER APPLICATION OF THE CALIBRATION RESULTS.

<table>
<thead>
<tr>
<th>TAIP3 CV</th>
<th>OPM7 (ns)</th>
<th>Standard deviation (ns)</th>
<th>OPM8 (ns)</th>
<th>Standard deviation (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPMT (start)</td>
<td>0.26</td>
<td>0.66</td>
<td>0.12</td>
<td>0.65</td>
</tr>
<tr>
<td>RO 5</td>
<td>-0.07</td>
<td>0.70</td>
<td>-0.02</td>
<td>0.70</td>
</tr>
<tr>
<td>RO 6</td>
<td>0.09</td>
<td>0.40</td>
<td>0.16</td>
<td>0.41</td>
</tr>
<tr>
<td>SP01</td>
<td>-0.25</td>
<td>0.50</td>
<td>-0.09</td>
<td>0.49</td>
</tr>
<tr>
<td>SP02</td>
<td>-0.14</td>
<td>0.50</td>
<td>0.02</td>
<td>0.60</td>
</tr>
<tr>
<td>PT07</td>
<td>0.32</td>
<td>0.61</td>
<td>-0.11</td>
<td>0.62</td>
</tr>
<tr>
<td>PT10</td>
<td>0.26</td>
<td>0.65</td>
<td>-0.18</td>
<td>0.65</td>
</tr>
<tr>
<td>PTBB</td>
<td>0.29</td>
<td>0.53</td>
<td>-0.14</td>
<td>0.53</td>
</tr>
<tr>
<td>GTRB</td>
<td>-0.05</td>
<td>0.63</td>
<td>0.10</td>
<td>0.63</td>
</tr>
<tr>
<td>GTRI</td>
<td>-0.02</td>
<td>0.65</td>
<td>0.13</td>
<td>0.65</td>
</tr>
<tr>
<td>IENG</td>
<td>0.06</td>
<td>0.47</td>
<td>0.20</td>
<td>0.48</td>
</tr>
<tr>
<td>OPMT (end)</td>
<td>0.00</td>
<td>0.61</td>
<td>0.07</td>
<td>0.61</td>
</tr>
</tbody>
</table>
UTC(ROA) – UTC(OP) as obtained from TAIP3 CV between all the receivers implemented in ROA during the calibration campaign and OPMT in OP compared to data issued from BIPM Circular T 321.

UTC(SP) – UTC(OP) as obtained from TAIP3 CV between all the receivers implemented in SP during the calibration campaign and OPMT in OP compared to data issued from BIPM Circular T 321.

UTC(IT) – UTC(OP) as obtained from TAIP3 CV between all the receivers implemented in INRIM during the calibration campaign and OPMT in OP compared to data issued from BIPM Circular T 323.

Finally Fig. 7 shows a similar plot for PTB, where PTBB measurements are considered, by using the former delay values or the newly calibrated ones. Similar results are obtained from PT07 or PT10 data. We notice here an offset of about 2 ns between the TAIP3 CV and BIPM Circular T 322. But this can be explained by the fact that, on the operational link between OP and PTB, the average offset between TAIP3 CV and TWSTFT deviated to about 2 ns around the calibration campaign period after having stayed below 1 ns for years. In other words, the mean offset appearing in Fig. 7 is clearly due to the operational offset observed on this link. This validates the GPS calibration, because the reference receiver in OP is in both cases OPMT. But the source of this offset between both techniques remains under investigation. In Circular T, the combined uncertainty on UTC – UTC(PTB) is given as $u = 0.7$ ns, PTB being the pivot for TAI computation.

UTC(PTB) – UTC(OP) as obtained by TAIP3 CV between PTBB and OPMT by using former delays or newly calibrated delays of PTBB, compared to data issued from BIPM Circular T 322.

C. Direct comparison with TWSTFT

This is confirmed by the direct comparison between TAIP3 CV and TWSTFT when using OP as pivot laboratory. In Fig. 8, we report the average differences between the TWSTFT and TAIP3 CV time transfer [3] on the links between OP and the four other laboratories, TAIP3 CV being based on former hardware delays. For the limited period of the actual calibration on each site, we also report in bold symbols these differences after implementation of newly calibrated delays. We selected data from only one receiver in each site, but the results are similar with any of the other remote receivers.

Differences between TWSTFT and TAIP3 CV on the links between OP and either INRIM (lilac), PTB (red), ROA (green) or SP (blue). Bold symbols are obtained after implementation of the newly calibrated delays.

For PTB and ROA, one can see that the newly calibrated delays do not significantly change the offset of about 2 ns between both time transfer techniques. For INRIM and SP, the obtained changes lead to offsets closer to 0 ns, from about 3 ns to about 1 ns for SP, and from about 5 ns to about 1 ns for INRIM. Clearly, the GPS calibration provided an improved convergence between both time transfer techniques.
VI. UNCERTAINTY BUDGETS

A. Terms to be taken into account

We consider that a delay measured by using a TIC can be achieved within an uncertainty of 0.20 ns (k = 1) by differential measurements. This is valid for 1 pulse per second (PPS) signal cable delay where the noise of the measurements typically stays below 10 ps. Such an uncertainty is applied for GPS receivers requiring a 1 PPS input signal only, like DICOM GTR50 or GTR51, and JAVAD receivers: RO_5, SP01, SP02, PT07, PT10, GTRB and GTRI. We also use a TIC to measure the 1 PPS output signal from a Septentrio receiver, which, when properly configured, allows getting access to the internal reference point of the main unit. But here the measurement noise is typically about 90 ps. The quadratic sum of both leads to a combined uncertainty of about 0.22 ns (k = 1). Such an uncertainty is valid for PolaRx3 and PolaRx4 receivers: OPM7, OPM8, and RO_6. The reference signal delay for ASHTECH Z12-T units is made of two different parts: a 1 PPS input signal, which is measured by using a TIC as discussed above, and the offset between this 1 PPS signal rising edge and the next zero crossing of an inverted 20 MHz signal based on the 10 MHz frequency distribution from the same reference time scale. When using an oscilloscope, this second measurement uncertainty is estimated to be about 0.30 ns (k = 1), and this reference delay uncertainty is then a combined uncertainty of about 0.36 ns. Potential instabilities of the reference receiver of the campaign might have an impact too. From OPM7 long-term analysis in general [6], and from OPM7 monitoring against OPMT during this calibration campaign, we estimate OPMT stability within 0.04 ns.

In addition, we also consider the noise of the differences between P1- and P2-code respectively when computing the calibrated delays between a local receiver and a traveling one. We used the upper limit of the statistical uncertainty on the TDEV we can compute for an analysis period the closest possible to 1 d. Similarly, we consider as time transfer noise the TDEV at 1 d for the TAIP3 CV between remote receivers.

The uncertainty budgets are made of quadratic sums of all the terms above. But in addition, the deviation of the traveling equipment during the campaign has to be considered too, which is one of the good reasons to travel with two main units apart from reliability issues. The largest average closure deviation between the start and the end of the campaign on the P-code delays of OPM7 or OPM8 against OPMT is 0.70 ns, as given in Table I. The largest average TAIP3 CV offset between OPM7 and OPM8 in all different locations is 0.60 ns peak to peak, as is obtained from Table II. And all averaged TAIP3 CV between receivers in common-clock set-up after application of the newly calibrated delays are staying below 0.32 ns with a noise below 0.70 ns, as given in Table III. Therefore, we have decided conservatively to consider the largest of these deviations as the traveling equipment stability over the campaign, that is 0.70 ns. Because inside the European Regional Metrology Organization EURAMET the uncertainties have to be published as expanded uncertainties for k = 2 (95 %), it had been proposed years ago [15] to consider half of such deviation, hence 0.35 ns here, to be added as simple sum to the quadratic sum of all other terms. When expanding the uncertainty for k = 2, the complete instability of the traveling equipment is adequately included.

B. Example of uncertainty budgets

Table IV provides the P1- and P2-code hardware delay uncertainty budgets of SP01 with respect to OPMT via OPM7, very similar results having been obtained via OPM8. Table V provides the TAIP3 CV uncertainty budgets for time transfer between OPMT and SP01 calibrated via OPM7 or OPM8. Note that the terms about the main unit connections to the external local references are common to both codes in Table IV or to both traveling receivers in Table V.

<table>
<thead>
<tr>
<th>Table V.</th>
<th>TAIP3 CV HARDWARE DELAY UNCERTAINTY BUDGETS OF TIME TRANSFER BETWEEN OPMT AND SP01 VIA EITHER OPM7 OR OPM8</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP01 against OPMT via OPM7</td>
<td>u(P1) (ns)</td>
</tr>
<tr>
<td>OP reference for OPMT</td>
<td>0.36</td>
</tr>
<tr>
<td>OP reference for OPM7</td>
<td>0.22</td>
</tr>
<tr>
<td>OPMT stability during the campaign</td>
<td>0.04</td>
</tr>
<tr>
<td>SP reference for SP01</td>
<td>0.20</td>
</tr>
<tr>
<td>SP reference for OPM8</td>
<td>0.22</td>
</tr>
<tr>
<td>TDEV(1 d) of OPMT – SP01</td>
<td>0.09</td>
</tr>
<tr>
<td>TDEV(1 d) of OPM7 – OPMT</td>
<td>0.06</td>
</tr>
<tr>
<td>Quadratic sum</td>
<td>0.53</td>
</tr>
<tr>
<td>Half closure deviation</td>
<td>0.35</td>
</tr>
<tr>
<td>Simple sum</td>
<td>0.88</td>
</tr>
<tr>
<td>k = 2</td>
<td>1.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table IV.</th>
<th>P1- AND P2-CODE HARDWARE DELAY UNCERTAINTY BUDGETS OF SP01 WITH RESPECT TO OPMT VIA OPM7</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP01 against OPM7 via OPM7</td>
<td>u(P1) (ns)</td>
</tr>
<tr>
<td>OP reference for OPMT</td>
<td>0.36</td>
</tr>
<tr>
<td>OP reference for OPM7</td>
<td>0.22</td>
</tr>
<tr>
<td>OPMT stability during the campaign</td>
<td>0.04</td>
</tr>
<tr>
<td>SP reference for SP01</td>
<td>0.20</td>
</tr>
<tr>
<td>SP reference for OPM7</td>
<td>0.22</td>
</tr>
<tr>
<td>TDEV(1 d) of OPMT – OPM7</td>
<td>0.06</td>
</tr>
<tr>
<td>TDEV(1 d) of OPM7 – SP01</td>
<td>0.09</td>
</tr>
<tr>
<td>SP reference for OPM7</td>
<td>0.22</td>
</tr>
<tr>
<td>SP reference for OPM8</td>
<td>0.22</td>
</tr>
<tr>
<td>TDEV(1 d) of OPM7/OPM8 – SP01</td>
<td>0.19</td>
</tr>
<tr>
<td>TDEV(1 d) of TAIP3 CV OPM7/OPM8 – SP01</td>
<td>0.16</td>
</tr>
<tr>
<td>Quadratic sum</td>
<td>0.58</td>
</tr>
<tr>
<td>Half closure deviation</td>
<td>0.35</td>
</tr>
<tr>
<td>Simple sum</td>
<td>0.93</td>
</tr>
<tr>
<td>k = 2</td>
<td>1.9</td>
</tr>
</tbody>
</table>

These hardware delay uncertainty estimates are for time transfer to OPMT. For time transfer against a receiver not part
of this small network, the uncertainty from OPMT calibrated delays should additionally be taken into account, except in the case OPMT was also the reference receiver for the calibration of this external receiver during another calibration campaign.

C. Summary of the expanded uncertainties

According to the computation mode described above, the estimates for this calibration campaign have to be given as \( k = 2 \) expanded uncertainties. In summary, the \( k = 2 \) expanded uncertainties \( u(P1) \) and \( u(P2) \) for either P1- and P2-codes or \( u(P3) \) for TAIP3 CV are provided in Table VI for all receivers in all stations, OPMT in OP being the reference receiver. As all uncertainties have been computed either via OPM7 or via OPM8, we have chosen each time the worst figure rounded to the upper value as conservative estimates. This relative calibration campaign lead to uncertainties close to the state of the art, which is mostly due to the very good stability of the traveling and reference equipment during the period.

### Table VI

**Summary of the Conservative Expanded Uncertainties for all the Receivers with OPMT as Reference**

<table>
<thead>
<tr>
<th>Expanded uncertainty ( k = 2 )</th>
<th>( u(P1) ) (ns)</th>
<th>( u(P2) ) (ns)</th>
<th>( u(P3) ) (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RO_5</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>RO_6</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>SP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP01</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>SP02</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>PTB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PT07</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>PT10</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>PTBB</td>
<td>1.9</td>
<td>1.9</td>
<td>2.0</td>
</tr>
<tr>
<td>INRIM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTRB</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>GTRI</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>IENG</td>
<td>1.9</td>
<td>1.9</td>
<td>2.0</td>
</tr>
</tbody>
</table>

VII. Conclusion

We reported about a GPS receiver relative calibration campaign, which took place between OP, where the reference receiver was located, SP, ROA, PTB and INRIM consecutively. The traveling equipment being based on two GPS receivers, it allowed to track any deviation between them during the campaign. We also have emphasized the validation aspects, internally first for consistency check, and externally either against BIPM Circular T data or directly against TWSTFT on the links to OP. We provided the elements used for the computation of the uncertainty budgets, leading to conservative \( k = 2 \) expanded estimates within 2.0 ns for time transfer against the reference receiver in OP. This result, close to the state of the art, could only be obtained thanks to the very good stability of the traveling equipment. A calibration report is in preparation to be forwarded to BIPM.

The fact that such hardware delay calibration activities are close to the 1.0 ns uncertainty limit, together with proven sub-ns calibrated cross-comparisons against a totally independent time transfer technique over three European links [16], call for an exploration of up to now unconsidered biases. In particular, continuous monitoring of direct comparisons between techniques on all possible links between reference time scales should become an operational basis, in order first to assess the claimed time transfer uncertainties, and second to follow closely potential deviations related to hardware delay changes with time.

## Acknowledgments

We acknowledge the support of the following people for the achievement of this GPS receiver relative calibration campaign: M. Abgrall, P. Blondé, O. Chiu and F. Meynadier (OP), E. Staliniere (PTB), R. Costa, I. Sesia and P. Tavella (INRIM), S. Binda and M. Sanchez Gestido (ESA). We are grateful to P. Defraigne (ROB) for having provided freely her TAIP3 processing software. We used International GNSS Service (IGS) products and National Resources Canada (NRCan) Precise Point Positioning (PPP) software for some computations.

## References


Abstract—Since 2010 ROA has supported the coordination of the EURAMET Technical Committee for Time and Frequency (TC-TF) Project 1156, a response from EURAMET TC-TF to Recommendation 2 of CCTF 2009: to study the characterization of GNSS equipment in use for establishing the time links between institutes contributing with their clocks to TAI. Starting that year, a GPS calibration campaign was organized between three contributing laboratories: ROA (Spain), PTB (Germany) and INRIM (Italy). The time transfer results were achieved by using the P3 method, and also carrier phase PPP comparison techniques. These results were also used to re-calibrate the TWSTFT (Two-Way Satellite Time and Frequency Transfer, TW for short) links between labs, with an uncertainty slightly higher than that of the GPS links.

During 2011 and 2012, the campaign was repeated, and in 2012 two other laboratories were included in the calibration trip: NPL (United Kingdom) and OP (France). In this paper we report the calibration results, with a focus on the long term stabilities of the GPS and TW links between the visited labs.

Keywords—GPS Calibration, TWSTFT Calibration, Time Transfer.

I. INTRODUCTION

The GPS time and frequency transfer using code and phase multichannel GNSS receivers, commonly known as geodetic receivers, is among the most useful tools for remote clock comparisons, in particular when receiver clocks are compared using Precise Point Positioning (PPP) software [1]. It is therefore one of the most precise techniques based on satellite measurements, if not the most, since the degradation in precision noted in the TW method due to, among other factors, a bandwidth reduction to 1.7 MHz. However, to provide accurate time transfer by means of a GPS link, it is necessary to carry out calibrations periodically to verify the long term stability of the equipment.

II. EQUIPMENT

The traveling receiver (TR) used in the first two calibrations comprises of a DICOM GTR50 receiver intended for time and frequency transfer, its NovAtel antenna GPS-702-GG with pinwheel technology for multipath rejection and stable phase center, and a 50 m low loss, easy to handle H155 antenna cable.

In the last calibration we used a PolaRxs2 receiver, a Leica AR25 GNSS choke ring antenna and the same type of antenna cable. To avoid the large systematic uncertainty inherent to any Time Interval Counter (TIC), an HP 53132A was included as traveling TIC.
At the beginning of the 2012 calibration campaign, we observed anomalous behaviour of the PolaRx2. But after a thorough examination, we concluded that the problem was caused by an improper selection of the 10 MHz input frequency level. The output frequency level for a frequency standard or a frequency distribution unit is usually around 1.0 Vrms (2.8 Vpp) at 50 Ω load impedance. But the specifications for this receiver indicate a frequency input level from 0.3 Vpp to 1.0 Vpp at 50 Ω.

If instead of the nominal voltage we use a higher one, the receiver shows its normal behavior, but the delay from the 1PPS output signal to the receiver measurement latching can change every time the receiver is restarted, to wit, the internal delay is no longer constant.

In Fig. 1 we show the common-view results of the PolaRx2 vs. a PolaRx3eTR receiver, in common clock setup. During a period of 60 days we have changed the 10 MHz input level from 3.3 Vpp to 0.3 Vpp, at intervals of 0.3 Vpp, only for the PolaRx2 receiver. The results should be centered around zero ns, however incursions up to 20 ns were detected. The receiver worked properly only when the frequency signal level was below 2 Vpp, and of course within the specified values.

Every receiver has its peculiarities, for example, the GTR50 receiver showed a small drift in phase together with a high noise level when the accuracy of the receiver board oscillator was less than 2 x 10⁻⁶. For this reason, a preventive adjustment to the nominal value was carried out before shipment.

III. CALIBRATION PROCEDURE AND RESULTS

The GPS calibration was carried out in differential mode or link calibration [3], where the GPS units involved in each lab (RI) and the TR were operated using common clock (UTC(Lab)) signals over a near zero baseline. The GPS link calibration value (GPSCAL) for a pair of receivers and labs is calculated by the simple difference of common clock difference (CCD) results. In simplified mathematical terms it can be read as follows for ionosphere-free P3 code data (a detailed example illustrated in Fig. 2):

\[
CCD1 = \langle \text{REFGPS}(TR) - \text{REFGPS}(R1) \rangle_{LAB1} \\
CCD2 = \langle \text{REFGPS}(TR) - \text{REFGPS}(R2) \rangle_{LAB2}
\]

\[
GPSCAL_{R1R2} = CCD2 - CCD1,
\]

where the \text{REFGPS} (or reference clock vs. IGST in the case of carrier phase PPP solution) value is the raw receiver reading in the CCTF CGGTTS GPS data, and \langle > is the average over a certain period, normally one day.

As shown above, calibration values are directly obtained in P3 code, instead of standard INTDLY(P1/P2). This is because the P3/PPP time transfer does not need these values, and the definition of the INTDLY is physically not rigorous, and practically unmeasurable in common cases, and the uncertainties of INTDLY(P1/P2) are no better than 3 ns [4]. Nevertheless, the calibration value can be easily obtained if needed by using equation (1) for P1 and P2, and the data collected in the CCTF CGGTTS GPS data files [5]:

\[
\gamma = (77/60)^2 \\
\text{REFGPS}(P1) = \text{REFGPS}(P3) + MSIO \\
\text{REFGPS}(P2) = \text{REFGPS}(P3) + \gamma \times MSIO = \text{REFGPS}(P3) + 1.647 \times MSIO
\]

It is assumed that this calibration value remains valid until any change or event happens in either of the installations, and it can thus be taken into account in calculations of the time scale differences between a pair of labs:

\[
\text{UTC(LAB1)} - \text{UTC(LAB2)} = \text{REFGPS}(R1) - \text{REFGPS}(R2) - GPSCAL_{R1R2}
\]
Once the calibrations started at ROA, the characterized TR was cyclically shipped to participating laboratories for 5-7 days of working operations with the local receivers. Finally the equipment was shipped back to ROA, to carry out the closure measurements with the initial set-up.

It is important to point out that when computing GPSCAL using eq. (1), any systematic error relative to the P3 CGGTTS header values of local receivers R1 and R2 cancels out, as well as the internal and antenna cable delays of TR [3]; the most important contributions to measurement error are those derived from the input time reference delay of the traveling receiver to the UTC(Lab), at both labs, which are the only two values that must be carefully determined for a GPS link calibration. For a PolaRx2 receiver, the input time reference delay is directly linked to the PPS (Pulse Per Second) output signal.

The GPS time link calibration values between all other labs and PTB are summarized in Table I. Only primary GPS links are shown. Annual differences are normally less than 2 ns (at least in PPP), except for the IT link, where an equipment setup change was performed.

### TABLE I.

<table>
<thead>
<tr>
<th>GPS LINK</th>
<th>GPSCAL P3/ns</th>
<th>GPSCAL PPP/ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>RO_6 – PTBB</td>
<td>-3.94 ± 0.95</td>
<td>-3.01 ± 1.00</td>
</tr>
<tr>
<td>RO_6 – PTBB</td>
<td>-1.11 ± 1.28</td>
<td>-1.27 ± 1.34</td>
</tr>
<tr>
<td>RO_6 – PTBB</td>
<td>-1.80 ± 0.79</td>
<td>-1.83 ± 0.80</td>
</tr>
<tr>
<td>IT1Z – PTBB</td>
<td>-5.36 ± 0.95</td>
<td>-5.25 ± 0.97</td>
</tr>
<tr>
<td>IT1Z – PTBB</td>
<td>-4.06 ± 1.29</td>
<td>-4.47 ± 1.36</td>
</tr>
<tr>
<td>IT1Z – PTBB</td>
<td>1.21 ± 0.82</td>
<td>1.39 ± 0.83</td>
</tr>
<tr>
<td>OP02 – PTBB</td>
<td>3.60 ± 0.80</td>
<td>3.54 ± 0.82</td>
</tr>
<tr>
<td>NP11 – PTBB</td>
<td>3.88 ± 0.79</td>
<td>3.88 ± 0.80</td>
</tr>
</tbody>
</table>

The receiver types are: RO_6 (PolaRx3eTR), PTBB, IT1Z and OP02 (ASHTECH Z-XII3T), NP11 (GTR50).

### IV. UNCERTAINTY EVALUATION

The overall uncertainty of the calibration value is estimated from the following expression:

\[ U_{GPS} = \sqrt{u_{A,1}^2 + u_{B,1}^2 + u_{B,2}^2 + u_{B,3}^2 + u_{B,4}^2 + u_{B,5}^2 + u_{B,6}^2} \]

where \( u_{A,1} \) reflects the statistical uncertainty of the CCD determination at the first lab, and \( u_{A,2} \) reflects the statistical uncertainty of the CCD measurements at the second lab. We have directly used the daily average SD, due to white phase noise exhibited up to an averaging interval of about one day [6].

The systematic uncertainty \( u_{B,1} \) represents the uncertainty of the 1PPS delay from the local reference time, connected to each pair of receivers at the first lab (and \( u_{B,2} \) the similar uncertainty at the second lab). The systematic contribution due to the use of an uncalibrated time interval counter cancels out, when a traveling TIC has been used at both labs, and using the same measurement procedure. We account from 0.2 to 0.51 ns for this contribution.

In \( u_{B,3} \) we have included the instability of the receivers and antennae. Environmental effects like humidity and especially temperature have been shown to be significant. We account 0.4 ns for this contribution [7].

In \( u_{B,4} \) we account for signal propagation effects, which mostly cancel in the chosen quasi zero baseline configuration. Only multipath errors might cause a small contribution to the uncertainty (0.3 ns for P3 [8] and 0.1 for PPP).

In uncertainty \( u_{B,5} \) we account for the closure measurement value, which is usually below 0.5 ns, but sometimes can exceed 1 ns.

Finally, \( u_{B,6} \) represents the uncertainty of the ambiguity estimation in the PPP processing, which can be taken as 0.3 ns [9].

The resulting final calibration uncertainty for a GPS link is typically from 0.8 to 2 ns.

### V. TW LINK CALIBRATION

Taking the values obtained previously, we used the GPS time transfer results to re-calibrate the TW links with PTB, using a 20-day data set for each lab, partly overlapping with the GPS calibration trip schedule.

We have initially derived the DCD (Double Clock Differences) between the GPS and TW links (TWCAL), as shown in Fig. 3. Each data point is the result of the exact difference of a TW value and the interpolation of the adjacent P3 values. In the case of PPP, calculated by NRCan PPP v. 1087 software, the interpolation is not necessary and the RINEX files have been processed at 1-day batch mode.

In this figure a more precise solution for PPP data can be noted, thanks to the incorporation of phase data with the code measurements, as well as a better estimation of atmospheric delay. PPP processing also applies special models for site displacement (e.g. Solid Earth Tides, Ocean Loading) and satellite attitude effects (e.g. Satellite Antenna offsets, Phase Wind-Up Corrections).
The uncertainty involved in this calculation is estimated from the following expression:

$$U_{TW} = \sqrt{u_{A,3}^2 + u_{B,7}^2 + u_{B,8}^2}$$

where $u_{A,3}$ reflects the statistical uncertainty of TWCAL values (< 0.4 ns).

In $u_{B,7}$ we have estimated the additional overall instability of the equipment involved in the calibration of the TW link: distribution of local UTC signals, the TW station components’ instabilities and environmental effects (0.3 ns).

In $u_{B,8}$ we have included the calculated uncertainty of the GPS link.

The final calibration uncertainty for a TW link is therefore between 1 and 2 ns. An independent study [4] using a completely different method gives a similar estimation.

VI. LESSONS LEARNED AND ENHANCEMENT OPTIONS

As stated in the analysis of uncertainty, one of the most important contributions to the uncertainty budget was due to the large systematic uncertainty inherent to any TIC, which was avoided by including one as a traveling TIC.

Nevertheless, the trigger level selected to measure the input PPS delay to local UTC is normally fixed at 1 V (or alternatively 0.5 V). But as illustrated in Fig. 4 it should be selected as a function of the source. Thus, 1 V is suitable for a Cs 5071A, but 0.5 V can be more appropriate for a Timetech PPS distribution unit.

This difficulty could be solved by adding to the traveling set a PPS distribution amplifier and a frequency distribution amplifier, both with constant output levels independently of the source input level. In this way, we could reproduce the same operating conditions in both laboratories, and we could solve, among others, the problem of providing an erroneous input frequency level to the PolaRx2.

Nevertheless the impact of using PPP or P3 data is normally less than 0.3 ns. Taking this into account, the TWCAL corrections (summarized in Table II), can be calculated by:

$$TWCAL = [TW(LAB1) - TW(LAB2)] - <PPP(R1) - PPP(R2) - GPSCAL_{R1R2}>$$

The time scale differences between a pair of labs have to be corrected according to:

$$UTC(LAB1)-UTC(LAB2)=TW(LAB1)-TW(LAB2)-TWCAL$$

<table>
<thead>
<tr>
<th>LINK</th>
<th>TWCAL PPP/ns</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROA – PTB10</td>
<td>-4.0 ± 1.4</td>
</tr>
<tr>
<td>ROA – PTB11</td>
<td>-4.4 ± 1.5</td>
</tr>
<tr>
<td>ROA – PTB12</td>
<td>-5.1 ± 0.9</td>
</tr>
<tr>
<td>INRIM – PTB10</td>
<td>-1.7 ± 1.4</td>
</tr>
<tr>
<td>INRIM – PTB11</td>
<td>-0.5 ± 1.5</td>
</tr>
<tr>
<td>INRIM – PTB12</td>
<td>-3.7 ± 0.9</td>
</tr>
<tr>
<td>OP – PTB12</td>
<td>-3.1 ± 0.9</td>
</tr>
<tr>
<td>NPL – PTB12</td>
<td>-0.1 ± 0.9</td>
</tr>
</tbody>
</table>

TABLE II. CALIBRATION AND UNCERTAINTY VALUES (TWCAL ± UTW) FOR TW LINKS WITH PTB, FROM 2010 TO 2012.

![Fig. 3. DCD of TW ROA-PTB and PPP/P3 (RO_6-PTBB) results.](image)

![Fig. 4. PPS output from three different devices.](image)
Currently the state of the art has probably been achieved with the BIPM GPS traveling calibrator (StdB), used for the METODE pilot project [10]. It is composed of 2 independent GNSS receivers (of different types), 2 antennas, 2 antenna cables, 1 TIC and 2 signal distribution devices. The only improvement would consist of replacing the frequency distributor by a constant output one, e.g. Symmetricom 58502A, thus the frequency attenuation could be fixed at the start of the campaign.

The calibration procedure to be followed and the operations of the traveling calibrator must be as simple as possible, to minimize the risk of mistakes.

Once a GPS link is calibrated, the recalibration of the TW link should be carried out with the PPP technique, especially for intercontinental baselines.

The calibrations were performed during the same time of the year, in order to avoid seasonal environmental effects. Nevertheless, the mechanism of the so-called long-term variation (understood as an instability or discontinuity, not periodic like the seasonal temperature variation but rather a one-way change) in the GNSS receivers or the TW ground stations is not clear yet [11].

The TR’s long term stability is not a critical issue, and the only requisite is that the equipment must be stable during the calibration tour. Since 2008 several GPS calibrations have been carried out, and although we have seen a number of problems such as water inside antennas, too sharply bent antenna cables, unexpected jumps in GPS receivers, drifts in CV results, corrosion in antenna cables and connectors, damaged shipment, etc., in general, a GPS calibration is an easy task. Nevertheless there are many factors that cannot be controlled, as well as the setup by local staff.

To verify possible inconsistencies, it would be very interesting to perform in parallel TW time link calibrations with a TW mobile station and with GPS traveling receiver. Nevertheless the only results documented are a bit dispiriting, e.g. the GPS parallel calibrations conducted by USNO with differences of about 7 ns [12].

An important issue to ensure the stability of both the TW link and the GPS link is to periodically verify the DCD series, in order to detect problems in the TW link or in the GPS one. Since the 2012 calibration, we have maintained a close monitoring of the DCD of the ROA-PTB links. As is shown in Fig. 5, during 2013 the changes of these values have remained below 1 ns, nevertheless during 2014 we have observed drifts larger than 2 ns. During the latest summer calibrations (one TW with a portable station and two GPS), the results were as expected from the DCD. The only possibility of disagreement is that both links were changing in the same way, but this point can be efficiently solved with multiple independent redundant time transfer GNSS systems.

This project has been stopped since 2012, but currently we are taking the appropriate steps to reactivate it, once we have solved some equipment limitations. The intention during this year is to extend the calibration to any other institute participating in the Project.

![DCD of TW ROA-PTB and PPP (RO_6-PTBB) links, from 2012 GPS calibration to 2014 TW/GPS calibration.](image)

**VII. CONCLUSIONS**

In this report we have summarized the GPS calibration trip experience between ROA, INRIM, PTB, OP and NPL, using a portable GPS receiver and a Time Interval Counter.

The estimated uncertainty indicates the real possibility of subnanosecond accuracy calibration of GPS links, which could be achieved by carrying out the improvements outlined in the last paragraph, at least for short/medium periods of time, neglecting possible long term delay changes in the GPS equipment. The uncertainty of a calibration increases with time, as an aging function, so frequent calibrations are the only way to ensure best performance.

The uncertainty associated with this kind of TW link calibration is mainly dominated by the GPS uncertainty, but for this link the estimated uncertainty is still below the nanosecond level. We plan to continue verifying these results by performing calibration campaigns annually.

**ACKNOWLEDGMENT**

The authors acknowledge the Geodetic Survey Division, Natural Resources Canada (NRCan), for providing the PPP software.

**DISCLAIMER**

GPS receivers and time and frequency devices are mentioned in this paper. The authors do not endorse any product or manufacturer.

**REFERENCES**


Comparison of Two Continuous GPS Carrier-Phase Time Transfer Techniques

Jian Yao¹, Ivan Skakun², Zhiheng Jiang³, and Judah Levine¹

¹ Time and Frequency Division and JILA, National Institute of Standards and Technology and University of Colorado, Boulder, CO, USA
² PNT Information and Analysis Center, Central Research Institute of Machine Building, Korolyov City, Russia
³ Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92312 Sèvres Cedex, France

E-mail: jian.yao@nist.gov

Abstract—Global Positioning System (GPS) carrier-phase (CP) time transfer, as a widely accepted high-precision time transfer method, frequently shows a data-batch boundary discontinuity of up to 1 ns, because of the inconsistency of the phase ambiguities between two consecutive data batches. To eliminate the data-batch boundary discontinuity, several techniques have been proposed in recent years. The question is how large the solutions of these techniques differ from each other and how well the solutions are faithful to clocks. To answer these questions, this paper chooses two techniques to study: Revised RINEX-Shift (RRS) technique [1-2], and Phase Common-View (Phase-CV) technique [3-4]. This paper shows that the time deviation of the difference between the two techniques is below 100 ps, for an averaging time of less than 10 days. Especially, for an averaging time of less than 1 day, the time deviation is less than 30 ps. We also find that both RRS and Phase-CV match TWSTFT (two-way satellite time and frequency transfer) and TWOTFF (two-way optical-fiber time and frequency transfer) quite well. The difference is typically within ±0.3 ns for more than 20 days. The above results are all based on a short-distance links (less than 2500 km). A long-distance comparison between these two techniques, such as a transatlantic link, has not yet been investigated.

Keywords: GPS, Carrier-phase time transfer, boundary discontinuity, Revised RINEX-Shift, Phase common-view, two-way satellite time transfer, two-way optical-fiber time transfer

I. INTRODUCTION

Global Positioning System (GPS) carrier-phase (CP) time transfer is a widely accepted high-precision time transfer method. This method provides much lower short-term noise than other time transfer methods, such as TWSTFT (two-way satellite time and frequency transfer) and GPS common-view (CV) time transfer. TWOTFT (two-way optical-fiber time and frequency transfer), as an emerging time transfer method, can potentially be more precise than GPS CP. However, a long-distance performance of TWOTFT, such as a transatlantic link, is unknown. Besides, it requires a lot of infrastructure work and is also expensive to maintain. Thus, GPS CP is and will continue to be one of the mainstream time transfer methods.

Along with the development of GPS CP time transfer method, the problem of data-batch boundary discontinuity attracts a lot of attention. The boundary discontinuity can quite seriously affect the long-term (e.g., > 1 day) time-transfer result. Studies show that the boundary discontinuity comes from the uncertainty in the phase-ambiguity estimation for each data batch. Fundamentally, this uncertainty further comes from the code noise, because the code measurements are used to estimate the phase ambiguity [5].

To solve the boundary-discontinuity problem, several techniques have been proposed in recent years [1-4, 6-7]. Each technique looks perfect on paper. However, there is little study on the comparison between these techniques. Thus, we have no idea of how large the results of these techniques differ from each other and how well the results are faithful to clocks. This paper focuses on answering these questions. Here, two techniques are chosen for comparison: Revised RINEX-Shift (RRS) technique [1-2], and Phase Common-View (Phase-CV) technique [3-4]. Section II provides the basic principles of these two techniques. The advantages and disadvantages of each technique are also discussed in this section. Section III compares their performance for baselines of 600 km – 2500 km, with TWSTFT as a reference. A three-station closure of Phase-CV is also done, to check its self-consistency. Section IV compares the results of these two techniques with a TWOTFT result for a baseline of ~268 km. These comparisons make us conclude that both techniques work well for a baseline of no longer than 2500 km.

II. PRINCIPLES OF RRS AND PHASE-CV

The RRS is actually an updated version of PPP (precise point positioning). It runs PPP for a data batch of multi-days (here, we choose 10 days) and extracts the middle epoch. Then it shifts the data batch by a small time step (here, we choose 10 min), runs PPP, and extracts the new middle epoch. It does the data-batch shift by 10 min again and again. The solutions at all middle epochs form the RRS result [2]. Here, we should mention, if there is a GPS data anomaly, a program is run to repair the anomaly and the RRS program uses the repaired GPS data [8]. Previous study has shown that the RRS technique can achieve the 10⁻¹⁷ level of instability for an
averaging time of 20 days with TWSTFT as a reference, while the conventional 30-days PPP processing is still $\sim 2 \times 10^{-16}$ for the same averaging time [2].

Phase-CV is similar to the traditional GPS CV time transfer, but using the phase data rather than the code data. Phase-CV is achieved by two steps. First, it uses PPP to estimate the absolute station coordinates and tropospheric zenith delays (TZD). Second, it does the single-difference of phase measurements between two stations, for the same GPS satellite. The single-difference recovers the integer property of the phase ambiguities. By using the coordinates and TZDs in the first step, we can resolve the integer ambiguities and clock difference between the two stations [3].

Before we study the technical performance of each technique in the later sections, we here want to address the advantages and disadvantages of each technique. These issues are often ignored, but they can sometimes be even more important than the pure technical performance.

First, RRS requires only a single station, while Phase-CV requires two stations. So RRS is still a type of PPP, while Phase-CV is not.

Second, RRS works for any baseline, short or long, since RRS does a time comparison between local time and the IGS time. The long-baseline performance of RRS (between NIST and PTB), with respect to TWSTFT, has been shown by Figure 11 of [2]. Phase-CV typically works worse as the baseline increases, because of few common-view GPS satellites and no common path. The network processing of Phase-CV, which is still under development, may help the long-baseline performance of Phase-CV.

Third, the solution of RRS is unique, no matter what the start date and the end date are. However, the solution of Phase-CV is not unique. First, it depends on the absolute station position, which can vary by ~1 cm when different GPS data batches are used. Thus, different people may use slightly different positions for Phase-CV. A slightly incorrect absolute position can lead to a small slope in the Phase-CV solution. The absolute position may also change as time passes. Second, Phase-CV is for frequency transfer. In order to achieve time transfer, we need to align the Phase-CV solution with the PPP solution on a long time interval (e.g., >10 days). However, the choice of a long time interval is arbitrary. Different long time intervals (e.g., MJD 56000.0 – 56000.25, and another core is used to compute MJD 56000.25–56000.50). Phase-CV requires more computation than PPP, but the increase is not big. Phase-CV is a sequential process.

Sixth, Phase-CV can work in real time or near real time, while RRS cannot. RRS has a latency of 5 days.

Lastly, Phase-CV sometimes cannot keep the integer ambiguity property, which leads to a re-initialization of the processing settings, while RRS does not have this problem.

III. Comparison between RRS and Phase-CV

In this section, we compare RRS and Phase-CV, for baselines of 600 – 2500 km. We will see that they agree fairly well.

PTBB is a GPS receiver at PTB (Physikalisch-Technische Bundesanstalt), Germany. The coordinates of this receiver are $X = 3844060.1$ m, $Y = 709661.2$ m, and $Z = 5023129.5$ m, in the ITRF (international terrestrial reference system) coordinate system. The reference time of PTBB is UTC(PTB) with a constant delay. OPMT is a GPS receiver at OP (Paris Observatory), France, with the coordinates of $X = 4202777.4$ m, $Y = 171368.0$ m, and $Z = 4778660.2$ m. The reference time of OPMT comes from a hydrogen maser, which usually has a non-zero slope. We should mention that the TWSTFT facilities at both PTB and OP share the same reference times as the GPS receivers. The baseline of the link of “OPMT-PTBB” is ~692 km.

We do the time comparison between OPMT and PTBB using RRS, Phase-CV, and TWSTFT, for MJD (Modified Julian Date) $56881.0 – 56905.0$ (Figure 1). Note, the slope from the hydrogen maser at OP has already been removed and some constant offsets are added to the three curves to overlap each other. Here, we should emphasize that we use exactly the same GPS data of OPMT and PTBB for both RRS and Phase-CV. From Figure 1, we can see that both RRS and Phase-CV provide continuous solutions. They match each other very well. They also match the TWSTFT result quite well, although there is an approximately 0.5 ns discrepancy during MJD 56887-56895. This discrepancy could come from either GPS time transfer or TWSTFT or both [9].

To investigate the agreement between RRS and Phase-CV, we do double-difference between RRS and Phase-CV for the link of “OPMT – PTBB” (Figure 2). The difference is within ±200 ps. This indicates a good match between the two techniques. Modified total deviation (Figure 3) and time total deviation (Figure 4) reveal the frequency stability of the double difference between RRS and Phase-CV. From Figure 3, we can see that the two techniques match with a fractional uncertainty of $\sim 5 \times 10^{-16}$ for an averaging time of 1 day, and $\sim 2 \times 10^{-16}$ for an averaging time of 10 days. Figure 4 shows that the time deviation of the double difference is below 100 ps for an averaging time of less than 10 days. Especially, the time deviation is less than 30 ps within 1 day. This indicates that even though we process the GPS code and phase data using
two different techniques, the time-transfer results are consistent with each other. This validates both techniques.

![Figure 1. Time comparison between OPMT and PTBB, using RRS, Phase-CV, and TWSTFT. Note, TWSTFT facilities at both OP and PTB share the same reference times as GPS receivers. Slope from the hydrogen maser at OPMT has been removed, and some constant offsets are added to the three curves to overlap each other.](image1)

To further verify the above conclusion that both techniques match each other very well, we also compute the double difference between the two techniques for other baselines.

MDVJ is a GPS receiver in Mendeleevo, Russia, with the coordinates of X = 2845456.3 m, Y = 2160954.3 m, and Z = 5265993.4 m. The baseline between PTBB and MDVJ is approximately 1778 km. And the baseline between OPMT and MDVJ is approximately 2457 km. The double differences between RRS and Phase-CV for these two baselines are shown in Figure 5 and Figure 6, respectively. Again, we can see that RRS is within approximately ±200 ps of Phase-CV. Here, we should mention that the Phase-CV has an average offset of about +0.35 ns for the link of “OPMT-MDVJ”. This constant offset leads to the curve in Figure 6 shifting down by 0.35 ns. The reason for the offset comes from the ambiguity of the absolute time in Phase-CV. Phase-CV itself can only provide the frequency transfer result. To provide the time transfer result, it requires the assistance of the conventional PPP solution. However, the boundary discontinuity in the conventional PPP can lead to a slightly biased time transfer result. That is why Phase-CV is 0.35 ns biased from RRS in Figure 6.

![Figure 2. Double difference between RRS and Phase-CV during 56881.0 – 56905.0, for the link of “OPMT-PTBB.”](image2)

![Figure 3. Modified total deviation for the double difference between RRS and Phase-CV, for the link of “OPMT-PTBB.”](image3)

![Figure 4. Time total deviation for the double difference between RRS and Phase-CV, for the link of “OPMT-PTBB.”](image4)
From the above discussion, we know that RRS and Phase-CV agree within ±200 ps. Now that we have done the time transfer between each two of the three stations, a three-station closure may tell us the self-consistency of a time transfer technique. Since RRS is a type of single-point technique, the time difference between two stations is achieved by introducing a common reference time. Often, we choose the IGS (international GNSS service) time (IGST) as the common reference time. Then, the three-station closure of RRS becomes

\[
\text{Closure} = (\text{PTBB} - \text{MDVJ}) + (\text{MDVJ} - \text{OPMT}) + (\text{OPMT} - \text{PTBB})
\]

(1)

Equation (1) indicates that the three-station closure of RRS is always exactly 0. The red curve in Figure 7 further confirms this conclusion. However, the Phase-CV is a type of common-view technique. It provides the time difference between two stations directly and no common reference time needs to be introduced in the Phase-CV. Thus, the three-station closure of Phase-CV is

\[
\text{Closure} = (\text{PTBB} - \text{MDVJ}) + (\text{MDVJ} - \text{OPMT}) + (\text{OPMT} - \text{PTBB}).
\]

(2)

Equation (2) cannot be further simplified. Thus, the closure of Phase-CV is not necessary to be exactly 0. The closure test for Phase-CV can show how well it is self-consistent. The black curve in Figure 7 shows the result of the Phase-CV three-station closure test. We can see that the closure is not around 0 ns. Instead, it is shifted by approximately -0.37 ns. As mentioned before, this offset comes from the ambiguity of the absolute time in Phase-CV. From Figure 7, we know that the peak-to-peak value of the closure is as small as ~60 ps. Besides, the closure does not change over time. These indicate that the Phase-CV processing is self-consistent for frequency transfer.

IV. COMPARISON OF RRS AND PHASE-CV WITH TWOTFT

TWOTFT is a fast-emerging time transfer technique. Many people have demonstrated its ultra-precise time transfer capability [10-12]. Thus, a comparison between GPS and TWOTFT can provide the instability of GPS time transfer, because the instability of TWOTFT is typically smaller or even negligible when compared to GPS.

There is an optical fiber link between AOS (Astrogeodynamical Observatory) and PL (Polish Atomic Time Scale) in Poland [13]. The length of the optical fiber is ~420 km. There are also two GPS receivers, i.e., AO_4 and GUM4, at AOS and PL, respectively. The coordinates of AO_4 are X = 3738358.4 m, Y = 1148173.7 m, and Z = 5021815.8 m. The coordinates of GUM4 are X = 3653847.0 m, Y = 1402629.2 m, and Z = 5019465.1 m. Thus, the baseline
between these two stations is approximately 268 km. The time references for the optical fiber link and the GPS receivers are the same at each station.

Figure 8 shows the time difference between AOS and PL using TWOTFT, RRS, and Phase-CV, for MJD 56902.0 – 56928.0. We make the three curves match at MJD 56928.0 for a better comparison. The TWOTFT result (blue curve) is hard to see in Figure 8, because it is almost completely covered by the red/black curve. This indicates that both RRS and Phase-CV agree with TWOTFT well over the entire 26 days.

The TWOTFT result (blue curve) is hard to see in Figure 8, because it is almost completely covered by the red/black curve. This indicates that both RRS and Phase-CV match TWOTFT very well.

To show the difference between GPS time transfer and TWOTFT, we do double difference between RRS/Phase-CV and TWOTFT (Figure 9). The BIPM 35-days PP (i.e., TAI PPP 35 days) result [14] is also provided in Figure 9, as a reference. There are two anomaly points at 56909.35 and 56921.81. The BIPM TAI PPP shows two jumps at both anomaly points. The jumps are 0.7 ns and 0.4 ns, respectively. Because of the jumps, the trend is changed significantly. For example, the time change from 56909 to 56922 is approximately 0.9 ns, which significantly affects the time-comparison result. In contrast, the RRS technique (red curve) performs very well at both anomaly points. It remains flat (within ± 100 ps) compared to TWOTFT, during 56903 – 56915. There is also no significant change around the second anomaly point (i.e., during 56921.5 – 56922.5). Over the whole 26 days, the difference between RRS and TWOTFT is less than ±250 ps. This indicates the correctness of RRS. The Phase-CV (black curve) does not do well at the first anomaly point. It initializes the filter and thus is very noisy during the whole day of MJD 56909. Actually, there was also a jump of about −1 ns on MJD 56909 in the original Phase-CV result, because we need to re-estimate the absolute time using PPP when a re-initialization occurs. We have already removed this jump in Figure 9. There was also a jump at the second anomaly point in the original Phase-CV result. We again removed the jump by a simple concatenation. From the black curve, we can see that the difference between Phase-CV and TWOTFT is also less than ±250 ps. Its slope is pretty small and is not affected by the jumps and the anomaly points. Especially, it keeps flat during 56917 – 56920, while there is a small dent in RRS. The reason why Phase-CV is so flat probably comes from the fact that Phase-CV uses phase only and thus the noise in code is well excluded. From the above analysis, Phase-CV is good for the frequency transfer. For the time-transfer purpose, a careful calibration or adjustment at each re-initialization point is required in Phase-CV. Next, let’s consider the long-term trend of the three curves in Figure 9. We can see that RRS and BIPM TAI PPP goes down by ~100 ps during the 26 days, while Phase-CV goes up by ~300 ps. The increase in Phase-CV is probably because station coordinates were not estimated in the same filter and was fixed for the whole 26 days. Note that the three GPS carrier-phase techniques use the same GPS data, but, unfortunately, the long-term trends are different. This indicates that different GPS CP techniques introduce different long-term trends. And it is hard to tell which technique is more correct. In this case, the long-term difference between RRS and Phase-CV is ~ 400 ps for 26 days, which matches our conclusion in Section III that the difference between RRS and Phase-CV is within ± 200 ps.

To study the frequency stabilities of RRS, Phase-CV, and BIPM TAI PPP, with respect to TWOTFT, we compute the modified total deviation of the double difference (Figure 10). Note, we have already removed the bad data of Phase-CV on MJD 56909. We can see that Phase-CV provides the smallest instability. RRS is better than BIPM TAI PPP after ~ 6 hours. Both RRS and Phase-CV provide ~1 × 10^{-16} level of instability after 5 days. The above results are only based on the fact that
the baseline is ~268 km. For a transatlantic link, the RRS performance has little change (see Figure 4.16 in [15]). However, the Phase-CV performance typically gets worse, if three bridge stations are introduced. We add the four short baselines (< 2000 km) together to achieve the transatlantic time transfer. Thus, the Phase-CV instability for the transatlantic link is increased to double of the instability for a short baseline ($\sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2 \times 2$). This theoretical frequency instability for a long-distance link is shown by the blue dotted curve in Figure 10. We can see that RRS becomes the best among RRS, Phase-CV, and BIPM TAIPPP for the case of a transatlantic link.

The three curves in Figure 10 can also be used to set the upper limit of the frequency instability of the time transfer techniques. For example, the upper limit of the RRS instability is $5 \times 10^{-15}$ at 3 hours, $9 \times 10^{-16}$ at 1 day, and $2 \times 10^{-16}$ at 5 days. The upper limit of Phase-CV instability (for ~268 km baseline) is similar to RRS instability, but with a significant improvement at 1 day (i.e., $6 \times 10^{-16}$).

Admittedly, both RRS and Phase-CV are still under development and they can be further improved. Nevertheless, even without any further improvement, both techniques are already better than the BIPM TAIPPP, based on the above comparison with TWOTFT.

In order to improve the performance of RRS, we adjust the weights of code and phase in RRS. The RRS is actually a phase time transfer technique with a long-term steering (e.g., > 1 day) to the code data. Since the code data are noisier than the phase data, we decrease the weight of code in RRS so that the long-term steering is not overreacting. For example, we change the weight ratio of code to phase from the default 1:10000 to 1:40000. We find that this change makes the dent during 56917 – 56920 and also other oscillations in the red curve in Figure 9 become smaller. Figure 11 shows the RRS result with the improvement of code&phase weights. In terms of frequency stability, there is an obvious improvement for the averaging time of ~1 day (see Figure 12). Now, the upper limit of the RRS instability becomes $7 \times 10^{-16}$ at 1 day.

![Figure 11. Double differences of “RRS with Improvement – TWOTFT” (red), for the link between AOS and PL. The black curve is the same as Figure 9. It is plotted in this figure as a reference.](image)

In Figure 12, we can see that the performance of RRS with improvement in code and phase weights (red curve) becomes better than the BIPM TAIPPP (green curve). The black curve is the same as Figure 10 and is plotted in this figure as a reference.

![Figure 12. Performance of RRS with improvement in code and phase weights (red curve), for the link between AOS and PL. The black and green curves are the same as Figure 10. They are plotted in this figure as a reference.](image)
V. CONCLUSIONS

In this paper, we compare two continuous GPS carrier-phase time transfer techniques: Revised RINEX-Shift (RRS) technique, and Phase Common-View (Phase-CV) technique. The time difference between these two techniques is typically within ±200 ps for baselines of less than 2500 km. This indicates a good agreement between the two techniques.

The double difference between these two techniques and other independent time transfer techniques, such as TWSTFT and TWOTFT, can reveal how well the two continuous solutions are faithful to clocks. We find that both RRS and Phase-CV match the long-term trend of TWSTFT quite well. However, RRS and Phase-CV can sometimes walk ~0.5 ns away from TWSTFT. This can come from either TWSTFT or GPS, or even both. Compared with a two-way optical fiber link with a ~268 km baseline, both RRS and Phase-CV vary less than ±250 ps during 26 days. This comparison confirms the correctness of both techniques. We find that Phase-CV can provide a slightly better frequency transfer result than RRS for the averaging time of around 1 day. However, this is only for the case of baseline = 268 km. Its long-distance (e.g., a transatlantic link) performance is unknown (typically worse with bridge stations introduced) and hard to verify, because of no such fiber link. However, a network processing of Phase-CV, which is still under development, may help the long-distance performance. The ambiguity of the absolute time and the problem of re-initialization in the Phase-CV solution also need to be solved, if the time transfer, instead of the frequency transfer, is our main concern. Our study also shows that the conventional BIPM TAIPPP can have an incorrect time-transfer slope due to the data-batch boundary discontinuity. With the advent of RRS and Phase-CV, the GPS time transfer becomes more faithful to clocks and thus can observe a remote clock behavior better.

Acknowledgements

The authors thank Francois Lahaye for sharing the NRCan PPP software. Jerzy Nawrocki at AOS, Poland and Albin Czubla at PL, Poland are specially thanked for providing the optical-fiber-link data and the GPS data. We also thank those people who maintain the GPS receivers of PTBB, OPMT, MDVJ, AO_4, and GUM4. IGS is gratefully acknowledged for providing GPS tracking data, station coordinates, and satellite ephemerides.

REFERENCES

Correction for Code-Phase Clock Bias in PPP

Pascale Defraigne
Royal Observatory of Belgium
Brussels, Belgium
p.defraigne@oma.be

Jean-Marie Sleewaegen
Septentrio
Leuven, Belgium,
sleewae@septentrio.com

Abstract—Precise Point Positioning (PPP) is a zero-difference single-station technique that has proved to be very effective for time and frequency transfer, enabling the comparison of atomic clocks with a precision of a hundred picoseconds and a one day stability below the 1e-15 level. It was however noted that for some receivers, a frequency difference is observed between the clock solution based on the code measurements and the clock solution based on the carrier phase measurements. These observations reveal some inconsistency between the code and carrier phases measured by the receiver. One explanation of this discrepancy is the time offset that can exist for some receivers between the code and carrier phase latching. This paper explains how a code-phase bias in the receiver hardware can induce a frequency difference between the code and the carrier phase clock solutions. The impact on PPP is then quantified. Finally, the possibility to determine this code-phase bias in the PPP modeling is investigated, and the first results are presented.

Keywords—GNSS; PPP; carrier phase; code;

I. INTRODUCTION

Precise Point Positioning (PPP) is a zero-difference single-station technique based on the joint analysis of dual-frequency ionosphere-free combinations of codes and carrier phases measured in one station, to determine its position and its clock synchronization error at each observation epoch [1]. PPP has proved to be very effective for time and frequency transfer, enabling the comparison of atomic clocks with a precision of a hundred picoseconds and a one day stability below the 1e-15 level [2]. It was however noted that for some receivers, a frequency difference is observed between the clock solution based on the code measurements and the clock solution based on the carrier phase measurements. This is reflected by a non-zero average of the discontinuities between successive batch solutions [3], and by systematic effects between the code and carrier phase residuals of the PPP solutions [4]. These observations reveal some inconsistency between the code and carrier phases measured by the receiver.

One explanation of this discrepancy is the time offset that can exist for some receivers between the code and carrier phase latching. Before looking at an offset between code and carrier phase measurements, we first analyze the effect of a time offset between the measurements collected by two receivers connected to a same antenna and a same atomic clock. The receiver clock offset between the two units is in the present case about 201 µs. The differences between the code measurements made by the two receivers, and between the phase measurements, are plotted respectively in Figures 1 and 2. These differences have been presented using a different color for each satellite track. Note that in Figure 1, some ns biases have been added between the difference tracks to improve the visibility, and in Figure 2 all the carrier phase differences have been corrected for the ambiguities, removing from each difference the average of the differences inside its track.

Apart from the bias in the pseudorange measurements, we additionally observe a positive slope inside all the tracks. In the carrier phase data, thanks to the lower noise level, we can see that this drift in fact has the same S-like shape for all the satellite.
Fig 2. Carrier phase differences between the two receivers desynchronized by 201 µs, plotted for each satellite track separately.

The explanation can be found in the variable radial distance between the satellite and the receiver during the track. If receiver A takes its measurements at \( t_0 \) while receiver B takes its measurements at \( t_0' = t_0 + \tau \), then the emission times of the received signals in \( t_0 \) and \( t_0' \) are \( t_e \) and \( t_e' \). The measured pseudorange is the radial distance \( R \) for receiver A and \( R' \) for receiver B (plus a clock bias term constant for all satellites). As illustrated in Figure 3, \( (R'-R) \) is going from negative to positive values during the satellite track, which explains the differences depicted in Figures 1 and 2. This variation of radial distance is directly linked with the Doppler frequency \( f_D \) which is the basis of the carrier phase measurement:

\[
0 = \frac{c}{R} f_D (R-R) - \phi_0
\]

where \( c \) is the velocity of light, and \( f_0 \) the emitted frequency (L1 or L2 for GPS).

A synchronization offset between two receivers will however not provide any drift in the difference between their clock solutions as the software will determine the emission time after having computed a first estimation of the receiver clock offset in the following procedure:

1. A first estimation of the emission time:
   \[
   (t_e')_1 = t_0' - \frac{P}{c}
   \]

   Where \( P \) is the pseudorange, and \( t_0' \) is the receiving time in the receiver time scale, and the Pseudorange is measured containing the desynchronization between the receiver and the satellite clock.

2. To get the emission time in the GNSS time scale, the satellite position \( X_s(t_e') \) is used to estimate the receiver clock with respect to the GNSS time scale:
   \[
   \Delta t_{rec} = \frac{P}{c} - ||X_s(t_e') - X_{rec}|| + \Delta t_{sat}
   \]

3. The emission time in the GNSS time scale can then be obtained as:
   \[
   (t_e'_{GNSS}) = t_0' - ||X_s(t_e') - X_{rec}|| - \Delta t_{rec}
   \]

4. Obtain the correct position \( X_s(t_{e'_{GNSS}}) \).

The impact of the receiver clock offset on the computed satellite position is therefore corrected for and the associated Doppler increment present in the pseudoranges is absorbed.

B. Unsynchronous Code And Carrier Phase Measurements

Consider now a receiver in which the phase measurements are made after the code measurements with delay \( \tau \). In this case, the differences between the code and carrier phase data will contain a term constant for all satellites (the clock bias term) and a term that is dependent on the Doppler and that is therefore satellite-dependent. More specifically, in case of a delay \( \tau \) between the latching of the code and the phase measurements, the difference between code and phase measurements for a satellite \( i \), in cycles, is given by the following formula:

\[
\frac{d\phi_i}{f_L} = \tau * f_L + \tau * f_{D,i}
\]

where \( f_L \) is the nominal carrier frequency (e.g. 1575.42MHz for GPS L1) and \( f_{D,i} \) is the Doppler of satellite \( i \), in Hz. Note that only the fractional part of \( d\phi_i \) is relevant as the integer part is absorbed by the carrier ambiguity.

This difference will however not be absorbed in the data analysis, because the satellite emission time is only determined from the code measurements, the carrier phase data being ambiguous. The same emission time, and hence the same satellite position, is therefore used for code and carrier phase data, and the Doppler-dependent term in the phase measurements due to the non synchronous measurement latching is not corrected for.

In order to illustrate this, we simulated the Doppler increment associated with a given delay between the code and carrier phase latching. For this, we used a 30-second RINEX observation file from the receiver PTBB, connected to a H-maser. We determined the Doppler frequency from the carrier phase measurements:

\[
f_D(t,sat) = \frac{\phi(t + 30s,sat) - \phi(t,sat)}{30}
\]

The Doppler frequencies so-estimated for several satellite tracks are plotted in Figure 4. The same shape is observed as in Figure 2, which confirms what was explained before, i.e. the differences observed between the measurements of two
unsynchronized receivers are a function of the Doppler frequency.

We then modified the carrier phase measurements as follows:

\[ \phi_{\text{new}}(t, \text{sat}) = \phi_{\text{raw}}(t, \text{sat}) + \tau . f_D(t, \text{sat}) \]  

(7)

where \( \tau \) is the delay between the code and carrier phase latching. The Doppler increment is added to the measurements on L1 and L2 separately, and the ionosphere-free combination is computed afterwards. The clock solutions obtained by PPP for different values of \( \tau \) are presented in Figure 5, and compared to the standard solution (in black), i.e. obtained from the raw code and carrier phase measurements.

Furthermore, a drift of about 60 ps per day was observed in the receiver NIST [5]. Contacts should therefore be taken with the receiver manufacturer to discuss about a possible delay of 2 \( \mu \)s between the code and phase measurement latching.

For a receiver of known delay between the code and phase measurement latching, the associated Doppler increment can therefore be corrected before the GNSS data analysis to avoid the artificial slope in the clock solution.

III. ESTIMATING THE CODE PHASE LATCHING OFFSET

The delay between the code and phase measurement latching can be estimated by computing carrier phase single differences in a zero-baseline setup (see Figure 2). However, it would be interesting to see if this delay can also be estimated in non zero-baseline configurations, as part of the PPP algorithm. The proposed approach consists in adding an incremental Doppler in the carrier phase measurements, and solve for the delay in the least square inversion. The PPP tool used for this experiment is Atomium [6]. The modeling of phase observations is therefore modified as follows:

\[ P_3 = \rho + \text{tropo} - \Delta t_{\text{sat}} + \Delta t_{\text{rec}} \]  

\[ L_3 = \rho + \text{tropo} - \Delta t_{\text{sat}} + \Delta t_{\text{rec}} + N_3 + w + \tau . \text{Dop} \]  

where \( P_3 \) and \( L_3 \) are the ionosphere-free combinations of the code and carrier phase measurements on L1 and L2. \( \rho \) is the geometric distance between the satellite and the receiver, \( \text{tropo} \) is the tropospheric delay, \( \Delta t_{\text{rec}} \) and \( \Delta t_{\text{sat}} \) are the receiver and satellite clock offsets, \( N_3 \) is the float ambiguity corresponding to the ionosphere-free combinations, \( \tau \) is the unknown delay between the code and carrier-phase measurement latching, \( w \) is the windup, and \( \text{Dop} \) is the ionosphere-free combination of the Doppler frequencies converted in meters for L1 and L2.

In order to validate the new inversion, the following strategy was used:

1. The value of \( \tau = -12.6 \mu s \) was determined for a given day and a given station using the modified Atomium described here above.

2. An additional delay of 5 \( \mu \)s was simulated between the codes and carrier phases of the same data set, using the same approach as in Section II, i.e. adding an incremental Doppler as in equation (6).

3. The value of \( \tau = -17.6 \mu s \) was again determined from these modified data.

We therefore retrieved in the inversion exactly the same delay as the one which was simulated, as the estimated \( \tau \) has increased from 1.26e-5 to 1.76e-5 second. Figure 6 presents the clock solutions obtained either with the classical PPP inversion (in black), or with the additional estimation of \( \tau \) either on the raw data (in green) or on the modified data where an artificial delay between code and carrier phase measurements of 5 \( \mu \)s...
has been added. The red curve has been slightly shifted downwards in order to make it visible, because it is exactly equal to the green curve. This result indicate that if a delay exists between the code and carrier phase latching, it is perfectly detected by the modified version of Atomium, and the clock solution is not affected by this additional parameter estimation.

The modified PPP software Atomium was then applied on several stations equipped with different receiver makes, during a 3 month period. Three of these stations are located in Brussels, and connected to the same cesium clock HP5071A. Two of them (ZTB1 and ZTB3) are additionally using the same antenna. The data where analyzed in daily data batches and the estimated delays for each day are reproduced in Figure 7. The first observation concerns the magnitude of the delays determined. It is indeed at the level of tens of micro-second which seems unrealistic. Furthermore, the delays determined are not constant but present strong variations from day to day, while if a delay exists between the code and carrier phase latching, then it is related to the receiver architecture and should be constant with time. The likely reason is that the estimation of \( \tau \) is highly sensitive to the colored measurement noise and multipath. Looking more specifically to the two receivers in zero clock and zero baseline, i.e. ZTB1 and ZTB3, on mjd 57060, there is a difference of 10 \( \mu s \) between the delay estimated for both receivers. However, the differences between the carrier phase measurements of both receivers do not show some specific Doppler-like behavior as shown in Figure 8.

As a last observation from Figure 7, there is a correlation between the values determined for the three receivers located in Brussels. A clock-related effect could therefore be investigated in the future.

IV. CONCLUSION

This paper proposed to explain some apparent differences between the frequency of the clock solutions obtained from the analysis of GNSS codes measurements or of the phase measurements. The proposed explanation of this discrepancy is the time offset that can exist for some receivers between the code and carrier phase latching. This paper explained how a code-phase bias in the receiver hardware can induce a frequency difference between the code and the carrier phase clock solutions. It was demonstrated that the effect can efficiently be modeled using the estimated Doppler frequency determined from the phase data. As a result of our simulations, it was shown that the drift in the clock solution is directly proportional to the delay between the code and phase measurements, and for a delay of 1 \( \mu s \), a daily drift in the PPP solution will appear with a magnitude of 30 ps. For a receiver of known delay between the code and phase measurement latching, the associated Doppler increment can therefore be corrected before the GNSS data analysis to avoid the artificial slope in the clock solution.

In a second part of the study, it was proposed to determine possible delay between the code and carrier phase measurements in the PPP inversion, as an additional parameter. The results obtained there were however unrealistic; most probably some noise in the data is absorbed in the estimated delay.
ACKNOWLEDGMENT

Some GNSS measurements from the National Metrology Institutes NIST, INRIM, OP and PTB, have been used in this study; the authors acknowledge them for making their measurements available.

REFERENCES


All Digital Frequency Synthesis Based on New Sigma-Delta Modulation Architectures

Paul P. Sotiriadis
Department of Electrical and Computer Science
National Technical University of Athens
Athens, Greece
E-mail: pps@ieee.org

Abstract—This work presents a general architecture for all-digital frequency synthesizers based on a frequency-shifted \(\Sigma/\Delta\) modulator. The synthesizers can achieve spurs free spectrum within the band of interest and high dynamic range. System level aspects of the synthesizers are discussed and MATLAB simulation results are presented.

Keywords—Dithering, Direct digital frequency synthesis, noise shaping, quantization, spectrum

I. INTRODUCTION

Over the past thirty years all-digital frequency synthesis has attracted the interest of the scientific community because of the advantages of the digital circuits like design-automation, testability, robustness to noise, temperature, supply and process variations, etc. Downscaling of integration technologies makes all-digital frequency synthesis even more desirable as digital circuitry becomes faster and smaller in area while analog RF becomes more challenging to design and with limited area-scaling capability.

Perhaps the most successful digital-intensive frequency synthesizers are the All-Digital PPLs (ADPLL) [1] in Fig. 1 and the Direct Digital Synthesizers (DDS) [2] in Fig. 2. Even though both have an enormous number of practical applications, they both use analog and mixed signal blocks like the time-to-digital converter [1], the VCO and the DAC, which need careful design in the particular IC technology the circuit is fabricated.

A way to get fully-digital frequency synthesizers is to remove the DAC in DDS (Fig. 2), i.e., the only analog (mixed-signal) block in DDS and use the MSB of the LUT as the output. This results in the Finite State Machine (FSM) with single-bit (SB) output in Fig. 3, sometimes called Pulse-DDS (PDDS).

Another successful fully-digital frequency synthesizer is the Flying-Adder [3] in Fig. 5. It is a period-synthesizer in the sense that the output average period is proportional to the frequency control word \(w\). The single-bit output of the Flying-Adder has similar spectral properties with that of the PDDS, i.e. the output spectrum is full of spurs for values of \(w\) not resulting to integer frequency division. A typical example of the simulated spectrum is shown in Fig. 6.
The use of dithering in both PDDS and the Flying-Adder synthesizers can alleviate or completely eliminate spurs. Random dithering is the only purely digital technique that can achieve this. The spectra of (dithered) PDDS and (dithered) Flying-Adder, corresponding to Figs. 4 & 6 are shown in Figs. 7 & 8 respectively.

Dithering may remove spurs but it introduces noise which typically appears as a noise floor, weighted by the spectrum of the shape of the output pulses. If white uniformly distributed dithering is used, the Dynamic Range, defined as the power of the carrier over the power spectral density of the near-in noise is in the order of $10 \log_{10} \left( \frac{f_s}{f_{fs}} \right)$ for both the PDDS and Flying Adder. This is because of the single-bit quantization with independent of the architecture.

Direct dithering techniques, e.g. use of colored random sequences, (without feedback) can improve the dynamic range but not a lot [4-5].

II. DIRECT FREQUENCY SYNTHESIS BASED ON SIGMA-DELTA MODULATION

The previous section indicates the difficulty of achieving high dynamic range and SFDR from direct all-digital frequency synthesizers with single-bit output that have a direct signal path. This translates to the difficulty of shaping the spectrum of the power of the strong 1-bit quantization noise. Feedback structures instead allow for very efficient noise shaping resulting in high dynamic range, and ideally spur-free band of interest.

To generate a single-bit sequence with sinusoidal like spectrum we use a phase accumulator and a cosine LUT to generate a sinewave in the digital domain exactly as in the standard DDS, as shown in Fig. 9.

Instead of the multi-bit DAC of the DDS we use a 1-Bit $\Sigma\Delta$ modulator to convert the multi-bit instantaneous value of the cosine to $\pm 1$, i.e. one Bit as illustrated in Fig. 10. We assumed a general form of the $\Sigma\Delta$ modulator with an extra gain $g$ in the feedback path. The sequence $d(k)$ is a dithering sequence to help suppressing spurious signals further.

Assuming for now (only) that the quantizer is equivalent to an addition with a quantization noise sequence $n(k)$, then $X(z) = STF(z) \cdot U(z) + NTF(z) \cdot D(z) + NTF(z) \cdot N(z)$, where $U(z)$ is the Z-transform of the input $A \cos(k\Omega)$,

$$STF(z) = \frac{F(z)}{1 + gF(z)}$$ is the signal transfer function and
$NTF(z) = \frac{1}{1 + gF(z)}$ is the noise transfer function of the $\Sigma/\Delta$ modulator.

Apparently the transfer function $F(e^{j\omega})$ should attain large absolute values within the band of interest, around $\Omega$, in order to make $|NTF(e^{j\omega})|$ small and $|STF(e^{j\omega})|$ close to one, as illustrated in Fig. 11 from [6].

![Figure 11: Typical Band-Pass NTF and STF.](image)

The poles of $F(z)$ must be accumulated in the band of interest around $\Omega$, and the zeros should be chosen to optimize noise suppression and ensure stability of the loop as shown in Fig. 12.

![Figure 12: Poles and Zeros of $F(z)$.](image)

A challenge with this choice is that $F(z)$ must be changed when the desirable frequency $\Omega$ changes (and the band of interest around it). To resolve this difficulty and allow wide range of change of $\Omega$ without any "tuning" of the parameters, the filter is implemented as follows.

### A. Frequency Shifted Filter

Since the desirable is the generated single-bit output to resemble the spectrum of $A\cos(k\Omega)$, it is reasonable to consider the band of interest centered at $\Omega$. To this end, we can implement $F(z)$ as a frequency shifted version of a Base-Band filter $H(z)$, i.e. as in Fig. 13.

![Figure 13: Frequency-Shifted implementation of $F(z)$.](image)

The input is down-converted by multiplying with $\cos(k\Omega)$ and $\sin(k\Omega)$, filtered through two independent paths, and, then up-converted by another set of multiplications. It can be shown that the input-output transfer function of the structure in Fig. 13 is

$$F(z) \equiv \frac{Y(z)}{X(z)} = \frac{1}{2} \left( H(ze^{-j\Omega}) + H(ze^{j\Omega}) \right).$$

### B. More Accurate Modeling and Stability

For low (single-bit) resolution quantizer, the accuracy of the $NTF$ and $STF$ we get under the assumption that the quantizer is equivalent to an addition of a quantization noise is inadequate. Instead, we prefer to use a quasi-linear model of the loop based on the Random-Input Describing Function method [7]. To this end, the 1-Bit quantizer is modeled as a set of parallel paths, one of gain $K$ for the sinusoidal signal and one of gain $L$ and the addition of quantization noise $n(k)$, for the (total) noise. Then the model in Fig. 10 is decomposed into two pieces: A) a loop capturing the behavior of the loop in Fig. 10 to the sinusoidal signal:

![Figure 14: Equivalent loop for the sinusoidal behavior of the loop in Fig. 10, based on the Random-Input Describing Function method.](image)

It gives directly $X_c(z) = \frac{KF(z)}{1 + gKF(z)} U(z)$, for which assuming that the magnitude of $|F(e^{j\omega})| \to \infty$ as $\omega \to \Omega$ we get that $x_c(k) = (A/g)\cos(k\Omega)$; and B) a loop capturing the behavior of the loop in Fig. 10 to the noise signal:

![Figure 15: Equivalent loop for the noise behavior of the loop in Fig. 10, based on the Random-Input Describing Function method.](image)

giving $X_n(z) = \frac{L}{1 + gLF(z)} D(z) + \frac{1}{1 + gLF(z)} N(z)$

where both the dithering (noise) and the quantization noise are propagated to the output.

The total output is the sum $X(z) = X_c(z) + X_n(z)$. The derivation of the quasi-linearization gains $L$ and $K$ is done in a way similar to [7]. For the input $A\cos(k\Omega)$, it is...
\[ STF(z) = \frac{KF(z)}{1 + gKF(z)} \] and \[ NTF(z) = \frac{1}{1 + gLF(z)} \]. Moreover we also need to define the Dither-transfer function \[ DTF(z) = \frac{L}{1 + gLF(z)} \] as it is different from \( NTF(z) \).

Then for the sinusoidal input \( U(z) \),

\[ X(z) = STF(z) \cdot U(z) + DTF(z) \cdot D(z) + NTF(z) \cdot N(z) \].

In terms of the stability of the loop, first we note that for the output it is \( x(k) \in \{\pm 1\} \) and since the maximum sinusoidal amplitude of the output is \( 4/\pi \), it must be \( A < 4g/\pi \). Also, there are power constraints since the total noise power plus the power of the sinusoidal at the output must equal 1. This provides necessary equations to solve for the gains and signal powers in the loop [7].

The next thing is to examine the pole location of all three transfer functions \( STF(z) \), \( DTF(z) \) and \( NTF(z) \). Note that the gain \( g \) was deliberately inserted in the loop to allow for tuning the poles' location. For the purpose of this paper, we discuss the following examples.

C. Example 1

Consider the case of a transfer function \( H(z) \) with poles and zeros as shown in Fig. 16. It has three pairs of complex conjugate poles and a pole at 1, all on the unit circle, and six zeros inside the unit circle.

The impact of the loop gain \( g \) on the stability of the loop is illustrated by the root-locus of the transfer function \[ \frac{1}{1 + a \cdot F(z)} \] in Fig. 20. Here \( a \) corresponds to the value of \( g \cdot K \) in the \( STF \) and to \( g \cdot L \) in the \( NTF \) and \( DTF \).

![Figure 16: Poles and Zeros of \( H(z) \) of the example](image)

![Figure 17: Poles and Zeros of \( F(z) \) of the example](image)

![Figure 18: Zoom in on the poles and Zeros of \( F(z) \) of the example](image)

![Figure 19: Amplitude graph of \( F(z) \) .](image)

![Figure 20: Root locus of \( 1/(1 + a \cdot F(z)) \) for \( a \) ranging from 0 to 1.](image)
The output spectrum of the frequency synthesizer is shown in Fig. 21. It corresponds to clock frequency \( f_{\text{clk}} = 1\,\text{GHz} \) and equivalent resolution bandwidth (on the spectrum analyzer), \( \text{RBW}=25\,\text{kHz} \). The Nyquist frequency is \( f_{\text{Nyquist}} = f_{\text{clk}} / 2 \). Therefore the noise level appearing at about 

\(-65\,\text{dBc}\) is down to 

\(-65\,\text{dBc/Hz}\). Whether there exist hidden frequency spurs below 

\(-65\,\text{dBc}\) needs to be investigated further via simulation or measurement.

\[10^{10}\log 6250 \, \text{dB} = -185 \, \text{dBc/Hz}.\]

D. Example 2

This example shows the output spectrum when a more narrow filter is used. The order of the filter \( H(z) \) is again equal to 7. Equivalent clock frequency and resolution bandwidth are \( f_{\text{clk}} = 1\,\text{GHz} \) and \( \text{RBW}=6.25\,\text{kHz} \), respectively.

A zoom in centered on \( \Omega \) is shown in Fig. 23. The noise level appearing at about 

\(-145\,\text{dBc}\) is in principle down to 

\(-145\,\text{dBc/Hz} - 10\log_{10} (6250) \, \text{dB} = -185 \, \text{dBc/Hz}.\)

\[\frac{f}{f_{\text{Nyquist}}} = \frac{f_{\text{freq}}}{f_{\text{Nyquist}}} \]

E. Final Remarks

A major aspect of the synthesizer is the complexity and power consumption. Multipliers are typically the most power consuming and slow elements. A variation of the architecture without the four multipliers in Fig. 13 is possible. The results are similar with the exception of some spurs -outside the band of interest. An example of output spectrum is shown in Fig. 24.

\[\frac{f}{f_{\text{Nyquist}}} = \frac{f_{\text{freq}}}{f_{\text{Nyquist}}} \]

III. CONCLUSIONS

This work presented an all-digital frequency synthesizer based on a frequency-shifted \( \Sigma/\Delta \) modulator. Design choices have been discussed along with performance and stability aspects. Output spectra demonstrated the capability of the synthesizer to achieve very high dynamic range within the band of interest. Typical spectrum of a variation of the synthesizer with multiplier-less structure has been presented.

ACKNOWLEDGEMENT

Described work was partially supported by Broadcom Foundation USA

REFERENCES

Noise in High-Speed Digital-to-Analog Converters

P.-Y. Bourgeois\textsuperscript{1}, T. Imaike\textsuperscript{2}, G. Goavec-Merou\textsuperscript{1} and E. Rubiola\textsuperscript{1}

\textsuperscript{1}FEMTO-ST Institute, Time & Frequency Dept, UMR 6174-CNRS, University of Franche-Comté, Besançon, France
\textsuperscript{2}Nihon University, Dept of Electronic Engineering, Japan
Email: pyb2@femto-st.fr

Abstract—We report on the measurement of phase noise of high speed analog to digital converters in a full digital measurement setup and for various development boards. The tested configurations ensures a Nyquist rate higher than 100 MHz suitable for conventional ultralow noise devices. Several analogous to digital converters featuring a SNR higher than 140 dB enable the measurement of AM and PM noise with a background noise of -185 dBC (floor) and -160 dBC (flicker, 10 Hz off the carrier).

I. INTRODUCTION

Digital tools are a mature technology to perform dynamic signal analysis on ultrasound clocks and devices presenting unprecedent levels of stability[1-3]. Jointly with the help of FPGAs used as real-time coprocessors and CPUs running multi-task operating systems with double precision, cross-correlation techniques on full samples phase times series benefit from high bandwidths up to 100 MHz off the carrier. The IF conversion is done after sampling, analysis resolution is limited in the measurement time by the resolution of the digital to analog converters. We present in this paper various tested configurations ensuring a Nyquist rate higher than 100 MHz, where $f_s$ represents the sampling rate.

III. Digital Architectures

In this project, we have operated a selection of various platforms with integration of FPGA and CPU cores. All algorithms have been developed from scratch in C as a library featuring basic blocks functions, double and single precision in order to predict correct behaviour of their hardware description counterparts. To our point of view, this is the correct way to fully master the full system flow and measurement chain at every stage. In this manner it will be possible in the future to develop accurate models including quantization noise processes. Indeed this excludes the use of any proprietary blackboxes.

We present here 3 digital architectures that are, amongst others, under test at FEMTO-ST.

We have selected a high-speed digitizing system, from Alazartech company, consisting of 2 synchronized boards embedding each 2 AD9467 (16 bits, 250 Msps), 2 Altera Stratix III FPGAs (main/coprocessor) and PCIe extension connected to multicore PC station running debian-based GNU-Linux kernel 3.16. The versatility of such a system enable ease of retrieving continuous samples at full speed, fast development and algorithms testing as the interfaces and communication parts are already provided.

Second, the recent multicore SoC FPGAs as Zynq/Cyclone V systems offer potential high-end features and are of growing interest. For we have conducted tests on Zynq-based platforms (ZC706 coupled to 2 dual-channel LTC2158 (14 bits, 310 Msps), and also tested the dual-channel LTC2145 (14 bits, 125 Msps) of the Redpitaya system, although this last platform is more dedicated to low-quality general purposes (small FPGA, only 2 channels, not Open Source). We report in this paper various tests performed on several kind of platforms based on FPGAs with deported CPUs or the latest SoCs embedding hard cores CPUs.

II. NUMBERS

Assuming a uniform quantizer, the quantum resolution step is defined as the ratio of the voltage full scale range and the number of bits $M$:

$$ q = \frac{v_{fsr}}{2^M - 1} \quad (1) $$

The associated noise is a statistical process representing the density of probability of states within a measurement length $\tau$:

$$ \sigma^2 = \frac{1}{\tau} \int_{-\tau/2}^{\tau/2} c_q(t)dt = \frac{q}{12} \quad (2) $$

Finally the total noise represents the integrated noise over the measurement bandwidth, directly related to the Nyquist frequency:

$$ N = \frac{\sigma^2}{f_N} \quad (3) $$

This last equation remains only when proper filtering and small fraction of aliasing occurs, thanks to Parseval theorem.

Eventually, for an incoming wave of peak voltage $a \sim v_{fsr}$, one may derive the signal to noise ratio:

$$ SNR = \frac{4 \cdot q^2}{3 \cdot v_{fsr}^2 \cdot a^2 \cdot f_s} \sim \frac{4}{3 \cdot 2^M \cdot f_s} \quad (4) $$
The digital phase noise measurement system[3] is depicted in the preceding figure. The phase modulated noise degrading a perfect sinusoid is downconverted to DC after sampling thanks to a numerically controlled oscillator (NCO) set up at the carrier frequency. Successive filtering/decimation stages allows to focus on lower decades off the carrier or examine the spectral measure at lower sampling rates by filtering out aliased noise while reducing the measurement bandwidth. Phase estimation is done by calculating the arctangent function of the in-phase and quadrature components of the demodulated and filtered signal. Eventually the amplitude is also estimated. From the phase time series (amplitude time series), the Fourier transform is computed thanks to the FFTW[4] algorithm. Other filtering/decimation stages completes the process of lowering the decades. Finally the spectrum of variances is reconstructed from these decades. The demodulation process and first decimation/filtering stages may be abusively called digital down converter (DDC).

V. SOFTWARE CALIBRATION

In order to verify the correct interpretation of the spectral measure (normalizations processes), a noise generator calibrated at -113 dBc mixed with a low noise 10 MHz reference signal is send to the analog to digital converter. The counterpart digital down converter/reconstructed from these decades. The demodulation process and first decimation/filtering stages may be abusively called digital down converter (DDC).

The power spectral density is evaluated from the collected samples at full speed with double precision (‘calculated pn from samples’ in the figure). The counterpart digital down conversion in single precision version embedded into the FPGA, by using a squarewave NCO (inverted samples every 25 points per period for a 10 MHz carrier at 250 MHz sampling rate) and a 127 coefficients Blackman-Harris windowed sinc filter to ensure sufficient aliased noise rejection, shows similar results. The first stage output effectively ensures that no extra quantization noise is induced by the chain, particularly the full bit width remains unchanged. This process is no longer true when multiple stages must be embedded while performing slice rescaling, and additive noise may be taken into account.

VI. ADC NOISE MEASUREMENTS

A. ADC noise measurement principle

The setup of ADC phase noise measurement is depicted in the following figure.

For the LTC2158, the input carrier was 6.6 dBm, with a voltage fullscale range of $v_{fsr} = 1.35$ V. Slices of $2^{17}$ points were taken at every output stage of the filters to reconstruct the spectrum. White noise is about 10 nV/$\sqrt{\text{Hz}}$ and flicker of 5.6 $\mu$V/$\sqrt{\text{Hz}}$. SNR=154 dB for white noise.

For the Redpitaya system, the input stage was bypassed and loaded to 50 $\Omega$. Only the LVDS amplifier was kept to prepare the ADC to be feeded with differential signal. The gain is about 2, for a 0 dBm signal and a $v_{fsr} = 1.25$ V at the input of the analog to digital converter (the measure was done in a differential mode). The onboard clock system was used to clock the ADC at 125 Msp/s, slices of $2^{14}$ data were taken, and because of spare space, only two filtering/decimation stages were performed within the FPGA. The obtained white noise
is $29\text{nV}/\sqrt{\text{Hz}}$ and $4.5\mu\text{V}/\sqrt{\text{Hz}}$ for flicker. SNR=$143.5$ dB (white) in a 1Hz bandwidth.

Results from the AD9467 of the Alazartech boards presented, for a 12 dBm carrier ($v_{fsr} = 2.5$ V), a white noise floor of about $-157$ dBV$/\text{Hz}$ (14 nV$/\sqrt{\text{Hz}}$) and a flicker of $-110$dBV$/\text{Hz}$ (3 $\mu$V$/\sqrt{\text{Hz}}$). Referred to the carrier, this is equivalent to a SNR of 156 dB (white) and 109 (1 Hz) in a 1 Hz bandwidth.

Taking into account the 1 dBc of carrier power, the presented results are compatible with phase noise measurements.

B. Effective number of bits evaluation

This technique is perfectly suitable to a fast evaluation of the effective number of bits (ENOB) of the analog to digital converters. One may just analyze the white noise floor for quick evaluation of the ENOB. Technically, this may only need 1024 samples for example and a FFT evaluation on 512 points, or even less ; to get better accuracy, simple averaging may help. We have derived the ENOB calculation assuming uniform quantization and the fact that it is directly related to the signal to noise ratio :

$$ENOB = \log_2 \left( 1 + \frac{v_{fsr}}{\sqrt{2 \cdot f_N \cdot S_{floor}}} \right)$$ (5)

where $S_{floor}$ is the measured white voltage noise floor, $v_{fsr}$ is the voltage full scale range and $f_N$ the Nyquist frequency.

Applied to the AD9467 system, the measurement of $S_{floor} \sim -158$ dBV$/\text{Hz}$ lead to an effective number of bits of about 12 in agreement with the technical datasheet.

C. ADC noise vs sampling rate

White phase noise floor directly depends on the sampling rate as shown on the following figure.

The measured noise floors for various sampling rates are in perfect agreement with their theoretical expectations.

VII. SINGLE CHANNEL NOISE MEASUREMENT

The digital single channel noise floor measurement obeys the setup described on the following figure. When the two converters are feeded with a full scale range 10 MHz low noise wave, we may have access to the single channel noise by differentiating the 2 arms after digital down conversion, phase extraction and spectrum calculation.

For the noise budget, the uncorrelated arms noises are suppressed and just remains the contribution of the converters noises.

The presented $\mathcal{L}_f$ spectrum shows it is possible to perform measurement of low noise oscillators up to -160 dBc without the need of more complex architecture.

VIII. APPLICATION TO THE MEASUREMENT OF CSO

It is possible to apply the 2 channels technique to the measurement of a pair of cryogenic sapphire oscillators (CSO) exhibiting a frequency instability in the $10^{-15}$ range.
beatnote of about 7.029 MHz of a pair of CSO feeds a two-channel digital phase noise measurement system. The beatnote is downconverted and successive filtering/decimation stages are applied.

The setup favorably compares with the indicated noise floor of the TSC5125 but with only a 2 channels configuration.

IX. 4-CHANNELS DIGITAL SIGNAL ANALYZER WITH CROSS-CORRELATION

We have applied the all-digital 4-channel cross-correlation technique as described in [3]. After sampling and gain adjustment, the digital down conversion is applied and phase extracted from the I/Q data flow at a rate of 25 Msps. The phases time series of 2 pairs of channels are differentiated and the cross-spectrum is calculated.

For the measurement of the system noise floor, a 10 MHz signal is send to each of the 4 channels. With $10^7$ correlations on a continuous data flow, the obtained floor quickly reaches -185 dBc and is below -160 dBc for Fourier frequencies at 10 Hz off the carrier (1000 correlations).

As an application, we have performed the measurement of the low noise synthesizer used in our experiments, the R&S SMA 100A with low noise option. The resulting plot is compared to the expensive Agilent PN5052B.

X. CONCLUSION

In this paper we have presented a correct all-digital technique for the evaluation of the noise of high-speed analog to digital converters. The setup is suitable for fast analysis of the effective number of bits of such converters, the main parameter related to quantization noise and signal to noise ratio, directly impacting the resolution of digital measurement systems. A 2 channel phase and amplitude noise measurement system has been developped and applied to the measurement of ultra-stable cryogenic sapphire oscillator. Also we have presented an extension of a 4-channels cross-correlation system resulting of noise floor of -185dBc, confirming the potential of such a technique.

ACKNOWLEDGMENT

This work is a part of the Programme d’Investissement d’Avenir at TF Dept of FEMTO-ST Institute (Oscillator IMP, First-TF, and Refimeve+), supported by the French ANR, and also supported the Region Franche Comté and by the Nihon University, Japan.

REFERENCES

[4] http://fftw.org/ ; FFTW was written by Matteo Frigo and Steven G. Johnson
Simple Method for ADC Characterization under the Frame of Digital PM and AM Noise Measurement

Cárdenas-Olaya Andrea C.*, Rubiola Enrico+, Friedt Jean-M.*, Ortolano Massimo+, Salvatore Micalizio*, Calosso Claudio E.*
* Department of Time and Frequency, INRiM, Turín, Italy
*Department of Electronics and Telecommunication, Politecnico di Torino, Turín, Italy
+ CNRS/UFc FEMTO-ST Institute, Besançon, France
*Department of Time and Frequency, FEMTOIST Institute, Besançon, France
E-mail: a.cardenas@inrim.it

Abstract—The last years improvements of electronic circuits has allowed the appliance of digital systems in phase noise measurement techniques where low noise and high accuracy are required, yielding flexibility in the implementation and setup of measurement systems. By definition, any measure performed is always affected and limited by the noise of the measurement instrument itself. Considering that the Analog to Digital Converter (ADC) is the core and front end of digital systems, its residual noise has an important impact on the system performance. Consequently, the selection of the proper ADC becomes a critical issue for the system implementation. Currently, the information available in literature deeply describes the ADC features mainly at frequencies offsets far-from-carrier. Nevertheless for time and frequency applications the performance close to the carrier is an important concern as well. In this paper, a simple method for ADC characterization is proposed based on the Phase Locked Loop (PLL) definition and on Phase and Amplitude Modulation (PM/AM) measurements, focused in obtaining the relevant information of ADC noise contributions for phase noise measurement applications. The purpose of such a method is to find the parameters of a state ADC noise model using a technique which avoids the use of complex hardware and allows having a low computational costs performance.

Keywords—phase noise; analog to digital converter; digital signal processing; phase modulation; amplitude modulation; PLL

I. INTRODUCTION

Phase noise measurement has been an important subject of study due to the serial implications that phase noise has on systems in which high source frequency stability is required in order to guarantee a correct and accurate performance, such as radar applications, ultra-stable oscillators, data communication links and multichannel receivers.

Phase noise represents the random frequency fluctuations caused by phase instability [1]. Therefore, the phase noise measurement of a system source provides the information about its frequency stability. Since the phase noise is a critical parameter in the selection of the appropriated system source(s), different techniques have been developed in order to measure it, most of them based on the use of spectrum analyzers and analog systems [2][3]. The classical phase measurement setup is the quadrature method depicted in Fig. 1.

![Fig. 1. Phase noise measurement. Quadrature method.](image1)

The measurement system works as an appropriate phase detector if the device under test (DUT) and the reference (REF) are in phase quadrature due to the mixer operating principle. Considering that meeting this requirement by manual configuration is not efficient neither easy, a feedback is added in order to keep the quadrature condition, as shown in Fig. 2.

![Fig. 2. Method with feedback to keep phase quadrature.](image2)
Furthermore, considering that analog systems are highly affected by mechanical noise (50Hz – 1 kHz) and that most of the system characterization of interest works from mHz to tens of MHz offset with respect to the carrier, the phase noise measurement has to be performed under excellent conditions of mechanical noise isolation (hardware design - PCBs, connections between stages) increasing the complexity of the implementation.

The problems previously mentioned are not only evidenced in the quadrature method. Parameterization issue and mechanical noise are problems that affect in general analog systems. Hence, in order to increase the system flexibility and reduce the effects of mechanical noise, temperature dependence, drift, aging and tuning, approaches as Software Defined Radio (SDR) [4] has started to be applied.

In the ideal SDR concept, the system architecture will consist in an analog to digital converter (ADC) and in the digital process block (Fig. 3).

![Fig. 3. Ideal SDR architecture.](image)

However, due to technological limitations of ADC and signal processing bandwidth, the real SDR system architecture may differ from the ideal solution depending of the application requirements. For the case of phase noise measurement, the ADC bandwidth could be an important constraint to the system performance, therefore, in order to reduce the bandwidth of the ADC input signal, the mixing and filtering stages can be performed before sampling.

Nevertheless, the analog mixers introduce AM noise to the system, adding a contribution to the phase noise characteristics not related to the DUT and challenging to subtract from the records. Thus, considering that the recent technological advances allow finding ADC components with wide operational bandwidth in the market, the possibility to implement phase noise measurement systems as the ideal SDR concept states is feasible and has been applied with different techniques [5, 6].

But, how to know which ADC suite better or more properly phase noise measurement requirements?

Generally as first stage, the selection of an ADC is based on the datasheet information, in particular number of bits, sampling frequency, bandwidth and signal to noise ratio, features that determine the main characteristics of the measurement outcome. In time and frequency applications and in particular for phase noise measurements, the ADC noise spectra provides additional information that allows the identification of ADC noise contributions to the data converted, and therefore the effects on the phase noise measured. Different techniques based on histograms and FFT analysis has been developed in order to characterize the ADC noise spectra [7]. Although such techniques are very useful for ADC characterization and test, they do not provide complete information about the ADC noise, due to the complexity of the hardware needed (components performances not still available, memory lack, etc.) especially at low frequencies, i.e., flicker noise, which may impact on phase noise measurements in oscillators at frequency offsets close-to-carrier.

Thus, the purpose of this work is proposed a method for ADC noise characterization that provides the ADC noise spectra at frequency from 1Hz or less using a simple system setup.

For accomplish this aim, a model of ADC noise is state, which sets the ADC noise as a function of two parameters or noise contributions, additive noise and jitter noise, as described in section II. Subsequently, based on this model, the method proposed will find such parameters through PM and AM measurements, tracking the relevant ADC information using a PLL (Section III). The implementation of this method was performed using Red Pitaya platform as explain in Section IV, which provides the hardware needed for a first approximation.

II. ADC NOISE MODEL DESCRIPTION

In order to discriminate the nature of the noise that can affect an oscillator signal, an ADC noise model was adopted (Fig. 4). It consists in three noise components. The first one, $n_a$, represents additive noise caused by thermal noise, circuit construction, voltage reference, which is presented mainly as amplitude modulation noise. The second one, $n_j$, represents the noise caused by the aperture jitter or aperture uncertainty in the sample and hold during the data acquisition, which is presented as phase modulation noise at the ADC output. The third component is the quantization noise, characteristic well known which is spread along the ADC bandwidth.

![Fig. 4. ADC noise model adopted.](image)

Hence, the model under study will depend only of the unknown components, being a function of the $n_a$ and $n_j$ parameters, as depicted by the eq. (1) that intrinsically will take into account the quantization noise effect.

$$n_{ADC} = f(n_a, n_j)$$ (1)
The main objective of using this model is to be able to predict the ADC noise contributes on the phase noise measurement based on the parameters information.

III. METHOD PROPOSED FOR ADC NOISE CHARACTERIZATION

According with the model described in the previous section, two parameters must be determinate, $n_a$ and $n_j$, which generate amplitude and phase modulation noise respectively. From Fig. 5 can be observed how amplitude fluctuations affect the points of maximum amplitude in an oscillator while phase fluctuations are easily detected at the zeros-cross points. Thus, tracking these points of the ADC output, having as input a sinusoidal signal from an oscillator will allow determine the parameters value but including the noise of the oscillator as well.

![Fig. 5. Oscillator noise model. Amplitude and Phase Noise.](image)

However, since the phase noise measurements under study are based on differential techniques, the digital instrumentation will used at least two independent ADC channels. Therefore, the ADC noise contribution to the phase-meter will be differential as well. In consequence, the common noise between the two ADC channels will be cancel and therefore the amplitude and phase fluctuations presented in the points mentioned above, will be traduced to ADC noise as described in Fig. 6.

![Fig. 6. Description of common noise cancellation using two ADC channels.](image)

In order to track the points with relevant information of amplitude and phase fluctuations, the method depicted in Fig. 7 is proposed.

![Fig. 7. Proposed method for ADC noise characterization.](image)

The synthesizer will generate the sinusoidal input signal for both channels of the ADC, but one of them will be used also to track the points of interest using a down-sampling block (DEC) and a PI controller. The output of the controller will be connected to a DAC which will provide the proper information to the synthesizer that also works as Voltage Controller Oscillator (VCO) correcting the signal generated in order to acquire the set point configured, maximum amplitude point or zero-cross point.

The ADC noise spectra will be obtained performing the proper decimation and filtering. This last stage will consist in an accumulator and an average filter.

IV. METHOD IMPLEMENTATION

As platform for implementation or test bench was used Red Pitaya [9] an open source embedded system that includes a 14 bits ADC of two channels at 125MSps, a dual DAC of 14 bits at 125MSps and a System On Chip (SoC) Zynq 7010 from Xilinx (FPGA+ARM). Additionally, this platform counts with a PID controller block ready to be used and all the drivers for the ADC and DAC operation. Fig. 8 depicts the block diagram of the method implemented.

As input of the system, a sinusoidal input signal was generated with carrier frequency ($\nu_o$) of 31.25MHz and 1.5Vpp. Due to the fact that the input frequency is four times the sampling frequency (125MHz), the down-sampling factor is set to four. With this configuration is acquired one sample per period, which can be the maximum amplitude point or zero-crossing point, depending of the system configuration. The data are store in memory blocks of 16384 samples each one of 32bits, i.e., blocks of around 65Kbytes.

The PID controller actually is a controller Proportional and Integral proper configured to track the points of relevant noise information.
V. RESULTS

A. PM measurement approach – $n_{\text{jitter}}$ parameter

In Fig. 9 is shown the noise spectra of the component $n_{\text{jitter}}$ measured using an input signal of $v_i = 31.25$ MHz (blue curve). The pink spectra is the same component measured without input, i.e., with a load resistor of 50Ω.

It can be observed how the jitter effect start to be evident increasing the white noise of the spectra when the frequency of the input signal increase, as was expected, due to the fact that the slope of the signal acquired increase being more sensitive to these phase fluctuations. The noise a low frequencies has a slope of -10dB/Hz than means flicker noise, which remains with the same value.

B. AM approach – additive noise parameter

Fig. 10 depicts the spectra of the additive noise component, $n_{\text{a}}$ refered to the spectra measured without input signal. In this case the white noise did not change because whether the amplitude fluctuations is low it will not affect the ADC spectra.

It can be observe noise around 80KHz that could be caused by the internal reference of the ADC whose effect was not cancel due to the fact that this is an ADC noise source that must be considered.
The white noise in both figures presented above does not take into account the aliasing contributions caused by the down-sampling process. Since it is the sampling frequency is divided by a factor of 4, the white noise is reduced 6dB.

The table reported below, depicts preliminary results about the ADC noise parameters. This first approximation can be compared with the performance of a generic analog mixer which has a flicker noise around -140dB/Brad2/Hz at 1Hz [10].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>White Noise</th>
<th>Flicker Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_a$</td>
<td>-149 dBV/Hz</td>
<td>-106 dBV/Hz</td>
</tr>
<tr>
<td>$n_{inter}$</td>
<td>-148 dB/Brad2/Hz</td>
<td>-105 dB/Brad2/Hz</td>
</tr>
</tbody>
</table>

VI. CONCLUSIONS

Preliminary results were obtained for the ADC noise parameters. It is imperative to validate the method under different conditions, using different ADC architectures and technologies in order to verify the accuracy of the method.

The measurements performed provide information along seven decades using approximately 40% of the FPGA resources, which implies not high computational resources.

The spectra obtained present low spurious along the bandwidth, it means not excess of noise caused by the digital processing.

REFERENCES

6/12-channel Synchronous Digital Phasemeter for Ultrastable Signal Characterization and Use

Massimo Caligaris, Costanzo Giovanni A.Á, Calosso Claudio E.∗
Department of Electronics and Telecommunication, Politecnico di Torino, Torino, Italy
∗Istituto Nazionale di Ricerca Metrologica, INRIM, Torino, Italy
ÁINRIM and Politecnico di Torino
E-mail: c.calosso@inrim.it

Abstract—Nowadays, in primary time and frequency laboratories we can find high spectral purity signals in the 10 MHz – 10 GHz range generated from cryogenic oscillators or ultra-stable lasers together with frequency combs. Their short-term stability surpasses by one to two orders of magnitude the performances of active hydrogen masers (AHM), while in the long-term AHMs still have a better behavior. The new technology can be considered mature for what concern spectral purity, but we cannot say the same about complexity, power consumption and reliability. In this sense, it is important to measure ultra-stable sources with respect to AHMs. First, to test their spectral purity or, at least, to give it an upper bound; second to have a continuous monitoring; finally, to combine them in order to get the best of all in term of phase noise and frequency stability. All of these requirements can be satisfied by the system we are developing. It is a multi-channel synchronous and real-time phasemeter based on Tracking Direct Digital Synthesizer (TDDS) technique. The results related to the first prototype are presented.

Keywords—phasemeter; digital electronics; DDS; FPGA; cross-variance;

I. INTRODUCTION

As anticipated in [1], the tracking DDS technique has many applications in time and frequency metrology. We firstly proposed it to generate a composite local oscillator for atomic fountains [2]. Then we used it with very noisy signals as the ones encountered in coherent fiber links [3]. Now we explore the possibility to use this scheme with ultra-low noise signals, as the one at 100 MHz, generated by a frequency comb referenced to an ultra-stable laser. In this sense, we implemented a new version of this phasemeter, with reduced residual noise and with synchronous channels, in order to take advantage of the cross-correlation and cross-variance techniques. After a brief recall of the principle of operation of the Tracking DDS, we describe the architecture of the phasemeter, its characterization in term of Allan Deviation and 2-sample cross-variance and its use to measure the comb versus our masers.

II. ARCHITECTURE

The system we are developing (Fig. 1) is a multi-channel synchronous and real-time phasemeter based on tracking Direct Digital Synthesizer (DDS) technique. The tracking DDS is a Phase Locked Loop (PLL) where the more usual Voltage Controlled Oscillator (VCO) is replaced by a Direct

![Fig. 1: block diagram of the 6-channel phasemeter. The tracking DDS is in the dashed rectangle.](image-url)
synchronous measures, at the level of microsecond. Each channel measures the phase of the input with respect to the phase of the local oscillator, a Voltage Controlled Saw Oscillator (VCSO) that runs at 1 GHz. It is convenient to use the phase time, to take into account the different frequencies involved automatically

\[ x_k^d = x_k^a - x_k^c + x_{TDDS} \quad k = 1 \ldots 6 \]  

where, for the channel \( k \), \( x_k^d \) is the digital output of the tracking DDS, \( x_k^a \) is the phase-time of the signal at the input, \( x_k^c \) the noise of the VCSO and \( x_{TDDS} \) the residual noise of the channel. The noise of the local oscillator \( x_k^c \) is in common mode and can be cancelled by considering the difference of two channels or of a weighed combination of them (i.e. the average of the first and the second channels minus the average of the third and the fourth ones).

III. NOISE BUDGET

The key element of each channel is the DDS, whose residual phase noise represents the ultimate limit for what concern the performance. We used the AD9912 from Analog Devices. It is a 48-bit 1 GHz DDS, with an output resolution of 14 bits. It was characterized in detail in our previous work [4]. Here we repeated the measure: the output of two DDSs with the clock in common mode are measured with a commercial phasemeter. The output frequency spans from 250 down to 1.95 MHz as shown in Fig. 2.

The AD9912 exhibits pure x-type noise characterized by a jitter of about 700 fs and a flicker of 5 fs, or, that is the same for 100 MHz, the frequency at which we use our system, by \( b_1 = -110 \) dBBrad and \( b_0 = -154 \) dBBrad/Hz. The Allan deviation is \( 2.1 \times 10^{-14} \) at 1 s (\( f_0 = 10 \) Hz).

At 100 MHz, the other components have a negligible noise: the flicker of the mixer increases, because the noise of the DDS is lower. The possibility to measure a wide range of frequencies allows, on the other hand, to use an external pivot oscillator to furtherly reduce the contribution of tracking DDS. Finally, the internal local oscillator has the possibility to be phaselocked and is externally available. It runs at 1 GHz and, if necessary, it can be directly multiplied to 10 GHz, allowing a direct comparison with the ultra-stable source(s) without any significant degradation. These features are particularly interesting to generate a real-time composite clock.

IV. RESULTS

To characterize the phasemeter and the ultrastable signal, in addition to the Allan variance (2), we used the 2-sample cross-variance (3) [5]. In the following, for brevity, we will omit the term 2-sample.

\[ \sigma^2_y(t) = \frac{1}{2(M-1)} \sum_{i=1}^{M-1} (y_{i+1} - y_i)^2 \]  

\[ \sigma^2_z(t) = \frac{1}{2(M-1)} \sum_{i=1}^{M-1} (z_{i+1} - z_i)(z_{i+1} - z_i) \]  

The cross-variance has the same formula as the Allan variance, except that it considers two signals: \( y \) and \( z \) instead of one. In the case, the two signals are uncorrelated then the cross-variance tends to zero when the number of samples \( M \) increases. In this manner, it is possible to reduce the contribution of the noise of the references and of the instrument in the final measure. Fig. 3 reports a classical scheme, where only the signal of interest \( c \) is squared, because it appears in both signals \( y \) and \( z \).

A. Resolution of the phasemeter

We tested the phasemeter by measuring the outputs of a 6-channel hybrid power splitter. The phase noise on each channel is the same and cancels by considering the difference of two channels, being in common mode. Data has been accumulated and then decimated to obtain a rate of 20 Sps and a measurement bandwidth of 10 Hz. Fig. 4 shows the Allan deviation of \( x_1^a - x_2^a \) (blue) and \( x_3^a - x_4^a \) (green curve). By subtracting 3 dB, we estimate the

![Fig. 2: residual phase noise of the DDS AD9912 clocked at 1 GHz with the frequency output that spans from 250 MHz down to 1.95 MHz](image)

![Fig. 3: block diagram of the measures done to characterize the phasemeter and the comb referenced to an ultra-stable laser](image)
resolution of a single channel: $2 \times 10^{-14}$ and $2 \times 10^{-17}$ at 1 s and 3000 s respectively. By using 4 channels, it is possible to calculate the cross-deviation between $x_1^y - x_2^y$ and $x_3^y - x_4^y$. In this case, the resolution improves by a factor 10 at 1 s where it reaches the interesting value of $2 \times 10^{-15}$, ten times better with respect to the one obtained with the Allan deviation. We notice that the improvement is proportional to $\sqrt{M}$ and scales as $\sqrt{\tau}$. Hence, it is higher for short $\tau$ and with longer measures.

B. Noise of the comb

We measured the comb with respect to four active hydrogen masers. The signal at the output of the comb has been split to feed two channels. This because the noise of the comb was expected to be of the same order of the single channel one. The masers, instead, have been connected without splitting, being their noise higher.

The red curve represents the comparison of maser 3 with maser 4. It has a typical behavior: at 1 s, the stability is about $10^{-13}$ and goes down, reaching $10^{-15}$ at 1000 s. The curves blue and green show the comparison of the comb with respect to the two masers. For measurement time greater than 10 s, they represent the noise of the comb: we can see the effect of the temperature of the laboratory that induces a bump at 200 s and then the effect of a $-7.4 \times 10^{-11}$/day drift. On the other hand, for measurement time lower than 10 s, they represent the stability of the masers, which is $1 \times 10^{-13}$ at 1 s for both of them. The cyan curve, instead, is the cross-deviation of the comb with respect of the two masers. Being the noises of the masers uncorrelated, their product averages down and the noise of the comb arises: there is a flicker phase noise at the level of $2 \times 10^{-14}$ at 1 s, then a floor of $1.3 \times 10^{-14}$ followed by the effect of the temperature and of the drift as we have already seen. The test report of the frequency comb shows a flicker PM noise $b_f = -109$ dB/6 which is in agreement with what we measured. To conclude, the frequency comb itself limits the ultra-stable signal in the short term while, for $\tau > 2$ s, the noise of the ultra-stable laser dominates.

V. CONCLUSIONS

We implemented and characterized a 6-channel synchronous phasemeter able to measure frequencies from 5 up to 400 MHz. It has been characterized at 100 MHz where it shows a resolution of $2 \times 10^{-14}$ at 1 s, limited only by the residual phase noise of the DDS we used. Thanks to the 2-sample cross-variance, we were able to increase the resolution of the instrument by a factor 10 at 1 s as well to compare an ultra-stable signal with respect to two references whose noise is higher. This allowed us to measure the stability of the 100 MHz signal generated by our frequency comb locked to an ultra-stable laser, without having other RF ultra-low noise signals, but only active hydrogen masers. The architecture of this system is based on FPGA and is very flexible. We plan to use it for many other applications in the field of time and frequency, by example in the frame of time-scale generation and laser locking.

ACKNOWLEDGMENT

This work has been funded by the EMRP program (IND55 Mclocks). A special thanks to all members of the Go Digital working group, at FEMTO-ST and INRIM, and, in particular, to E. Rubiola.

REFERENCES

Time Signals Converging within Cyber-Physical Systems*

Marc Weiss
Time and Frequency Division, NIST
Boulder, CO, USA
E-mail: mweiss@nist.gov

Sundeepe Chandhole
Research & Development
National Instruments, Austin, TX, USA

Hugh Melvin
National University of Ireland
Galway, Ireland

Abstract—Time is central to predicting, measuring and controlling properties of the physical world, and is one of the most important constraints distinguishing Cyber-Physical Systems (CPS) from distributed computing in general. However, mixing the cyber and the physical presents a fundamental challenge, since computers and communications systems have abstracted away the physical layer and timing is fundamentally a physical signal. While such abstractions have yielded significant benefits, time has been a casualty. CPS used in industry today achieve time-awareness by making use of time-aware fieldbuses and devices with specialized proprietary software. However, this approach has proved restrictive in both the topologies achievable and the scalability of networks beyond a certain size. The new era of the Internet-of-Things and the Industrial Internet is paving the way for convergence, where time needs to be an integral part of the cyber, making integration of cyber and physical seamless. However, this requires successful research in a number of different areas.

The National Institute of Standards and Technology (NIST) has formed a CPS Public Working Group (PWG), with members from global industry, academia and government. This CPS PWG is tasked with creating a set of frameworks and reference architectures for CPS, to promote proper function and interoperability. Public documents from this effort will soon be available. We discuss the timing section of the CPS PWG document and focus on the status of challenges and efforts to integrate time-sensitive with best-effort processes in CPS nodes and the networks that connect them.

Keywords—Cyber-physical systems, time-stamp, Internet of Things, time-sensitive networks

I. INTRODUCTION

We stand at the advent of a revolutionary new economy fueled by the global Internet of Everything, IoE, a combination of the traditional telecom system with its growing need for wireless technology, and the emerging Internet of Things, IoT, [1] [2], including Machine-to-Machine (M2M) technology [3]. Cisco, among others, predicts that there will be a trillion endpoints connected to the Internet by 2022, with $14.4 trillion in value at stake [4]. General Electric, GE, says "about 46% of the global economy or $32.3 trillion in global output can benefit from the Industrial Internet” [5]. The National Institute of Standards and Technology (NIST) has formed a Cyber-Physical Systems (CPS) Public Working Group (PWG) to bring together experts to help define and shape key aspects of CPS, and to create a framework and reference architectures to encourage interoperability and appropriate designs [6]. One fundamental enabler of this revolution will be a better marriage of timing signals and data that otherwise will limit this growth. Currently, optimal use of data in computing and networking is anathema to optimal use of timing signals. Computer hardware, software and networking all isolate timing processes, allowing the data to be processed with maximum efficiency due in part to asynchrony. Yet, coordination of processes, time-stamping of events, latency measurement and control, and optimal use of precious spectrum are enabled by timing.

Timing is critical for the future development and improvements to several current high value applications. For example, smart transportation involving the exchange of information between vehicles, highways, and perhaps civil authorities will depend on a robust ubiquitous timing system to ensure the availability and integrity of the data. Similar requirements are found in the operation of the power grid, especially now that wind farms, solar arrays and the like, which will require different control strategies, are becoming an important part of the system. Medical applications such as tele-surgery, and regulating fairness in financial systems are other important examples.

II. NIST CPS PUBLIC WORKING GROUP

In 2014, NIST convened the CPS PWG with a kick-off webinar in June and a face-to-face meeting in August. This grew out of a recognition that, while companies are already building CPS, there lacks a unified technical foundation for broad collaboration. Missing are a consensus definition and taxonomy, reference architecture, and a shared understanding of the essential roles of timing and cybersecurity. The good news in the CPS field is that there is substantial growth of applications in many sectors ranging from energy to health, disaster resilience, transportation, manufacturing, building management, and others. However, these deployments are often sector-specific and are not designed for interoperability.

* Contribution of U.S. government not subject to copyright
across sectors. Further, individual communities, states, and countries are implementing their own, unique solutions that are also not designed for interoperability with their neighbors. The resulting landscape of isolated, legacy systems will only continue to grow, making solutions to create interoperability only more difficult with time, and thus limiting the potential benefits of CPS.

The increasing complexity of a 21st century society demands systems-of-systems solutions that require integrating CPS across domains and at multiple scales. This requires developing a common technical foundation that will enable us to work together to achieve this potential. That’s the goal of the CPS Public Working Group.

Participation in the PWG is open and free to everyone, anywhere in the world. Most of the sub-group work is done in virtual meetings and using web collaboration tools allowing participation from anywhere. All of the products of the PWG will be openly available online to anyone. The output of the PWG to the public will be two documents developed in sequential phases: a CPS framework that describes best practices and options using current technology, and a CPS Technology Roadmap identifying opportunities for a coordinated effort on key technical challenges. The CPS framework will be released as a draft for public review soon, in the spring of 2015, from [7].

The PWG is organized into five subgroups each of which is led by a collaboration of three co-chairs: one from each of NIST, academia, and industry. The five subgroups are reference architecture, use cases, cyber-security, timing, and data interoperability.

The timing part of the CPS framework document consists of three major sections. First, the time-awareness section examines the components of a CPS from the perspective of the presence or absence of explicit time in the models used to describe, analyze, and design CPS and in the actual operation of the components. Next the time and latency section addresses the use of time to provide bounded latency in a CPS. Thirdly, the section on secure and resilient time addresses the special security problems associated with timing.

We focus in this paper on the time and latency section, discussing the need for and status of convergence between time-sensitive and best-effort processes in CPS nodes and interconnecting networks.

III. TIME AND LATENCY IN CPS

The aim of this section in the CPS PWG framework is to provide reference architectures/frameworks that enable building time-aware CPS to solve control and measurement applications.

Given the diversity in CPS applications and scale, it is not surprising that temporal considerations vary considerably over the range. For example, in small closed systems such as a packaging machine, the primary temporal concern is that all components respect a self-consistent timing design. In such systems, networking temporal considerations, e.g. design of a TDMA scheme, are part of the design itself. However in large scale, and more critically, in environments characterized as “System of Systems”, timing issues are more difficult. For example “smart highways” will involve many different systems, some in the vehicle, some in the infrastructure, some in a traffic management center, etc. Each will have its own temporal requirements which must be met while sharing network bandwidth and in some cases computation bandwidth on servers. Many technological challenges remain in managing the timing in such systems. The remainder of this section discusses both the general issues as well as some of the current thinking on these issues. Some of these can be applied to smaller systems. There is no doubt that the work on larger systems will result in improvements, e.g. in time-sensitive network technology, that will make small system temporal design much easier and more robust.

CPS are used in both control and measurement applications. The requirement of bounded latency is obvious in control systems where the latency from when a physical input is read to when a physical output is written has to be proven by timing and schedulability analysis. In large-scale control systems this requirement becomes even more challenging since the input, computation and output may be occurring on different nodes that are spatially distributed. The challenges of predictability in software are added to by the non-determinism provided by layers of software managing data-transfer on the network connecting these nodes. As the scale of CPS expand to Systems of Systems, the impact on timing of Cloud Computing and Networking concepts such as Software-Defined Networking (SDN) and Network Functions Virtualization (NFV) need to be carefully considered.

In CPS-based measurement systems, the deterministic relationship between acquired data (e.g. simultaneity) is of paramount importance. However, what is typically overlooked is the efficiency and complexity of transferring the acquired data from thousands of nodes to one or more aggregating units, where analytics or logging is being performed. Misaligned data can result in faulty conclusions. In many CPS-based applications, the data measurements are used for asset or structural-health monitoring and in many cases a timely response based on real-time analytics is required. Time, when applied to data-transfer can enable bandwidth reservation in networks used in these measurement applications, thereby enabling faster analytics, a smaller memory footprint, and increased efficiency in data-reduction techniques (for logging). Moreover, bounded latency is extremely useful in distributing triggers to multiple nodes inside a CPS.

Similar to CPUs, computer networking has traditionally been optimized for “best effort delivery”, and that has worked extremely well in the past and will continue to do so in the future for many uses. However, a challenge exists when the same networking technology is used for time-sensitive applications that are served by CPS. There is much work being done for enabling time-based CPS, using standard Ethernet technologies to enable seamless integration with the Internet. This “Time-Awareness” in standard Ethernet is paving the way to enable time-sensitive (bounded latency) traffic to coexist on the same network as traditional best-effort (no latency guarantees) traffic. Further details of this work relating to
networks, FPGAs and computers can be found in [8] [9] [10] [11] [12] [13] [14].

A. CPS Domain and Network Managers

A time-aware CPS should guarantee bounds on latency of data delivery and guarantees on synchronization accuracy as it applies to timing correlation of physical I/O. To build such large-scale systems with these guarantees the following two concepts of CPS Domain and CPS Network Manager (CNM) are defined.

CPS Domain: A CPS domain is a logical group of CPS nodes and bridges which form a network with their own timing master. The master may synchronize to a globally traceable time source (e.g. GPS). Each CPS domain has its own primary (or self-consistent as described earlier) time-scale. This time-scale provides a strong monotonically increasing clock to applications for performing input/output functions and time-based scheduling. The timing master of a CPS domain should not produce a discontinuity of time once time-sensitive data transfer within the domain has commenced, even if the master loses connectivity to its global source (e.g. GPS) sporadically.

If a global traceable time is required inside a CPS node, then the node can implement a second time-scale called the Global Traceable Time-Scale. This time-scale can be managed independent of the CPSs primary Time-scale. To correlate the CPS’s primary time-scale to the Global Traceable Time-Scale, the offset of the primary time-scale from the Global Traceable Time-Scale can be maintained at all times by the CPS node. The Global Traceable Time-Scale can be used to correlate CPS Time-Scales from multiple CPS domains. This is illustrated in Fig. 1.

Many CPS will be small enough that they don’t need an external time-scale and the primary time-scale will suffice. However, significant benefits can accrue from such systems being, and some level of traceable timing may be available, though perhaps not at the needed stability or accuracy.

• Control and manage the state of all CPS nodes in a CPS domain.
• Coordinate with a centralized network controller to configure bridges in a CPS domain.
• Configure transmission schedules on CPS nodes
• Monitor the health of the CPS domain (for handling errors, changing schedules and bringing new CPS nodes online, etc.).
• Configure application and I/O timing on each CPS node
• Configure any static timing requirements for time-based synchronization

Fig. 1. Domains and Multiple Time-scales in Time-aware CPSs

The functions of a CNM vary depending on the size of the system. These functions include:

Fig. 2. CPS Network Manager configuring a CPS

Either the CNM or the centralized network controller has to gather performance metrics and determine the topology of CPS nodes in a CPS domain in order to create a schedule. The relevant performance metrics include Bridge Delays, Propagation Delays, and Forwarding/Transmission delays. There are multiple ways to detect topology. For example, one approach to Software Defined Networking (SDN) defines a “Packet-In” “Packet-Out” protocol which uses Openflow [14] with Link Layer Discovery Protocol (LLDP) [15]. Some other protocols like PROFINET [16] use Simple Network Management Protocol (SNMP) [17] along with LLDP. The Centralized Network Manager computes the topology for the CPS domain using these mechanisms, and determines the bandwidth requirements for each time-sensitive stream based on application requirements. The bandwidth can be specified by the period and the size of the frame. Optionally the application can also specify a range <min, max> for the offset from start of a period. This information is provided to the Centralized Network Controller. The Centralized Network Controller computes the path for the streams and gathers performance metrics for the stream (latency through the path and through the bridges). This information is then used to compute the schedule for the transmission time of each time-sensitive stream and the bridge shaper/gate events to ensure that each time-sensitive stream has guaranteed latency through each bridge. Additionally, queues in bridges are reserved for each stream to guarantee bandwidth for zero congestion loss. It should be noted that schedulability analysis and computation
is the subject of continuing research as the problem becomes intractable for large systems.

It should also be noted that there is considerable activity in the IEEE 802.1 and other standards communities in providing additional tools for controlling network temporal properties.

B. Converging Time-Sensitive and Best-Effort Processes

Many CPS nodes will need to combine time-sensitive with best-effort processes. Such a time-aware node will have separate streams for the two types of data and applications. An illustration of a possible device model for a time-aware CPS node is shown in Fig. 3.

![Fig. 3. Time-Aware CPS Device Model](image)

The physical layer receives data units from the data link layer and encodes the bits into signals and transmits the resulting physical signals to the transmission medium connected to the CPS node. If the physical layer supports a time stamp unit (TSU) then its management interface should be connected to the data link layer so that a time stamp can be retrieved as and when required by the timing and synchronization protocol (e.g. IEEE Std. 1588TM [10]).

The data link layer provides time-sensitive data communication among devices in a CPS domain. The data link layer implements a set of dedicated buffer pairs (Tx and Rx queues) for time-sensitive data. At a minimum two pairs of buffers are required so that time sensitive data can be managed independently from best effort data. The time-sensitive transmit buffer is connected to a scheduled (time-triggered) transmit unit. This unit uses a schedule provided by the CPS Network Manager and reads data from the application and copies it into the time sensitive transmit frame and transmits the frame on to the CPS domain.

- The application layer consists of two parts:
  - Application-support protocols: These are the protocols that support the conveyance of time sensitive data at the user’s application level.
  - Time-Sensitive Data Mapping: Protocol to manage the mapping of application data to time sensitive data exchange frames between devices. An example can be CANopen [18] which is used as a data-mapping protocol by multiple industrial protocols.
  - Best-Effort protocols: Used for standard internet access, non-time-sensitive streams.
  - Timing and Sync Protocols: These include protocols which propagate synchronized time from the network to the application (including I/O functions). Some examples of such protocols are IEEE 1588, IEEE 802.1AS [19], etc.
  - User application: User defined applications accessing time sensitive and best effort data, and time-sensitive I/O interfaces to allow decoupling of logical and physical time with enforcement only at the boundary to physics. An example of a realization of this capability is inherent in the design of the Texas Instruments DP83630 Ethernet PHY1.

Currently time in a CPU is implemented via time-stamp counters (TSC) that increment time using the local clock driving the CPU. This clock does not maintain network time. The TSC can be disciplined via software to slave it to network time. However this leads to significant loss of precision and accuracy. For CPS nodes that synchronize to a single external clock source, it may be desirable to have the TSC driven directly by the network time. This may be implemented by linking the registers of the TSC with the timekeeper in the network interface or by providing a common time-base which can be atomically captured by the network interface before propagating the network time to the CPU or any peripheral device. More generally, CPS applications may choose to maintain offset/PPM state for each derived clock and translate on-the-fly as needed without physically disciplining the TSC. This is especially useful in cases where the applications care about multiple time sources.

Languages used for modeling and programming of time-aware CPS need time as a fundamental programming semantic. Time in the language is required when interfacing to physical I/O and the network. Functions that take future time events to read physical inputs and write physical outputs can enable coordination of physical I/O with scheduled data on the network. Additionally, time-triggered loops can enable...
coordination of logic execution with schedule of transmission of data. PTIDES1 [20] and LabVIEW1 [21] are two examples of system design tools which implement these time-based programming semantics.

CPS can employ operating systems with a wide range of complexities, from a simple application-level infinite loop (e.g. the Arduino platform) to a virtual machine hypervisor running several instances of virtualized systems on a multi-blade, multi-core hardware platform. The issues that arise throughout these systems with respect to time-awareness are how to get time to the application with a bounded latency and with accuracy, and how to schedule tasks with time accuracy and bounded latency.

At the application layer, the introduction of explicit time will have a profound impact on the conception, design, execution, and robustness of CPS applications. This is a very active area of research, but there are hints of things to come. For example the concept of decoupling of logical and physical time with enforcement only at the boundary to physics mentioned above has yet to be fully exploited. In some cases, tradeoffs can be exploited by applications between message passing, which consumes network bandwidth, and reasoning about timestamps, which can in some cases eliminate some of the messages. An example of this in database management is the Google Spanner system [22].

Building CPS using the above mentioned techniques will make it easier to characterize systems, which is a key requirement of safety-critical systems. CPS with scheduled converged networks built with FPGAs and time-aware CPUs will provide static guarantees and always satisfy timing requirements for their time-sensitive traffic. Architecture-specific analysis tools can derive these guarantees in the form of upper and sometimes also lower bounds on all execution times, since time is foundational in all elements of the CPS.

C. Needed research

We identify a number of areas where research on timing is needed to ensure that the full potential of CPS is realized.

Further research in languages used for modeling and programming time-aware CPS is desired that will allow an application written on a CPS node to be represented as one or more timed-functional modules which can be shifted in time by a Schedule Generator to align production and consumption of timed-functional modules which can be shifted in time.

Increasing precision of timestamps will not only improve application timing but allow better utilization of bandwidth on the network. Currently, asymmetry of delay in networks and phase errors due to asynchronous clocks driving the transmissions on a network are the major causes of inaccuracy in time transfer. Research aimed at increasing precision of timestamps by enabling new hardware and software methods to correct these asymmetries and phase errors will improve clock accuracy by an order of magnitude.

Better precision time-based synchronization [23] in IEEE 802.11 will enable time-awareness in wireless access points and stations. Research into mechanisms that use these synchronized clocks to create a TDMA-based scheme that can coexist with best-effort traffic similar to wired Ethernet (802.1Q) will enable reduced cost of infrastructure for many CPS applications.

WANs currently offer QoS over dedicated connections for all forms of real time communication (RTC) such as audio and video streaming. Developing synchronized time and the methods described in this paper, the same QoS would be possible over standard networks, thereby reducing costs and increasing accessibility.

IV. CONCLUSIONS

The expected massive growth in the new Internet of Things, encompassing Cyber-Physical Systems and the Industrial Internet, will require a convergence of time-sensitive systems with best-effort systems. Much work is already underway, though new research remains, which will require collaboration among different fields. The extent to which these timing challenges are met and surpassed will dictate the success of emerging CPS applications and others as yet unheard of.

REFERENCES


1 Commercial equipment is identified in order to describe a concept adequately. It is not intended to imply recommendation or endorsement, nor is it intended to imply that these entities, materials, or equipment are necessarily the best available for the purpose.


ns-level time transfer over a microwave link using the PTP-WR protocol

Ecole Polytechnique Fédérale de Lausanne, Electronics and Signal Processing Laboratory, Neuchâtel, Switzerland
mathieu.rico@epfl.ch; cyril.botteron@epfl.ch; jean-pierre.aubry@epfl.ch; pierre-andre.farine@epfl.ch

Abstract—today it is possible to achieve sub-ns level time synchronization on a wireline network while only us-level synchronization can be achieved on a wireless (microwave) link. In this paper we will, first, study the performances of different time synchronization wireline based protocols, such as Precision Time Protocol (PTP), Synchronous Ethernet (SyncE) and PTP White Rabbit (PTP-WR). And then, we will present our results using a wireless link, and determine which radio technology can achieve ns range time synchronization. Our motivation is to qualify a time transfer process operating over microwave link and offering secured GNSS-like time performance.

Keywords—microwave time transfer; PTP; IEEE1588; SyncE; PTP-WR,

I. INTRODUCTION

An increasing number of infrastructures require an accurate time reference, such as telecommunication base stations, energy industries (smart grid synchronization), and electronic intelligence systems. A GNSS receiver can provide an accurate time and frequency reference, but it requires a line-of-sight link to the satellites, and it has acknowledged vulnerabilities [1-4]. Therefore in some critical applications, a “GNSS-free” reference becomes mandatory [5].

Our goal is thus to identify a “GNSS-free” time transfer technology able to provide “some ns” level accuracy, traceable to UTC or to a private time reference, and to disseminate such a secure and accurate time towards a fixed installation, ELINT or SIGNIT. The transport media should be wireless microwave on medium distance (10-50 km) to avoid fiber deployment within a given territory.

We have selected three time transfer protocols available on wireline: Precise Time Protocol (PTP); PTP with Synchronous Ethernet (PTP-SyncE); and PTP white rabbit (PTP-WR) [10-14], which is an enhanced version of the PTP protocol [9] introduced by CERN providing time transfer in the sub-ns range over optical fiber links. Other physical approaches [7,15-16], involving GNSS based, SDH specific time stamp or amplitude modulation have been reported for time transfer. However they were not selected at that stage, because these techniques require specific HW/SW and do not seem ready for standardization. On the other hand, there is a significant amount of work dedicated to frequency accuracy and frequency transfer over fiber, mainly for primary clock comparison purpose [18-19]. These experiments are targeting the ultimate performance on frequency, and they can support dedicated HW/SW. Nevertheless, PTP-based protocols are preferred here because of equipment availability, Ethernet and IP network compatibility, and the launched standardization process of PTP-WR [17].

Currently, there is an increasing interest for PTP-WR over fiber, for short distance (less than 10km), as well as some very interesting developments on PTP-WR over long distance fiber [24]. However, while there exists some literature on PTP over wireless (see, e.g., [27]), there is not much published work yet on PTP-WR over wireless.

Therefore, our plan of work will be to confirm PTP-SyncE over wireless, PTP-WR over wireline, and then deploy PTP-WR over wireless, our final goal.

A. Time transfer criteria

Our purpose is to provide time transfer accuracy using, as much as we can, existing devices and protocols, already or close to be standardized.

The key criteria regarding time transfer functionality is accuracy, stability and precision (using wording definition of VIM, International Vocabulary for Metrology [6]). For the operational point of view, a time transfer process must be easy to calibrate and such calibration must be precise (in the sense of fidelity). Furthermore, that means that calibration must survive any switch off / switch on operation.

B. PTP principle

The PTP protocol, known as IEEE 1588, is an Ethernet based protocol, mostly introduced by Prof. Weibel [9], which is widely used in telecommunication and automation infrastructures. It can provide both time synchronization and frequency syntonization, and aims to provide sub-us level accuracy.

PTP protocol uses a “two-way time stamping” (similar as the one illustrated on the left of Fig. 2). Knowing those timestamps, the slave can identify its clock offset. The “two-way time
stamping” process is used also in other protocols such as NTP (Network Time Protocol), and PTP-WR. In PTP the propagation channel is assumed to be symmetrical, and time stamping can be done either in software or hardware (more accurate). PTP can be deployed on a classical IP network, but performance are highly improved by using PTP compliant intermediate devices such as Boundary Clock (BC) or Transparent Clock (TC) switches.

C. SyncE concept

Synchronous Ethernet is a standard that provides syntonization through an Ethernet network. It uses Ethernet idle patterns to encode a frequency reference on the physical layer. Specific hardware is needed to recover the frequency reference by using a PLL and all network devices must be SyncE compliant. Those constraints imply an upgrade of the whole network. It can achieve a ± 4.6ppm syntonization accuracy (ITU G.8262) (when the clock is not driven by an external reference) whereas standard Ethernet (IEEE 802.3) allows ±100ppm.

Therefore, the syntonization process through SyncE might be highly efficient. SyncE is physically generated by 8B/10B conversion (i.e., adding one bit every 4 bits), avoiding too many consecutive 1 or 0’s, and allowing to get a physical signal reference when data are processed and also when there is no data to convey. SyncE is a “master slave syntonization”, like the good old time of SDH, while PTP is a two way process.

D. White Rabbit presentation

PTP-WR consists of three main aspects: it employs PTP to exchange timestamps and coarse synchronization; SyncE for syntonization; and a Dual Mixer Time Difference (DMTD) [10, 14]) for fine estimation synchronization. PTP-WR has a more accurate delay model which takes into account the hardware delay and the asymmetry of the propagation channel.

PTP-WR operation is summarized on Fig. 2:

A SyncE link is established forth and back, a first step synchronization using PTP is then performed (coarse acquisition) and a final step is applied (see Fig. 3), using the D.DMTD (Digital DMTD [14]) embedded in master and slave to perform a very accurate (sub wavelength) of the phase-time offset, by comparison between the phases of the outgoing optical signal and of the return signal (SyncE loop back).

The PTP-WR delay link model, as define over wireline propagation, is defined on the following graph:

Under wireless media, the “link” delay $$\delta_{\text{in}}$$ and $$\delta_{\text{out}}$$ will take into account the propagation delay and the time for media conversion interface [18, 19].

E. Protocoles comparison

PTP works only at the level of “clock” synchronization, time stamp, detecting the time of flight and assuming a ToF symmetry. PTP-SyncE uses “one way syntonization”, operating on a synchronous link, allowing a better time adjustment because of a common frequency reference. PTP-WR uses a two way syntonisation, applies PTP “coarse” synchronization step, the two ways SyncE and the two D.DMTD allow to perform very accurate “phase detection”.

F. Goals and Methodology

Our goal is to identify a time transfer technology adequate to disseminate some ns level accuracy over a wide area from a sync master to fixed equipement users. Having selected an IP based generic solution, we will then compare the behavior of the various IP variants (PTP; PTP-SyncE; PTP-WR) and we will also compare the behavior of each technology over various propagating media (wire, fiber, microwave).

A wireless IP packet radio introduces flight delay & jitter, due to:
- Modulation scheme (X-PSK, X-QAM…), data rate (PTP is a 166 bytes process) due to latency and jitter in modulation / demodulation.
- Frequency/Time selective channel.
- Time or frequency spectrum allocation: FDD or TDD.
- Half duplex / full duplex configuration.
II. **TIME TRANSFER OVER WIRELINE**

**A. PTP over wireline**

For comparison purpose, we have done some time transfer tests over wireline/fiber, using PTP, PTP-SyncE and PTP-WR. The devices we used to qualify the PTP protocol are a PTP80 Grand Master Clock and PTP Slave from TIME & FREQUENCY Solutions. The tests were done on a 100base-TX RJ45 8-meter copper cable after a transient phase of 2 hours to let the master stabilize its OCXO.

Fig. 5 shows the results we obtained with PTP over wireline.

On steady state, we can get an average pps (pulse per second) delay of 47 ns and a standard deviation of 3.7 ns.

Those results seems to reach the lower bound of PTP performance, are well below the PTP protocol specification.

**B. PTP-SyncE over wireline**

PTP-SyncE was tested on a 5 km optic fiber. The test results are shown below:

SyncE over 5 km optical fiber (using optical SFP) shows a “bimodal” status with 2 lobes offset by around 10 ns, which is close to one period of the SyncE carrier signal. The distribution is not Gaussian, and the “weighted average time offset” is about of ~2 ns, while the dispersion (min/max range) is no more than 10 ns.

Thus it seems that, at that stage, while PTP seems to be limited just below 100’s ns range, the PTP-SyncE over fiber may reach some ns time transfer accuracy, and is able to exhibit a quite good stability.

Fig. 7 shows the initial time transfer series of Fig. 6. One can notice the initial transient time of the pps delay (in blue) oscillating within +50 ns and -60 ns, and stabilizing after 200 s.

After that, the oscillation remains within a 10 ns interval centered on -2ns. This oscillation seems to be generated by +/- 1 cycle of the SyncE carrier. This behavior could be a consequence of the hardware implementation of the PTP protocol on the PTP-SyncE master and slave devices we used.

**C. PTP-WR over wireline**

The PTP-WR devices are supplied by 7solutions, and consist in white rabbit switches and spec boards (PCI cards). We have performed two sets of characterization. One with a short fiber link (2m) and one using a 5 km long fiber. The following test results displayed on Fig. 8 were obtained by using PCI “spec boards” as master and slave. The distribution in each case is clearly Gaussian and the performance are well far under the sub-ns range, with an average delay of 120 ps and a standard deviation of 20 ps.

**D. Time transfer noise analysis tools**

The Gaussian shape in Fig.8 shows that we are facing mostly white phase modulation (PM) noise, and this noise can be averaged out with integration time. The record of successive (# 6000 points) pps offsets using WR over fiber is given in Fig. 9.

Since late 90’s, we know that the classical standard variance, \( \sigma^2(\tau) = \langle (y_i - \langle y_i \rangle)^2 \rangle \) is not applicable on clock signals because
of the environmental and inherent frequency drift, meaning that the average $<y_i>$ may drift with time, and the calculation will not converge with time and/or increasing number of samples. This is why D. Allan introduced the so called Allan variance [22], defined by the following equation:

$$\sigma^2_y(\tau) = (1/2)(y_{i+1} - y_i)^2$$

The Allan variance removes drift and is able to converge under many conditions. Such analysis provides information on the shape of the noise contribution, and linear “asymptotic” fit in terms of $\sigma^2_y(\tau) = \sum k_0 \tau^q$ can be used to define the noise behavior. Allan variance is not able to isolate white PM from flicker phase modulation, and a modified Allan variance was later introduced [23]. When using time difference rather than relative frequency variation, one has to replace $y(t)$ in the previous equations by: $y(k) = (x_{k+1} - x_k)/\tau$, $\tau$ being the equally spaced measurement time interval.

In these formulas, we are using the classical definition of $x(t)$, $y(t)$ and $\phi(t)$, i.e. $x(t) = \phi(t)/(2\pi\nu_0)$, $y(t) = (1/(2\pi\nu_0))d\phi(t)/dt$, and $y(t) = dx(t)/dt$. Overlapping samples was introduced to increase the time base of analysis despite a low number of samples. When computing samples for any $\tau = m\tau_0$ from time series collected every $\tau_0 (1s)$ we get [23] an equivalent expression:

$$\text{Mod.}\sigma^2_y(\tau) = (2m^2\tau^2(N - 3m + 1))^{-1} \sum_{j=1}^{N-3m+1} \left( \sum_{i=j}^{j+m-1} (x_{i+2m} - 2x_{i+m} + x_i) \right)^2$$

Samples are collected every seconds, and aggregate are calculated at overlapping $\tau = m\tau_0$ with $m = \{2, 5, 10, 20 \ldots \}$. In Fig 10, the “Modified Allan variance vs time” $\sigma^2_y(\tau)$ of the time offset provided by PTP-WR over fiber exhibits a $\tau^{-3}$ slope. This is in agreement with similar WR over fiber analysis done by CERN team [26]. The time transfer process seems to be mainly affected by white noise PM.

![Fig. 10: Mod.Allan variance time offset: $\sigma^2_y(\tau)$ PTP-WR over wireline – blue curve: raw data – red: $\tau^{-3}$ slope](image)

This preliminary comparison between the various protocols and variants allowed us to qualify each protocol capability.

### III. TIME TRANSFER OVER WIRELESS

As in the wireline case, we want to compare the time transfer behavior between the PTP variants (PTP, PTP-SyncE and PTP-WR) over various radio link configurations.

![Fig. 11- Time transfer over wireless test configuration](image)

The different radio link configurations are defined by the following parameters:

- Modulation scheme and rate
- FDD or TDD spectrum occupation
- Carrier frequency and bandwidth
- IP radio configuration: PTP compliant, PTP “transparent clock”, one way SyncE, two way SyncE

The impact potentially generated by the microwave link might include some packet losses and a significant impact on Packet Delay Variation (PDV) at radio interface, coming from the conversion delay from wireline media to microwave, and queuing jitter.

#### A. PTP over wireless

To qualify PTP over wireless, we used the Cambium PtP 650 radio device [20], which provides several interesting features such as TDD (time duplex division), SyncE and PTP transparent clock feature (TC).

1) Impact of the Transparent Clock feature

With the TC feature disabled, we obtained when considering 40’000 pps delay measurements an average delay of 1.8 us and a standard deviation of 867 ns, which is not within the PTP specifications. However, with the TC feature enabled (see Fig. 12), we obtained more interesting results: a time synchronization in the sub-100 ns level with an average delay of 50.5 ns and a standard deviation of 6.6 ns. Note that a transient phase (see the circles on Fig.12) can also be observed. Such performance can be explained by the process employed by the PTP TC device. Actually, a TC device adds automatically residence time data to every PTP packet passing through, thus the PTP slave can take into account the jitter due to random queuing of the PTP packet in the network device.

![Fig.12. PTP over TDD - TC enabled: <delay> 50.5 ns – std dev 6.6 ns](image)
2) Modulation impact BPSK vs. QAM

The microwave link used previously can be adapted from 256QAM (previous configuration) to BPSK. The observed performance with PTP, no SyncE, TC, is provided in Fig. 13:

![Fig. 13: PTP over BPSK modulation – <delay> 67ns – std dev 9ns](image)

From this figure, we measured with BPSK modulation a mean offset of -38 ns and a standard deviation of 9 ns, very similar as under 256QAM configuration (for which the average delay was 50.5 ns and a standard deviation of 6.6 ns).

3) FDD versus TDD

The AirFiber module from Ubiquity allows us to configure the access technique either in FDD or TDD. In both conditions we observed a transient initial drift “time to set” greater than 1 μs, up to a thousand seconds, and a steady state offset and fluctuations of some hundreds of ns.

At that stage, we can conclude that PTP over wireless may allow a time sync less than 100 ns with a dispersion below 10 ns, but only with the TC enabled configuration. No significant impact of radio link configuration (FDD/TDD, BW, modulation, data rate…) is observed under PTP operation.

B. PTP-SyncE over wireless

We have performed time offset measurement based on PTP-SyncE over a FDD microwave link provided by Bridgewave [21].

![Fig. 14: PTP SyncE over FDD – average delay 18ns – std dev 4.5 ns (calculated on the first 135000 points)](image)

The 240 ns pps delay jump (circled in Fig. 14) at 135’000th event is not yet understood. Domains before and after the delay jump, present a “Gaussian” population with a standard deviation below 5 ns. On each domain the time transfer offset exhibits two stable states separated by 10 ns (shown on the yellow histogram).

We have analyzed the 40’000 initial points along with Mod.Allan variance.

![Fig. 15: PTP-SyncE over FDD - pps offset: Mod.σ(τ)² plot](image)

In Fig 15, the Mod.σ(τ)² plot shows a τ⁻¹ slope between 10 and 1000s, indicating a white phase modulation contribution, and behaves as τ⁻¹ (or τ⁻¹½) below 10s and higher than 1000 s.

C. PTP-WR over wireless

As previously discussed, PTP-WR uses PTP to perform a raw estimate of time offset, and use DMTD to determine fine time offset, by measuring the phase on the optical link from master to slave and from slave to master. This is the reason why we need to have a “forth & back” SyncE compliant radio link, and a very stable propagation delay. Therefore, PTP-WR synchronization has not yet been obtained on wireless, because the propagation delay (calculated by using the timestamp exchanged) introduced by the radio is too jittery, as we can see on Table 1 that presents the calculated propagation delay on fiber and wireless link, using the white rabbit timestamps, t1,…,t4 which are defined on Fig.2.

<table>
<thead>
<tr>
<th></th>
<th>Average t2-t1</th>
<th>STD dev t2-t1</th>
<th>Average t4-t3</th>
<th>STD dev t4-t3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2m fiber</td>
<td>0.3 μs</td>
<td>3.4 ps</td>
<td>0.2 μs</td>
<td>4.0 ps</td>
</tr>
<tr>
<td>Wireless</td>
<td>38.4 μs</td>
<td>1.6 μs</td>
<td>25.5 μs</td>
<td>1.6 μs</td>
</tr>
</tbody>
</table>

We have collected approximately 1300 timestamps exchanges between the PTP-WR master and slave, and the calculation shows that the delay variation and standard deviation due to the radio are significantly larger than in the wireline situation.

To overcome the packet delay variation issue, we are planning to design our own radio interface. It will have the following characteristics:

- FDD access technique to avoid the complexity of TDD radio synchronization
- Bidirectional SyncE feature (mandatory for delay estimation).
- IP packet queuing with fixed delay to avoid jitter.

It must be noted that the latency specified in the FDD link we are working with (see [23]) is given as less than 30 μs, which would be equivalent to a time of flight of 10 km over fiber. Also FDD in the microwave technique is the equivalent of the WDM (Wavelength Division Multiplexing) in optic fiber domain.
IV. DISCUSSION

We have shown that accurate and stable time transfer may be obtained through proper wireless link.

We have identified the contribution of SyncE vs basic PTP, and the advantage of PTP-WR vs. PTP-SyncE. The synchronization accuracy is highly dependent to the Master and Slave PTP protocol implementation. Proper radio link can be specified as PTP compliant (supporting TC feature [25]), and SyncE.

The impact of modulation (256QAM, BPSK) and various bandwidth does not reveal a major impact at PTP level. This should be updated under PTP-WR configuration, expected to be more sensitive. So far, we have not made any attempt to control or modify calibration figures in the link delay model.

V. PRELIMINARY CONCLUSION

Our aim is to qualify a “GNSS-free” time transfer solution, allowing microsecond and sub microsecond accuracy.

We have realized time transfer over various protocols (PTP, PTP-SyncE and PTP-WR), and we have tested various microwave link configurations.

So far we have been able to get time transfer accuracy under 50 ns (1σ 10 ns) based on PTP-SyncE over low jitter FDD and TDD radio links, under the condition of a “transparent clock” PTP behavior. This is still far away from PTP-WR or PTP-SyncE performance over optical fiber, as shown on fig 8.

At the accuracy of 50 ns, we have not seen any significant impact of TDD/FDD, spectrum allocation, radio modulation, or bandwidth. SyncE has a major impact on time transfer performance, both on accuracy or on stability.

We have been able to quantify the basic requirements of a PTP-WR microwave link. Next steps will be to define an HW (FPGA) radio interface between PTP-WR devices and a dedicated microwave link.

ACKNOWLEDGMENTS

We warmly thank people and companies who have provided us hardware and support. Time&Frequency Solution–T&FS (UK) providing the PTP-80 master/slave devices, and Meinberg (G) providing the PTP-SyncE master / slave devices. The PTP-WR are from 7Solutions. The main wireless results were obtained over Cambium Networks and Bridgewave radio link.

Thanks to all of them for their support.

We warmly thank our sponsor, Armasuisse, under the contract 8003505738 ARAMIS-No. 041-21 for funding this project in EPFL ESPLAB.

REFERENCES


Precise UTC Dissemination through Future Telecom Synchronization Networks

Wen-Hung Tseng, Sammy Siu, Shinn-Yan Lin, and Chia-Shu Liao
National Standard Time and Frequency Laboratory
Telecommunication Laboratories, Chunghwa Telecom Co., Ltd.
Taoyuan 32661, Taiwan
E-mail: whtseng@cht.com.tw

Abstract—Precise time has become an essential service for most critical infrastructures. The development of telecommunication synchronization networks may facilitate the use of precise time, which will be a great benefit to a large number of devices connected to the Internet. This paper introduces our preliminary work to evaluate the dissemination of precise time reference through a potential telecom synchronization network. With synchronous Ethernet and IEEE1588v2 on-path support (including Telecom Boundary Clocks), time transfer can be accurate at the level of 100 ns. The other challenges of compensating network asymmetry and establishing traceability to Coordinated Universal Time (UTC) are discussed.

Keywords—IEEE 1588; Sync-E; UTC; network Synchronization

I. INTRODUCTION

Precise time synchronization of clocks has become a challenge not only for scientific researches but also for industrial networks. For emerging critical infrastructures, the demands of precise time synchronization have been increasing rapidly, e.g., synchronization for the next-generation mobile system, financial market for recording low-latency trading data [1], the electric smart grid for synchronized phasor measurement units [2], and cloud computing for time-stamping of events. In order to satisfy various applications, it is necessary to provide the standard Coordinated Universal Time (UTC) through protocol level support at the network nodes.

Meanwhile, owing to the need of broadband Internet access for smart phones, the mobile system operators have to manage wireless spectrum more efficiently. For the fourth generation (4G) or the coming future 5G mobile networks, the accurate time synchronization is vital for a larger number of base stations. The investment of synchronization networks is inevitable for today’s telecommunications industries. Towards the best investment-effectiveness, disseminating precise UTC time through commercial telecommunications network has become attractive for both telecommunications industries and official regulations. Deutsche Telekom has proposed a coherent frequency and phase synchronization network suitable for providing UTC traceable services [3]. In U.S., time transfer through a commercial telecommunication network is being investigated for the need of a backup critical timing infrastructure [4]. And, the precision of future critical timing infrastructures would ideally be less than 100 ns at core centers and a few μs at end nodes.

The Precision Time Protocol (PTP, IEEE 1588) is a protocol level solution for clock synchronization in networks [5]. The two-way time-stamps are exchanged between the master and slave clocks using packets containing timing messages. The supporting clocks in the distribution chain include Telecom Boundary clock (T-BC), Telecom End-to-end transparent clock (T-E2ETC), and Telecom Peer-to-peer transparent clock (T-P2PTC) mode [6-8]. Since both T-TCs (T-P2PTC, T-E2ETC) and T-BCs are hardware time-stamps with frequency support (Synchronous Ethernet (Sync-E)), one can accurately measure and correct the delays. If the network asymmetry compensation is further considered and designed, the telecom synchronization network can be a good alternative way to disseminate the UTC time. This paper begins with an introduction to the development and recent status of telecom synchronization networks, and then demonstrates the measurements of time delivery over mobile backhaul networks conducted in Chunghwa Telecom. Finally, the extended discussions are given.

II. EVOLVING TELECOM SYNCHRONIZATION NETWORKS

Traditional broadband telecommunication networks require precise frequency synchronization to ensure the quality of data streams, e.g., the timing of equipment in the Synchronous Digital Hierarchy (SDH) network needs to be synchronized for suppressing jitter and wander and reducing buffer memory size. In ITU-Recommendation G. 811 [9], a stability of $1 \times 10^{-11}$ is recommended for the primary reference clocks (PRCs) for the connection of international digital links. A few refined synchronization networks had been realized over very wide area, as demonstrated by AT&T [10]. Fig. 1 illustrates a simplified frequency distribution chain [11]. Through the help of synchronization supply units (SSU) and SDH equipment clocks (SEC), a reliable synchronization distribution can be ensured in a wide telecommunication network.
provide high data rate and good indoor coverage. In order to
Evolution (LTE), small cells are designed to operate in
coverage in population-dense areas. In 4G Long Term
challenge to serve the increasing demand for capacity and
can be threatened by jamming and spoofing of the signals [15].
A more serious concern is that the reliability of GNSS systems
view of sky. This restricts the use in the indoor environments.
GNSS signals, the reception of GNSS antenna needs a good
synchronization requirements can be easily satisfied by using
less than 5 ns with respect to UTC with dual-frequency
(GNSS) equipment, which can support sub-microsecond
usually equipped with the global navigation satellite system
Coordination (eICIC), and dynamic point blanking (or
Techniques used to coordinate cell sites for inter-cell
interference cancellation and to manage wireless spectrum
more efficiently, such as enhanced inter-cell interference
coordination (eICIC), and dynamic point blanking (or
coordinated scheduling), are required more stringent time
synchronization [12].

In Taiwan, Chunghwa Telecom began to provide the first
4G LTE services on May 30, 2014. In 2015, Chunghwa
telecom has provided 4G coverage to 99 percent of the
population in Taiwan, shown in Fig 2. Current operating
system is the LTE Frequency-division duplex (FDD), which
requires frequency accuracy of ±50 ppb. Both the GNSS and
the mobile backhaul packet based method, e.g., timing packets
of IEEE1588v2, are adopted for implementing frequency
synchronization [17]. The amount of 4G base stations is about
5,000 to 6,000. Chunghwa Telecom plans to deploy small cells
in a small volume in Taiwan from 2015. A regular 4G base
station may work in accordance with about 10 small cells
to achieve the best performance. Cooperating with China Mobile,
Chunghwa Telecom had conducted the first successful
FDD/Time division duplex (TDD) converged field trial at
Global TD-LTE Initiative (GTI) summit in 2014. For the
coming mobile networks, the time accuracy of ±1.5 µs is
essential for a larger number of base stations [12]. Fig. 3
illustrates a modified reference distribution chain to include
synchronous Ethernet clocks (EEC) and packet-based timing
[17].

III. DISSEMINATION OF UTC THROUGH SYNCHRONIZATION NETWORK

A. Evaluating the dissemination of time reference

To evaluate the performance of time delivery over
IEEE1588v2-aware networks with physical layer frequency
support (i.e., Sync-E), we use the hypothetical reference
model-2 (HRM-2) configuration based on ITU-T G.8271.1
[18]. In field trial, the mobile backhaul network is composed
of one telecom grandmaster clock (T-GM), nine telecom
boundary clocks (T-BCs), and one telecom time slave clock (T-
TSC). The network, which covers a range of about 30 km in a
metro area, is in ring structure using packet transport network
(PTN) with Sync-E frequency support, as shown in Fig. 4(a). A
primary reference time clock (PRTC) [19] provides the
reference time to T-GM. The time delivery from T-GM to T-
TSC over single T-BC and 9 cascaded T-BCs are measured,
respectively.

The measurements are performed by using a cesium clock
as the reference clock, as shown in Fig. 4(b). Since the wander
of a cesium clock is extremely low, it provides stable frequency
(i.e., precision better than 10^{-11}) and phase source for phase
wander measurements. A time interval measurement
equipment (HP 53131A) is employed to compare the time
signals between the original 1 PPS of the PRTC and the
terminal output 1 PSS of the T-TSC. All the time stamping
signals between the original 1 PPS of the PRTC and the
end of the measurement are recorded.

The measurement results are shown in Fig. 5. For the link
through single T-BC, the peak-to-peak time difference is 3 ns
over the measured period of 4 h, as in Fig. 5(a). The time
deviation (TDEV) is below 0.1 ns for the observation times
larger than 100 s. The mean time difference is -76.2 ns, mainly
contributed by the node induced asymmetry (e.g., T-BCs). For
the link through a chain of 9 cascaded T-BCs, the peak-to-peak value is 4 ns over the measured period of 19 h, as in Fig. 5(b). The TDEV is below 0.2 ns for the times larger than 100 s. The mean time difference is -112.9 ns, contributed by both node and network link asymmetries. The results show that very stable time transfer can be realized through a mobile backhaul network with the support of IEEE1588v2 packets and Sync-E.

B. Network Asymmetry

Because the round-trip measurement scheme of the IEEE 1588 cannot distinguish the propagation time and clock offsets, the asymmetry of network propagation time will cause the constant bias error of time transfer. We may classify the sources of asymmetry as those induced by the node (e.g., T-BCs) and by transmission links. The tolerable range of constant time asymmetrical error for T-BCs is now strictly required [20]. The mitigation of link delay asymmetry relies on the network paths being carefully designed and compensated [21, 22]. The GNSS timing can also be used to calibrate the asymmetry. One may put GNSS on some critical nodes between core and backhaul networks to help the compensation of asymmetry. For a few 100 ns level accuracy, the asymmetry issues still need to be well considered.

C. Traceability to UTC

Fig. 6 shows a prospective architecture for time synchronization distribution through mobile networks. For being traceable to UTC, the primary reference time clock (PRTC) of a core network needs to be calibrated and monitored in real-time by a national timekeeping institute (e.g., Telecommunication Laboratories (TL) in Taiwan). The GNSS common-view time transfer can provide near real-time monitoring [23]. Now, we are planning to establish the comparison link between UTC(TL) and the PRTC of core networks. And, a dedicated fiber link would be a future solution to verify the PRTC with sub-nanosecond accuracy [24].

IV. CONCLUSION

We have introduced the current development of telecommunication synchronization networks, especially for LTE mobile networks. To prepare for migration from frequency synchronization to time synchronization, we have experimentally evaluated the performance of time transfer
through a mobile backhaul network. With the support of IEEE 1588 v2 packets and Sync-E, very stable result can be realized.

For the future network, more data traffic may arise from sensor networks or the Internet of Things (IoT) [25]. We are looking forward to a wonderful daily life with the help of smart devices, which have their own artificial intelligence and can arrange all things orderly by communicating with each other. When "everything, connected" becomes a reality, we need new concepts of managing systems and optimizing processes. For a system with a large number of devices, time synchronization is usually an essential tool to keep things (e.g., sensors, communications, algorithms, and actuators) in orderly. Accordingly, the telecommunications industry can play the role of carrying IoT services based on its well-developed synchronization networks.

REFERENCES

Gap Reduction Based Frequency Tuning for AlN Capacitive-Piezoelectric Resonators

Robert A. Schneider, Thura Lin Naing, Tristan O. Rocheleau, and Clark T.-C. Nguyen
EECS Department, University of California, Berkeley, CA/USA
Email: schneid@berkeley.edu

Abstract—A voltage controlled resonance frequency tuning mechanism, capable of effecting 1,500 ppm frequency shifts or more, is demonstrated for the first time on an AlN capacitive-piezoelectric resonator. The key enabler here is a compliant top electrode suspension that moves with applied voltage to effectively vary capacitance in series with the device, hence changing its series resonance frequency. Capacitive-piezoelectric AlN micromechanical resonators, i.e., those with electrodes not directly attached to the piezoelectric material, already exhibit high Q-factors compared to attached-electrode counterparts, e.g., 8,800 versus 2,100 at 300 MHz; are on/off switchable; and, as shown in this work, can exhibit electromechanical coupling $C_x/C_0$ of 1.0%. This new ability to tune frequency without the need for external components now invites the use of on-chip corrective schemes to improve accuracy or reduce temperature-induced frequency drift, making an even more compelling case to employ this technology for frequency control applications.

I. INTRODUCTION

RF-MEMS resonators have seen widespread adoption in timing and wireless communications, for which the high $Q$'s and compact form factors of such devices have enabled transformative improvements in low-power integration. Capacitive-gap resonators achieve the ultra-high $Q$'s (exceeding 100,000) and voltage controlled tunability desired for direct RF-channelization and low phase-noise reference oscillators, yet they typically exhibit weaker electromechanical coupling than is desired at frequencies exceeding 100 MHz. Piezoelectric resonators, on the other hand, achieve the strong coupling desired for low-power oscillators and wide-bandwidth filters; however, they do so with reduced $Q$'s of $\approx$2,000 and a lack of built-in tunability. An RF-MEMS resonator technology possessing integrated tuning that is capable of both very high $Q$ and strong coupling is of great interest. In particular, commercially successful aluminum nitride (AlN) resonators would benefit from two key improvements: higher $Q$ and integrated tuning. Voltage controlled frequency tuning of AlN micromechanical resonators allows one to compensate unwanted resonance frequency shifts due to either temperature or manufacturing variations, thus ensuring proper operation. Although numerous frequency tuning methods have been demonstrated, none simultaneously achieves integration of the tuning element and the resonator, negligible DC power consumption, and continuous (non-discrete) tuning.

Frequency tuning through microscale ovenization, even using optimized thermally isolating supports, consumes significant DC power. Such devices also require thermal codeign of a heater and resonator. The work of [1] consumes up to 2.8 mW of DC power to achieve a broad tuning range of $\approx$4,500 parts per million (ppm) on an AlN micromechanical resonator. Although the tuning range is sufficient, the mW-level power consumption of ovenization is undesirable for ultra-low power portable applications.

As another approach, AlN FBAR oscillators can be tuned through adding capacitance in parallel with the FBAR to reduce parallel resonance frequency, $f_p$. The work of [2] uses two parallel switched capacitor banks in a Pierce oscillator topology to tune $f_p$ according to a digital control word to effect very wide frequency shifts, up to 7,250 ppm, while sustaining oscillation, attributable to an AlN FBAR’s large $C_x/C_0$ of $\approx$5%. Although switched capacitor tuning in CMOS is readily suitable for tuning a large FBAR with a low capacitance tuning sensitivity, attofarad-level unit capacitance would be needed for VHF contour mode resonators given their much smaller typical sizes.

Varactor based tuning, in contrast, offers a continuous range of voltage controlled tuning capacitance, albeit with a smaller tuning ratio, $C_{max}/C_{min}$, than is achievable in CMOS using switched-cap tuning. The work of [3] reports varactor based frequency tuning of AlN as an improvement over ovenization citing the advantage of virtually zero power consumption. Here, an off-the-shelf varactor is connected to one or more electrodes of a 13 MHz resonator to effectively stiffen the device and increase its series resonance frequency $f_s$. The varactor-tuned resonators achieve $\approx$600 ppm of frequency tuning when one tuning electrode is used (out of four) and $\approx$1,500 ppm when two tuning electrodes are used. Notably, these tuned devices have rather large motional capacitances $C_x$’s and tuning capacitances $C_{min}$’s of several fF and several pF, respectively, due to their operating frequency. To increase resonance frequencies towards VHF and UHF, such resonators, and hence their varactors too, would need to be scaled down in size dramatically, and would thus need very tight integration of tuning elements and resonators to mitigate parasitics, e.g., either by flip-chip bonding or co-fabrication on the same die. Since contour mode resonators are advantageous primarily due to their layout-defined resonance frequencies, one can envision a multi-frequency system utilizing corrective capacitive tuning on many such resonators, for which numerous variable capacitors, varactors or otherwise, would be needed, consuming considerable space or chip area. Indeed, it would be preferable if the variable capacitor were integrated into the resonator itself.
II. INTEGRATED CAPACITIVE FREQUENCY TUNING OF CAPACITIVE-PIEZO RESONATORS

Capacitive-piezoelectric AlN micromechanical resonators, i.e., those with electrodes not directly attached to the piezoelectric material, offer an important value proposition: higher Q’s by eliminating electrode damping [4] [5]. Such devices already exhibit high Q’s compared to attached-electrode counterparts, e.g., 8,800 versus 2,100 at 300 MHz; are “on/off” switchable [6]; and (as shown in this work) can exhibit respectable electromechanical coupling $C_{g}/C_0$ of 1.0%.

Fig. 1 depicts a tunable 300 MHz radial contour mode capacitive-piezo resonator, including its mode shape and dimensions. During operation, an electric field due to an ac voltage applied across top and bottom electrodes induces strain in the piezoelectric disk, which, for an electrical signal at the resonator’s natural frequency of vibration, excites the device into resonance. Interestingly, introducing top and bottom gaps $g_t$ and $g_b$, and hence series capacitances $C_{gt}$ and $C_{gb}$, above and below the resonator also imparts a gap dependent frequency shift to the resonator, as is discussed in Section III. This work reports the first experimental demonstration of frequency tuning via the voltage controlled reduction of either $g_b$ or $g_t$. The key enabler here is a new, more compliant top electrode suspension, the design of which is discussed in Section IV, that allows for vertical actuation of the top electrode.

Fig. 2 presents the voltage-controlled top electrode pull-down phenomena used to tune series resonance frequency via the cross sections of (a) an untuned and (b) a tuned capacitive piezoelectric AlN radial contour mode disk resonator. The AlN disk, shown in dark gray, is suspended between a pair of top and bottom polysilicon electrodes, shown in light gray, by a central stem. The top electrode, which appears to be floating and bottom polysilicon electrodes, shown in light gray, by a disk, shown in dark gray, is suspended between a pair of top and bottom electrodes that pulls down the top electrode, as is illustrated in Fig. 2b. The reduced gap increases series capacitance $C_{gt}$, thereby decreasing the series resonance frequency, $f_s$.

The integrated tuning method of this work eliminates the need for external tuning components and intrinsically provides the appropriate variable capacitance range necessary to properly tune a device, even as devices are scaled in size and frequency. This new ability to tune frequency without the need for external components now invites the use of on-chip corrective schemes to improve accuracy or reduce temperature-induced frequency drift, making an even more compelling case to employ this technology for frequency control applications.

III. LUMPED ELEMENT MODELING

Fig. 3 presents two equivalent circuit models for a capacitive-piezo resonator. First, in Fig. 3a, sandwiched between the two gap capacitors, $C_{gt}$ and $C_{gb}$, is an electromechanical representation of an AlN resonator, consisting of resonator capacitance $C_{0,ng}$ ($ng$ denotes “no gaps”); an electromechanical transformer with transduction factor $\eta$; and mass, inverse stiffness, and damping elements $l_x$, $c_x$, and $r_x$, respectively. One can show that this circuit is behaviorally equivalent to the circuit of Fig. 3b having Butterworth Van Dyke (BVD) circuit component values, $C_0$, $L_x$, $C_x$, and $R_x$. To account for the changes introduced by capacitive gaps, two factors $\alpha$ and $\beta$ are introduced to modify the component values, given by:

$$\alpha = \frac{V_{AIN}}{V_{in}} = \frac{C_g}{C_g + C_{0,ng}} = \frac{t_{AIN}}{t_{AIN} + \epsilon_{r,AIN}(g_b + g_t)}$$

(1)

$$\beta = \sqrt{\frac{C_g + C_{0,ng} + c_x \eta^2}{C_g + C_{0,ng}}} = \sqrt{1 + \frac{C_{x,ng}}{C_g + C_{0,ng}}}$$

(2)

Here, $\alpha$ in Eq. (1) is the factor by which input voltage in the piezoelectric is reduced by the gaps through capacitive division. In Eq. (1), $C_g$ is the series equivalent of $C_{gt}$ and $C_{gb}$. In our devices, for which $g_b$ and $g_t$ are 120 nm and AlN thickness, $t_{AIN}$, is 1.7 $\mu$m, $\alpha \approx 43\%$. $\beta$ in Eq. (2) is the factor by which $f_s$ is tuned, affecting the motional capacitor, $C_x$, of Fig. 3b. Here, $\beta$ is slightly greater than one and depends on total gap size. Both $\alpha$ and $\beta$ approach one as $g_t$ and $g_b$ drop to
zero. Solving the BVD circuit for the frequency of minimum impedance, \( f_s \) takes the form:

\[
f_s = \frac{1}{2\pi} \sqrt{\frac{1}{L_x C_x}} = f_{s,ng} \beta = f_{s,ng} \sqrt{1 + \frac{C_{x,ng}}{C_{0,ng} + C_g}}
\]

(3)

Note that as \( C_g \) drops from infinity to zero, \( f_s \) rises between the natural eigenfrequency of the resonator \( f_{s,ng} \) and the gap independent parallel resonance frequency \( f_p \). The electromagnetic coupling, \( k_{eff}^2 = C_x/(C_0 + C_x) \approx C_x/C_0 \) of this circuit, a performance metric gauging both maximum fractional resonator tuning range and fractional filter bandwidth, derives from the relationship between \( f_p \) and \( f_s \) as follows:

\[
f_p = f_s \sqrt{1 + \frac{C_x}{C_0}} = \frac{f_s^2 - f_p^2}{f_p^2} = k_{eff}^2
\]

(4)

IV. TOP ELECTRODE SUSPENSION DESIGN

The top electrode suspension is designed to do the following: (a) have a low enough stiffness to enable reasonably low tuning voltages; (b) allow for vertical movement of the top electrode while ensuring uniform displacement over the plane of the electrode; (c) have minimal degrees of freedom to avoid unwanted rotations; (d) have a low enough electrical resistance to avoid loading \( Q \); (e) be compact enough to allow multiple \( \lambda/2 \)-coupled devices to exist side-by-side with each having its own suspension; and (f) have lithographically resolvable critical features. Meeting these requirements, the design of Fig. 4a was chosen and its suspension was FEM simulated to determine its displacement profile in response to a distributed force. One notes that design criterion (b) is met since the coloring of the circular plate, which indicates its vertical displacement, is quite uniform. The suspension consists of four axially-symmetric folded-beam springs and is entirely contained within the desired compact form factor. Using finite element analysis, the vertical stiffness of the device, \( k_z \), is calculated to be 89 N/m. To predict gap tuning as a function of \( V_{tune} \), one solves Eq. (5):

\[
F_z(u_z) = \frac{1}{2} \frac{dC_g}{du_z} V_{tune}^2 = k_z u_z
\]

(5)

Through plotting gap size vs. \( V_{tune} \) as is shown in Fig. 4b, the voltage at which electrode to resonator contact occurs is predicted to be 24V. Notably, over half of the gap tuning occurs within the narrow range of 20-24V.

V. TOP GAP TUNING VS. BOTTOM GAP TUNING

The device shown in Fig. 2 is an instance of a top gap tuned resonator, where the top electrode moves relative to the resonator. One can also make a bottom gap tuned resonator by anchoring the resonator to the top electrode rather than the substrate. Thus, when the top electrode descends, the resonator moves with it, and the bottom gap \( g_b \) changes, as is illustrated in Fig. 5. Here, a top electrode supported resonator allows for maximal electrode coverage and hence better electromechanical coupling and increased tuning range.

VI. RESONATOR TUNING RANGE

The available capacitive frequency tuning range for this device is the frequency difference between its minimum possible \( f_s \), and maximum possible \( f_p \) values. Thus, to determine the largest tuning range one can attain, one can simulate the performance of a capacitive-piezo resonator with gaps reduced below 1 nm, yielding the gapless series resonance frequency, \( f_{s,ng} \). As Eq. (6) shows, \( C_x/2C_0 \) for a resonator is
Tuning range.

A convenient expression for the maximum possible fractional sacrifice layer pinholes, and/or surface roughness. Also, fabrication difficulties related to stiction, post-release cleaning, reducing both if electrode contact is to be avoided. Beyond stress concerns, Such strain gradients set a minimum limit on top gap sizes films can arise, typically causing structures to bend upward.

imperfect deposition conditions, bending moments in AlN thin small gaps also involves undesirable tradeoffs. First, under stress concerns, reducing both \( g_b \) and \( g_t \) very aggressively may also lead to fabrication difficulties related to stiction, post-release cleaning, sacrificial layer pinholes, and/or surface roughness. Also, due motion would likely result in reduced fine gap tuning accuracy.

Gap tuning efficiency drops from 60 ppm/nm at 0 nm to only 3.8 ppm/nm at 500 nm. Third, accommodating a large range of gaps predicts frequency tuning range of 1.1% (11,000 ppm) is available. To illustrate the wide available tuning range, Fig. 6 presents multiple simulated \( S_{21} \) frequency characteristic plots for a capacitive piezo disk resonator having \( g_b = 1 \) nm and a widely actuable top electrode gap, \( g_t \). Here, \( g_t \) is varied from 1 nm to 10 \( \mu \)m. The plot legend lists the nine different values of total gap size included in the simulation and their corresponding frequency shifts from the zero gap state.

**A. Simulation of Theoretical Maximum Tuning Range**

Our model of the 300MHz 11.2-\( \mu \)m-radius capacitive-piezo disk of this work having full electrode coverage and 1 nm gaps predicts \( C_x/C_0 = 2.2\% \). Thus, a maximum fractional frequency tuning range of 1.1% (11,000 ppm) is available. To illustrate the wide available tuning range, Fig. 6 presents multiple simulated \( S_{21} \) frequency characteristic plots for a capacitive piezo disk resonator having \( g_b = 1 \) nm and a widely actuable top electrode gap, \( g_t \). Here, \( g_t \) is varied from 1 nm to 10 \( \mu \)m. The plot legend lists the nine different values of total gap size included in the simulation and their corresponding frequency shifts from the zero gap state.

B. Constraints on Making Large Gaps for Tuning

Tuning with large gaps involves several undesirable tradeoffs. For example, excessive values of \( g_{total} = g_t + g_b \), e.g., beyond 500 nm, can greatly diminish electromechanical coupling. Also, gap tuning efficiency, \( \Delta f_{gap} \), drops as \( g_{total} \) increases, eventually becoming quite small at large gap sizes. Gap tuning efficiency drops from 60 ppm/nm at 0 nm to only 3.8 ppm/nm at 500 nm. Third, accommodating a large range of motion would likely result in reduced fine gap tuning accuracy at small gap sizes. In view of these considerations, one should keep gap sizes below a judicious value of \( g_{total,max} \).

C. Constraints on Making Small Gaps for Tuning

Choosing to fabricate capacitive-piezo resonators with very small gaps also involves undesirable tradeoffs. First, under imperfect deposition conditions, bending moments in AlN thin films can arise, typically causing structures to bend upward. Such strain gradients set a minimum limit on top gap sizes if electrode contact is to be avoided. Beyond stress concerns, reducing both \( g_b \) and \( g_t \) very aggressively may also lead to fabrication difficulties related to stiction, post-release cleaning, sacrificial layer pinholes, and/or surface roughness. Also, due to the need to overetch AlN with the current fabrication process, a thick bottom sacrificial oxide eases concerns of excessively overetching the bottom electrode interconnect layer. One should thus avoid scaling \( g_{total} \) too aggressively simply to ensure proper operation.

D. Optimal Sizing of Gaps

In this work, top and bottom gaps of 120 nm were chosen to attain a respectable tuning range of \( \approx 1,600 \) ppm without risking device failure. Here, the chosen tuning ratio, \( C_{max}/C_{min} \), is 2. Larger tuning ratios are certainly possible. To examine the relationship between \( g_{total} \) and \( \Delta f_s = f_s - f_{s,ng} \), refer to Fig. 7. Overlaid on the plot are annotations showing the designed gap range for our device and its associated tuning range of 1600 ppm. Considering possible future performance improvements, through using a wider range of \( g_{total} \) values, e.g., with a fixed \( g_b \) of 70 nm and a variable \( g_t \) between 0 nm and 430 nm, 5,200 ppm of tuning would be attained, representing a potential threefold improvement. Additional design iterations will determine the true limits of this technology.

**VII. DEDICATED TUNING TRANSDUCERS**

Two gap reduction based capacitive tuning methods are explored in this work: tuning through an input/output (I/O) transducer, which affects the I/O coupling, \( \alpha \), of Eq. (1); and tuning through a separate tuning transducer, which does not. Our analysis so far has emphasized tuning for single transducer devices for which I/O transduction and tuning transduction occur through the same electrode pair. Single transducer capacitive-piezo resonators advantageously minimize \( R_x \) and maximize \( C_x/C_0 \); however, the achieved values do not stay constant over the tuning range. If \( R_x \) and \( C_x/C_0 \) change too much through tuning \( f_s \), some oscillator or filter designs may not function properly. Additionally, single transducer tuning
through proper chamber conditioning and heating the wafer the reactively sputtered AlN film was dramatically improved, µ the thick polysilicon was mitigated using 2.5 and 50% to dedicated tuning, hence reducing µ modied to instead allocate 50% of its area to dedicated I/O and from I/O gap reduction based tuning to separated tuning, if µ calculated the frequency tuning for the overall resonator, simply substituted µ, tuned of Eq. (7) for µ in both Eq. (3) and β of µ: µ, tuned = cµ(cµ + C0,n0) cµη 2 + (cµ + C0,n0) (7) As an example of the performance tradeoff of switching from I/O gap reduction based tuning to separated tuning, if the electrode configuration of a single transducer device is modified to instead allocate 50% of its area to dedicated I/O and 50% to dedicated tuning, hence reducing µ and µ, holding all else constant, Rµ will increase by a factor of four and Cµ/C0 (and hence tuning range) will drop by a factor of two.

VIII. DEVICE FABRICATION

To validate the design and operation of the tunable capacitive-gap piezo resonators presented here, a number of such resonators were fabricated using a fabrication process similar to that of [6], though with several key improvements. Fig. 9 presents a cross section of a capacitive-piezo disk resonator. Here, the deposited film forming the bottom electrode was changed from the previously-used 150 nm thick molybdenum film to a 2µm-thick doped polysilicon film to avoid galvanic corrosion during a wet hydrofluoric acid release. The doped polysilicon is made relatively thick to reduce its sheet resistance to under 10 Ω/□. The large topography of the thick polysilicon was mitigated using a 2.5 µm oxide deposition which was then chemical mechanically planarized. The sacrificial oxide thicknesses were reduced down to 125 nm to improve Cµ/C0. Most importantly, the crystallinity of the reactively sputtered AlN film was dramatically improved, through proper chamber conditioning and heating the wafer during deposition, further increasing Cµ/C0. Using an X-ray diffractometer, the measured full-width half maximum of the AlN film used was measured at 1.60°. Finally, the top electrode thickness was reduced, allowing for lower actuation voltages of the top electrode due to the strong cubic dependence of vertical stiffness on material thickness.

IX. EXPERIMENTAL RESULTS

We report experimental tuning demonstrations at 300 MHz of (a) a two transducer top-gap-tuned resonator to verify constant Rµ tuning and (b) a single transducer bottom-gap-tuned resonator to verify maximum tunability and minimal Rµ. Both devices were tested in a vacuum probe station using a properly calibrated network analyzer with a 50Ω termination at both ports. Transmission scattering parameter (S21) measurements were collected and plotted for various values of Vtune to verify operation.

Fig. 10 demonstrates tuning in a top-supported single-disk device utilizing one electrode pair for I/O transduction and a separate electrode pair for top gap reduction tuning. As shown in the SEM, two radial contour-mode resonators are mechanically coupled using an extensional mode 800-nm-wide coupling beam of acoustic length λ/2 to form a single degree of freedom resonant system. The upper resonator is surrounded by an I/O transducer electrode pair through which its transmission vs. frequency is measured. Surrounding the lower resonator is the same tuning transducer shown in Fig. 8, consisting of a grounded bottom electrode and bias voltage (Vtune) actuated spring-supported top electrode. The measurement plot includes a series of 2-port transmission (S21) measurements taken with various values of Vtune ranging from 0 to 24V. The entire frequency characteristic is shown to uniformly shift downward in frequency as Vtune is increased, with no changes in Rµ, Q, and Cµ/C0, as expected.

Fig. 11 demonstrates tuning in a top-supported single-disk device with one transducer for both tuning and I/O in order to achieve maximum total tuning range. The measurement scheme for this device uses a bias tee to add a bias voltage to the RF input signal. As shown in the measurement plot, one achieves improved Rµ, Cµ/C0, and tuning range through using a single transducer, as expected. Here, the measured tuning range is increased to 1547 ppm. The Rµ of this device is showed to decrease as Vtune is increased, agreeing with the model. Note that fµ stays relatively fixed, as is simulated in Fig. 6.
A summary of capacitive-piezo tunable resonator model parameters that correctly predict tuning behavior of our demonstrated devices is included in Table I.

TABLE I: Tunable Resonator Model Parameters

<table>
<thead>
<tr>
<th>Model Parameter</th>
<th>Single Disk Resonator</th>
<th>Two Disk Resonator</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_x )</td>
<td>( 9.3 \times 10^{-7} ) kg/s</td>
<td>( 2.75 \times 10^{-6} ) kg/s</td>
</tr>
<tr>
<td>( l_x )</td>
<td>( 1.535 \times 10^{-12} ) kg</td>
<td>( 3.070 \times 10^{-12} ) kg</td>
</tr>
<tr>
<td>( c_x )</td>
<td>( 1.840 \times 10^{-7} ) s^2/kg</td>
<td>( 9.20 \times 10^{-8} ) s^2/kg</td>
</tr>
<tr>
<td>( \eta_{\text{fr}} )</td>
<td>( 4.90 \times 10^{-5} ) C/m</td>
<td>( 4.50 \times 10^{-5} ) C/m</td>
</tr>
<tr>
<td>( C_{0,\text{ng}} )</td>
<td>( 1.99 \times 10^{-14} ) F</td>
<td>( 1.88 \times 10^{-14} ) F</td>
</tr>
<tr>
<td>( g_{\text{tot},\text{tune}} ) [min, max]</td>
<td>[136 nm, 240 nm]</td>
<td>[240 nm, 240 nm]</td>
</tr>
<tr>
<td>( \alpha ) [max, min]</td>
<td>[0.56, 0.42]</td>
<td>[0.42, 0.42]</td>
</tr>
<tr>
<td>( \beta ) [min, max]</td>
<td>[1.00488, 1.00639]</td>
<td>[1.0029, 1.0029]</td>
</tr>
<tr>
<td>( c_{x,\text{tuned}} / c_d ) [min, max]</td>
<td>N/A</td>
<td>[0.9943, 0.9956]</td>
</tr>
<tr>
<td>( \Delta f_x )</td>
<td>-1510 ppm</td>
<td>-650 ppm</td>
</tr>
</tbody>
</table>

X. CONCLUSIONS

Voltage controlled gap reduction has now been shown as an effective means for controlling the resonance frequencies of capacitive-piezo AlN micromechanical resonators. Introducing integrated variable capacitances into VHF AlN resonators is also shown to enable higher Q’s while maintaining strong coupling. We have demonstrated 1547 ppm of tuning at 300 MHz using a single transducer and 630 ppm of tuning using a dedicated tuning electrode to maintain constant \( R_x \). Though these initial results are impressive, analysis shows that through optimizing gap sizes, a further threefold increase in the tuning range will be feasible. We believe that the tuning method presenting in this work will be useful for improving accuracy or reducing temperature induced frequency drift in our resonators, for which measured TCF’s of 15 ppm/K are typical. Frequency modulation, mixing, passband correction of filters, and filter tunability may also be implemented using this technique.

ACKNOWLEDGMENT

The authors would like to thank DARPA, the Berkeley Sensor & Actuator Center, and the staff of the Berkeley Nanofabrication Laboratory for supporting this work.

REFERENCES

Switchable 2-Port Aluminum Nitride MEMS Resonator Using Monolithically Integrated 3.6 THz Cut-OFF Frequency Phase-Change Switches

Gwendolyn Hummel and Matteo Rinaldi
Electrical and Computer Engineering
Northeastern University
Boston, Massachusetts, USA
Email: hummel.gw@husky.neu.edu, rinaldi@ece.neu.edu

Abstract—This work presents the first experimental demonstration of an intrinsically switchable Aluminum Nitride (AlN) 2-port MEMS resonator using 3 monolithically integrated chalcogenide phase change material (PCM) switches.

Keywords—phase change materials; reconfigurable resonators; aluminum nitride resonators; piezoelectric resonators.

I. INTRODUCTION

One of the main issues with current radio frequency (RF) wireless systems is the fact that the modern commercial and military wireless spectrum is rapidly changing and extremely crowded, requiring devices to use multiple communication standards, along with upgrading and changing system designs frequently. Because of this, the demand for RF components with specific attributes has been growing steadily. These special attributes include: miniaturized scale (to reduce the space required for the system), low cost, low power consumption (to allow components to consume low amounts of battery charge in technologies such as mobile phones), and high reconfigurability. In particular, highly reconfigurable RF systems are desirable because they can dynamically adapt to the rapidly changing RF spectrum satisfying the requirements of multiple current and future wireless communication standards.

In order to cater to this increased demand for highly reconfigurable RF systems, the monolithic integration of piezoelectric MEMS resonators and phase change material (PCM) RF switches has been proposed as a solution towards the development of miniaturized, low cost, high performance, and highly reconfigurable RF components. The monolithic integration of high performance RF switches in the resonator design enables effective ON/OFF switching of the resonant response as well as reconfiguration of the device electrical impedance and operating frequency [1], [2].

Chalcogenide PCMs are materials that change state between a high resistivity amorphous phase and a low resistivity crystalline phase. Such phase change is triggered by heating the PCM to a crystallization temperature (~200 °C) to obtain the crystalline (ON) phase or heating the material to the melting point in order to obtain the amorphous (OFF) phase. PCMs such as Ge2Sb2Te5 (GST) have previously been used in non-volatile memory applications. Recently PCMs have also been investigated for use in RF systems showing great potential thanks to their ease of fabrication, high ON/OFF resistance ratio (~10^5), and low ON resistance (~1-2 Ω) [3], [4]. Furthermore, PCMs do not require constant power to maintain either the ON (low resistance) or the OFF (high resistance) state. PCM (specifically Ge50Te50 [5]) RF switches have been demonstrated showing figures of merit (FOM) superior to the ones of other widely used solid-state switch technologies [6] and have been integrated in the designs of reconfigurable inductors [7], and programmable 1-port AlN MEMS resonators [1], [2].

In this work, for the first time, 3 PCMs switches, with ~3.6 THz cut-off frequency (at least 3X larger than aforementioned solid-state switches [6]), are monolithically integrated in the design of a 2-port Aluminum Nitride (AlN) MEMS resonator. The resulting 2-port reconfigurable MEMS resonator demonstrates low electrical loss in the ON state (~2 Ω) and high input impedance in the OFF state (input port capacitance as low as 70 fF), setting a milestone towards the development of AlN/PCM single-chip multi-band RF systems with the highest level of reconfigurability and minimum possible effect on the RF performance.

II. DESIGN AND FABRICATION

A. Design

Differently from previous demonstrations, a 2-port configuration is chosen since it enables the synthesis of reconfigurable narrow-band filters by simply electrically cascading multiple switchable resonator stages. The lateral-extensional mode resonator is composed of an AlN thin-film (500 nm) sandwiched between a bottom (Pt) plate electrode connected to electrical ground and top interdigital electrode (Al) patterned in 3 parallel fingers: 2 of which are connected to form the input port and 1 is connected to form the output port. 3 programmable PCM vias are employed to connect each of the metal fingers to the corresponding device terminal through a 250 nm SiO2 layer (Figure 1).
Fig. 1. Scanning Electron Microscope (SEM) image of (a) reconfigurable 2-port resonator, (b) switchable electrodes, (c) PCM via, and (d) 3D schematic of resonator design.

B. Fabrication

Fabrication of this switchable 2-port piezoelectric MEMS resonator was completed with a post-CMOS, 6-mask fabrication process (Figure 2). Starting with a high resistivity silicon wafer ($\rho > 20,000 \ \Omega$), 100 nm of platinum (Pt) was deposited and patterned to form the bottom electrode. 500 nm aluminum nitride (AlN) was deposited and etched to form the vias to the Pt and the dimensions of the resonant micro-plate. The interdigital electrodes (100 nm aluminum (Al)) were deposited and patterned. An electrical insulator layer of 250 nm silicon dioxide (SiO$_2$) was PECVD deposited and etched, followed by the deposition of the PCM layer (Ge$_{50}$Te$_{50}$) to form the via switches. The top electrode of the switches and the probing pads (100 nm copper (Cu)) was deposited and patterned, and finally the silicon was isotropically etched in XeF$_2$ to release the device.

This simple and CMOS-compatible fabrication process lends itself very well to the monolithic integration required to design highly reconfigurable resonator structures.

III. EXPERIMENTAL RESULTS

The transition temperature needed for the switching of the PCM vias was achieved by passing current directly through the PCM (direct heating [3]). Voltage pulses of 1.5 V amplitude and 200 µs duration were used to turn the PCM switches ON while pulses of 4.7 V amplitude and 2 µs duration were used to trigger the ON-to-OFF transition.

Effective ON/OFF switching of the device transmission (~42 dB variation for fixed 50 Ω termination) was demonstrated (Figure 3), due to the monolithic integration of 3 ultraminiaturized (2 µm × 2 µm) PCM switches with radio frequency (RF) performance superior to the one of more conventional RF switch technologies: an ON-state resistance of 2 Ω (a low $R_{ON}$ minimizes the insertion loss of the resonator design) with an OFF-state capacitance and resistance of 22 fF and ~20 MΩ, respectively, were measured for the PCM switches of this work resulting in an RF switch cutoff frequency of 3.6 THz and an improved figure of merit ($FOM=R_{ON}C_{OFF}^{-1} \approx 44$ fs) compared to the ~100s fs of typical solid-state RF switches [6].

The device transmission (at the resonance frequency) was attenuated by ~42 dB when all three PCM switches were turned OFF. In the ON state (when all three PCM switches were turned ON), an insertion loss of ~10 dB, for a fixed 50 Ω termination, was recorded; which corresponds to a ~2.5 dB insertion loss for a matched termination of 571 Ω (this could be decreased with optimization of the resonator design, as a relatively low Q value of ~420 was achieved for this first prototype). A large input impedance variation between the ON and the OFF state ($C_{ON}/C_{OFF} \approx 21X$) was also recorded (Figure 4). Such low value of the device OFF state capacitance is

---

*This work was funded by DARPA MTO (N66001-14-1-4011) under the RF-FPGA program (Dr. Troy Olsson).*
highly desirable as it enables the implementation of large programmable filter arrays with reduced capacitive loading, enabling the achievement of the highest level of reconfigurability with the minimum possible effect on the RF performance of each individual element.

IV. CONCLUSION

This work presents effective ON/OFF switching of a 2-port AlN contour-mode resonator using monolithically integrated phase change material switches (2 μm x 2 μm) with a 3.6 THz cut-off frequency (switch ON-state resistance ~2 Ω and OFF-state capacitance ~22 fF). Effective ON/OFF switching of the device transmission (~42 dB variation with ~21X higher input impedance in the OFF state), are demonstrated without increasing the complexity of the device fabrication process (only 2 additional masks compared to a static resonator) or requiring substantial modification of the device layout (only an additional probing pad per via). The proposed technology platform enables dense integration of resonators and switches with reduced resistive losses and capacitive loading effects setting a milestone towards the development of reconfigurable RF micro-systems capable of adapting to the rapidly changing RF environment and satisfying multiple wireless communications standards.

ACKNOWLEDGMENT

The authors wish to thank the staff of the George J. Costas Nanoscale Technology and Manufacturing Research Center at Northeastern University for assistance with device fabrication.

REFERENCES

Analysis of the Impact of Release Area on the Quality Factor of Contour-Mode Resonators by Laser Doppler Vibrometry

Brian Gibson, Kamala Qalandar
and Kimberly Turner
Department of Mechanical Engineering
University of California, Santa Barbara
Santa Barbara, USA
Email: briangibson@engr.ucsb.edu
kamala@engr.ucsb.edu
turner@engr.ucsb.edu

Cristian Cassella
and Gianluca Piazza
Department of Electrical and Computer Engineering
Carnegie Melon University
Pittsburgh, USA
E-mail: ccassell@andrew.cmu.edu,
piazza@ece.cmu.edu

Abstract—Energy dissipation through the anchors of an aluminum nitride (AlN) MEMS contour-mode resonator (CMR) plays an important role in setting the device quality factor at frequencies under 500MHz [1]. The acoustic energy leaving the resonator is affected, not only by the anchor, but also the portion of the surrounding device layer that has been released from the substrate during fabrication. Typical device simulations used for design do not take into account the motion of this additional area. Using laser Doppler vibrometry and COMSOL simulations we show a variation in device $Q$ by 28% as a result of the motion of this released region.

I. INTRODUCTION

MEMS based resonators are an active field of research due to great interest in their commercialization into frequency sources and filtering operations. Their CMOS compatibility makes them an excellent replacement for traditional quartz-based resonators by reducing the cost and size of these components. An important figure of merit for resonator performance is the quality factor, $Q$. This is an inverse measure of the energy dissipation per cycle of the resonator. In an effort to maximize the device $Q$ of MEMS contour-mode resonators (CMRs), extensive research has been conducted in setting the electrode size and placement [1], inactive region dimensions [2], and anchor dimensions [3], [4], [5]. In fact, [3], [4] have shown that, below 500MHz, anchor loss is a dominant source of damping in piezoelectric CMRs.

Yet to be considered is the effect of the substrate at the anchor attachment points on device $Q$. When modeling these devices, the assumption is always made that the outer end of the anchor is fixed with respect to the substrate. This simplifies analysis but is not completely accurate due to the fact that an isotropic etch step is used to release the device itself from the substrate [6]. This etch process also undercuts the region of the device layer outside of the anchor region (Figure 1). This region will undergo greater displacement as a result of the resonator motion. This energy leaves the body of the resonator through the anchors and is largely unrecoverable.

To determine the effects of this substrate motion, the out-of-plane displacement of a 220MHz aluminum nitride (AlN), piezoelectric contour-mode resonator (CMR) was measured using laser Doppler vibrometry (LDV) [7] along a path shown in Figure 1. Electrical measurements were also taken to find $Q$ and $k^2$. The device then underwent a XeF$_2$ etch to release additional substrate area. The measurements were then taken again. This sequence was repeated eight times. We show that there is a variation in $Q$ of 28% due to the effects of the released region and that this variation occurs as a function of $L$, the distance from the active region of the device to the edge of the released region of the substrate (Figure 1).

In a CMR, the primary motion is in-plane which cannot be directly measured with LDV. Despite this, there is out-of-plane motion due to the Poisson effect and due to the piezoelectric effect. To this end, a finite element model was built which included perfectly matched layers (PMLs) [8], [9] to simulate this device. This model was validated by matching its out-of-plane motion with that of the LDV measurements. The energy flux through the resonator anchors produced by FEA could then be compared to the released substrate length, $L$.

II. ALUMINUM NITRIDE CMR RESONATORS

A. Design and Fabrication

The resonator used for this study is an AlN based CMR [2]. It is a piezoelectric device consisting of an AlN device layer sandwiched on the top and bottom by two metal electrode layers. The top electrode layer consists of interdigitated electrodes connected to opposing voltage polarities and the bottom metal layer is a solid floating electrode used to contain the electromagnetic field. By applying an AC voltage to the electrodes at the device’s resonant frequency, in-plane Lamb waves are set-up as a result of the strain induced by the $d_{31}$ coefficient in the AlN. The resonant frequency is set by [10]...
Fig. 1: A 220MHz AlN contour-mode resonator. The device layer immediately around the CMR has been released from the substrate and exhibits strain as the result of energy dissipated out of the anchors. The released length, $L$, was found to have a direct effect on device $Q$.

$$f_0 = \frac{1}{2W} \sqrt{\frac{E_{eq}}{\rho_{eq}}}$$

(1)

where $W$ is the electrode spacing, and $E_{eq}$ and $\rho_{eq}$ are the equivalent Young's modulus and density, respectively, of the combined three layers. They are fabricated by first depositing a 10nm layer of titanium onto a silicon substrate followed by 100nm of aluminum. A 1µm layer of AlN is deposited for the device layer and then 100nm of aluminum is patterned to form the top electrode layer. A Cl$_2$-based dry etch is used to form the device shape by etching the AlN. The final step uses an isotropic XeF$_2$ etch to release the device from the substrate. A Cl$_2$-based dry etch is used to form the device shape by etching the AlN. The final step uses an isotropic XeF$_2$ etch to release the device from the substrate. Subsequent XeF$_2$ etch steps for this study were carried out using an Xactix series X3 etcher using 3 Torr of XeF$_2$ and 2.5 Torr of N$_2$ in 60 second cycles.

B. Experimental Measurements

Admittance data was collected using a Rhode&Schwarz ZVL vector network analyzer (VNA). The device was probed using an Infinity probe from Cascade Microtech and the VNA was calibrated using a Cascade Microsystems impedance standard. From this data, resonant frequency, $f_0$, $Q$, and $k_i^2$ were extracted.

The out-of-plane displacement measurements were taken using a Polytec UHF120 LDV through a 50X objective. The spot size is approximately 2µm. The device was driven using a Rhode&Schwarz SMBV100A function generator. The device was probed in the same setup as the VNA measurements. The device was scanned along the path shown in Figure 1 with a spatial resolution of 366nm. The device was driven at -6dBm and its resonant frequency determined by the electrical measurements. This power level was chosen to prevent the resonator from operating in a nonlinear regime while still inducing enough out-of-plane motion to be detectable with the LDV. Displacement data was collected with an out-of-plane resolution of 4.2pm and with a noise floor of 4pm. The resolution bandwidth was 19.5kHz.

C. Finite Element Modeling

In order to determine the energy loss through the anchors, a COMSOL Multiphysics model of the resonator body, anchors, and surrounding substrate was made. The geometry of the model was taken from the fabricated device, and PMLs were used to model the semi-infinite substrate [8], [9]. The PML layers were used at the edges of the silicon and AlN layers and on the upper electrode layer near the region of interest (shown in black in Figure 2). The PMLs were placed 300µm from the outer edge of the anchor in the direction of the scan path. This distance is far enough from the edge of the fixed substrate region so that any reflections at this interface could be taken into account in the simulations. The simulated magnitude and wavelength of the out-of-plane displacement were compared to that of the LDV measurements along the same path (Figure 1) to validate the model. The model was then simulated for each release distance, $L$, from the experiment and the energy flux through the outer cross-section of the anchor (Figure 2) for each release distance was calculated.

III. Results and Discussion

Figure 3 shows the LDV measurements and COMSOL simulation of the displacement of the substrate region shown in Figure 1 before any additional etching. This quality of fit of the simulation to the LDV data validated the COMSOL model. To provide a value that is proportional to strain energy, the magnitude of the displacement profiles from Figure 3 were squared and then the area under the resulting curve was integrated. This was done for each etch step and the data is plotted in Figure 4.

As discussed in [11] the amount of energy released by the active region towards each anchor is a periodic function of
the width of the bus bar, \( d \), and the anchor length, \( L_a \), and is governed by the equation:

\[
E_{\text{anchor}} = \text{abs} \left( \frac{\sin(2kD)}{4Z_0} \right)
\]  

where \( D = (L_a + d) \), \( k \) is the wave number, and \( Z_0 \) is an equivalent characteristic impedance of the active region of the resonator. Equation 2 shows that the minimum energy is transmitted from the active to the inactive regions when the dimension, \( D \), is a multiple of \( \lambda/4 \). In the current study \( D \) was considered to be equal to the entire released length \( L \) from Figure (1) and this equation was fit to the displacement data in Figure (4). When \( L \) is a multiple of \( \lambda/4 \) the inactive region acts as a \( 1/4\lambda \) transformer to place either a virtual fixed or stress-free boundary at the edge of the active region. Either of these conditions confines the maximum amount of acoustic energy into the active region and minimizes the energy lost to the anchor/substrate region. As shown in figure (4), when this occurs the displacement in the anchor region increases but the corresponding ratio between the displacements in the active and inactive regions increases as well. This overall decrease in energy flux through the anchors can be seen in the FEA results in Figure (5). Minima occur at \( n\lambda/4 \), corresponding to the clamped condition at the edge of the anchor region.

This minimization of energy loss corresponds to maxima in \( Q \). Table I gives the values for \( Q \), \( k^2_t \), and the resonant frequency that were extracted from electrical measurements collected at each etch step. Figure 6 shows the quality factor as a function of \( L \). Device \( Q \) varied by 28% as the released distance increased. A similar equation from [11] was modified to account for the released area:

\[
\tilde{Q} = \frac{Q_{\text{eff}}}{1 + 2\alpha \left( \frac{2\alpha h \sin(2kL)}{4k} \right)}
\]  

where \( Q_{\text{eff}} \) is the ideal \( Q \) with no anchor loss, and \( \alpha \) is proportional to the amount of energy lost to the anchor from the active region of the resonator. An identical \( \lambda/4 \) variation can be seen for \( Q \) with peaks at \( n\lambda/4 \). The \( k^2_t \) values did not show a dependance on \( L \) due to the fact that the static capacitance, \( C_0 \), and motional capacitance, \( C_m \), have a similar variation with \( L \). Since the measured device \( k^2_t \) was independent of \( L \), we can confirm that the out-of-plane motion in the inactive region must vary proportionally to \( Q \) due to conservation of energy. This fact explains the same periodicity of displacement and \( Q \) with respect to \( L \).

IV. CONCLUSION

It has been shown that anchor loss represents the major portion of the energy dissipation in contour-mode resonators at frequencies below 500MHz. We have shown that a substantial portion of that energy loss can be contributed to the portion of the device layer outside of the anchors that has been released from the substrate. This dimension is typically not considered during the design and fabrication process but can represent an improvement of up to 28% in quality factor. Future designs
Fig. 5: The simulated energy flux through the anchors of the resonator over a complete cycle. A minima occurs every $n\lambda/4$. This occurs as a result of the inactive region acting as a $1/4\lambda$ transformer and placing a virtual fixed-boundary condition at the edge of the active region. This serves to contain the acoustic energy within the active region of the resonator, improving $Q$.

Fig. 6: The quality factor at each etch step. $Q$ varies by 28% as the etch distance is increased. The maxima at $n\lambda/4$ correspond to points of minimum energy loss through the anchors.

TABLE I: The released distance, $L$, $Q$, $k_2^2$, and resonance frequency for each release step.

<table>
<thead>
<tr>
<th>$L$ [µm]</th>
<th>$Q$</th>
<th>$k_2^2$</th>
<th>$f_0$ [MHz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>46.15</td>
<td>1866</td>
<td>.3949</td>
<td>218.86</td>
</tr>
<tr>
<td>49.81</td>
<td>1631</td>
<td>.3943</td>
<td>218.85</td>
</tr>
<tr>
<td>54.21</td>
<td>1517</td>
<td>.4148</td>
<td>218.76</td>
</tr>
<tr>
<td>57.88</td>
<td>1897</td>
<td>.4056</td>
<td>218.43</td>
</tr>
<tr>
<td>61.17</td>
<td>1621</td>
<td>.3858</td>
<td>218.58</td>
</tr>
<tr>
<td>63.0</td>
<td>1403</td>
<td>.3760</td>
<td>218.35</td>
</tr>
<tr>
<td>64.5</td>
<td>1419</td>
<td>.4156</td>
<td>218.36</td>
</tr>
<tr>
<td>66.7</td>
<td>1676</td>
<td>.4255</td>
<td>218.37</td>
</tr>
<tr>
<td>69.2</td>
<td>1951</td>
<td>.4140</td>
<td>218.57</td>
</tr>
</tbody>
</table>

will need to include this if MEMS based resonators are to become the replacement for traditional quartz designs.

REFERENCES


Accurate Removal of RAM from FM Laser Beams *

Locking Accuracy at the ~1E-6 Level

Hall, John L; Zhang, Wei; Ye, Jun
JILA, University of Colorado, and NIST
Boulder, CO USA
jhall@jila.colorado.edu

Abstract—We demonstrate here a RAM cancellation method of reaching locking accuracy at shot-noise sensitivity level, and with a reserve precision sufficient to still be free of RAM-induced problems when the bandwidth has been narrowed to some tens of milliHz. Non-optical rf pickup sets the current limit at 2.8 ppm. Basically this paper announces the RAM-Buster approach needed to achieve the ideal spectroscopic accuracy, shot-noise-limited, as had been anticipated in the 1983 paper.

I. INTRODUCTION

The use of balanced rf-sidebands to serve as a Local Oscillator for observing phase variations of the carrier component has a long history, and the microwave frequency discriminator developed by R. V. Pound was an important application in the context of WW II. Eventually good ideas are rediscovered by later workers in other fields, and certainly it was exciting in the early 1980’s for the laser community to have the rf sideband methods applied to laser tuning[1] and spectroscopy[2] and frequency locking[3]. In a single step one went from the baseband frequency region, where amplitude noise may be measured in percents, to using the capability of optical phase modulation to encode the desired resonance information to a higher frequency, where the laser noise is vastly lower, usually near to the shotnoise level. This signal/noise (S/N) increase has enabled much of the remarkable progress in laser spectroscopy during the ensuing 3 decades. But soon spectroscopists found their spectra, while nearly free of laser noise, were importantly degraded by the unstable baselines observed. This transfer of low frequency noise up to rf center frequencies was the result of Residual Amplitude Modulation (RAM), which is produced by the Phase Modulator itself, as well as by unwanted spurious reflections in the optical setup[4]. So the first factor of 100 noise reduction came immediately with the rf sideband method, but this paper presents the first deterministic approach to capture the full remaining potential noise suppression, an attractive additional factor of >1000-fold useful dynamic range, for enhancing the precision and accuracy of laser spectroscopy into the shot-noise-limited S/N domain >1E6. We demonstrate here a RAM cancellation method of reaching shot-noise sensitivity level, but with a reserve accuracy sufficient to still be free of RAM-induced errors when the bandwidth has been narrowed to some tens of milliHz, by using extended averaging times, co-addition of multiple scans, or FFT-based direct spectral resolution. Basically this paper announces the RAM-Buster approach to achieving the shotnoise-limited spectroscopy as had been anticipated in the 1983 paper[3].

II. HISTORICAL ANTI-RAM METHODS

Some innovative early anti-RAM efforts involved modulation strategies which changed the spectral composition of the light, so that an included modulation sideband either was or was not absorbed by the sample, but at constant total laser power. Of course the baseline noise was relatively huge, dependent on the technology level of the laser. A two-tone modulation helps by sensing an upset of the symmetry of the modulation, allowing a low detection frequency, but with less noise, but also less signal[5]. A direct engineering approach measured the light beam after it was rf-modulated, but before it entered the sample of interest[6]. Ideally one could improve the modulation process[7] or servo-control it so that no RAM changes were visible on the sampled beam. These results were recently improved to nearly within the shotnoise level of the sampled light[8]. An important earlier advance minimized the AM noise at some chosen rf frequency, but the method apparently was not studied relative to RAM suppression[9]. The method described here has the larger objective of accurately cancelling the RAM, using a strategy which is fundamentally advanced from the earlier “measure-and-quickly-feedback” strategies, in that we now take advantage of the relatively slow changing rate of the RAM.

III. ORIGINS OF RESIDUAL AMPLITUDE MODULATION

RAM is produced by misalignment of optical polarization in the modulator, or by inhomogeneous rf fields in the crystal, which can lead to synchronous beam steering or synchronous displacement of the beam exiting the Electro-Optic Modulator (EOM). Usually the beam needs to be “misaligned” relative to the crystal faces to avoid forming weak Fabry-Perot etalon fringes. An attractive case uses a crystal in which the modulation occurs for the Extra-Ordinary (e)-polarization, as the double-refraction “walkoff” for this utilized polarization will further suppress etalon fringes. A problem arises by the action of Snell’s law in dynamically deviating the beam entering the crystal face at a non-normal angle, but this situation of two competing deviations can be beneficially used to cancel to zero the otherwise deleterious synchronous deviations arising from our inclining the crystal faces, inclined...
so that direct Fabry-Perot resonances are not formed in the modulator crystal. The modulator used in these studies is indeed of this category, using an ADP crystal 7 x 10 x 70 mm$^3$ in r41 transverse modulation geometry. The ADP material has only a weak Piezoelectric response. Another issue, angular deviation, arises from the Electro-Optic Prism formed if the light beam is not passing through an essentially uniform region of rf field in the crystal. We find fiber out-coupling the modulator is essential to homogenize the modulation across the beam’s aperture[7].

IV. PRINCIPLE OF THE CANCELLATION METHOD

The important new contribution of this present approach to the “Anti-RAM” universe is that our controller is based on the “just-fast-enough” control of the Negative RAM that is deliberately introduced in a Feed-Forward manner. Very likely, the amount of Negative RAM needed to bring the total RAM to zero is almost the same millisecond by millisecond. So a “brick on the gas pedal” approach can maintain a basically zero RAM without needing to be updated by a full-bandwidth feedback system. This cancellation concept breaks the otherwise ubiquitous requirement of a giant bandwidth in exchange for any giant feedback-generated reduction of the RAM. In a normal feedback system, to have very low residual RAM at the modulation frequency f we have chosen, we will need a servo bandwidth much larger than just f, in order to maintain servo stability. In this approach, time delays will quickly form a limit of the feasible stable gain. All feedback methods cost 3 dB in our S/N if the same optical power is detected in the RAM-control and signal-detection photodiodes.

Most importantly, and to emphasize, the “measure-at-f-and-control-at-f system” does not lend itself to closed loop gains of many thousands- or a million-fold, due to the time-delay issues encountered with such a huge servo bandwidth, which is required to provide servo loop stability.

Now consider the slow feedback control of the “gas pedal” that is otherwise noiselessly adding negative RAM to the laser beam. We have the leisure to integrate while checking to see if any RAM is yet visible in our corrected beam’s aperture[7].

V. TOPOLOGY OF ANTI-RAM SERVO SYSTEM

As noted, the current success with the RAM-Buster project can be attributed to its unusual servo topology, in which the amplified and processed error signal is not used as a subtractive input to an error summer. Instead, the amplified error signal is used as a gain control operating on a FeedForward waveform source, which can either cancel – or double – the system’s RAM error, depending on its phase. We know the sign needed by consideration of the sign of the amplified error signal. As usual in servo designs, the desired zero input error is a special place, and roughly corresponds to zero output error. Our error signal, however, is here controlling the “gas pedal”, so a zero input error will correspond to a special magnitude of the FeedForward Cancellation signal. Such an ideal insertion value will be efficiently estimated by feeding the reference wave into an analog multiplier, which will function as our attenuator –plus-sign control for the servo, acting under the control of an integrator, which is a normal part of a servo PID design. This servo integrator’s output will take on a dc offset over time so that the injected Anti-RAM signal exactly cancels the optically-generated RAM, as diagnosed by our photodetector, via its amplifiers and DBM phase-sensitive detector. The tracking gain needed to be increased by an additional integration below 1 Hz, as noted earlier.
VI. TESTING THE PERFORMANCE

A. De-based RAM tests - Methods

How can one test a spectrometer to establish that the RAM is below 1 ppm? Of course an "in-the-servo-loop" measurement will show essentially zero, since the servo will be trying its best to make this always be true. Instead we must look “out-of-loop” to confirm the performance. An ideal such out-of-loop viewpoint is provided by the dc baseline coming from the cavity frequency-locking setup, in the so-called PDH setup[3]. When detuned from any cavity resonances this dc should be zero. If it is not, the laser will be locked with an offset relative to the cavity. Of course for our tests, the laser is not locked to the cavity, and so it may drift near the frequency of some weakly excited higher-order spatial modes of the cavity, and provide huge bursts of nonzero output. But when the room temperature variations are minimal, one can see an extended interval within which the discriminator “zero” is actually challenging to quantify. We use the regular PDH setup, which has ~10 dB “excess” rf gain before the Doubly-Balanced-Mixer (DBM) to minimize unavoidable dc offsets, mostly of thermal origin. Following the DBM and its rf impedance-matching lowpass filter, we use the Input Monitor output of the cavity Servo Loop Filter to increase the scale of the offset voltage, typically a dc monitor gain of 500 was used. An RLC lowpass filter at 10 Hz further reduced the ac noise presented to the DVM for digital recording.

The PDH loop uses a level 7 DBM (Minicircuits ZAD-3+), which has a full dynamic range of ±258 mV (182 mV rms). The slope of the response would correspond to 576 mV rms full scale due to the extra gain, although saturation intervenes. So our 100% AM mark can be taken as 576 mV equivalent. The 1 ppm level would then be 407 µV at the monitored output, reduced by 1.73x due to conversion loss in this DBM output is the same as the one in the PDH locking system, and so the calibration is again 576 mV rms for 100% optical AM. Figure 2d shows such data, taken using 0.25 Hz BW resolution. But even at this narrow BW, shotnoise is totally dominating. So 25 successive sweeps were co-added, giving a noise averaging approximating that of the shotnoise spectrum. So our first calculation might be that the RAM is 1.5 µV rms, and is not particularly notable among the other noise peaks within the shotnoise spectrum. We can conveniently inspect the PDH Photodetector output for RAM at the 1.180 MHz locking frequency by down-shifting the rf frequency with a DBM driven at 1.181 MHz. The so-obtained beat frequency at 1 kHz can then be studied shifting the rf frequency with a DBM driven at 1.181 MHz.

B. dc ResultsUsing the PDH Discriminator

Returning to Fig 2b, we see 21 ¾ hours of data showing the RAM signal in the cavity PDH locking loop, without any interruption by laser drifting to encounter weak modes. The drift slope of the RAM equivalent signal is 94.6 ±1.2 nV per point (5.6 s). During the entire 21¾ hours run this would amount to 1.3 mV, under conditions where 1 ppm RAM would be represented by 235 µV. So the apparent RAM drift is 5.5 ppm over this time: we certainly would have less than 1 ppm change of the RAM-induced offset for ~4 hours.

C. Accuracy results from FFT analysis of the PDH ac error signal

We can conveniently inspect the PDH Photodetector output for RAM at the 1.180 MHz locking frequency by down-shifting the rf frequency with a DBM driven at 1.181 MHz. The so-obtained beat frequency at 1 kHz can then be studied at high resolution using an FFT spectrum analyzer. The conversion loss in this DBM output is the same as the one in the PDH locking system, and so the calibration is again 576 mV rms for 100% optical AM. Figure 2d shows such data, taken using 0.25 Hz BW resolution. But even at this narrow BW, shotnoise is totally dominating. So 25 successive sweeps were co-added, giving a noise averaging approximating that which would have resulted with a 10 mHz resolution. In this case the “RAM” signal was 1.5 µV rms, and is not particularly notable among the other noise peaks within the shotnoise spectrum. So our first calculation might be that the RAM is 1.5 µV, compared to 576 mV for 100%, suggesting the “RAM” signal is 2.6 ppm of the light.

However, it is important to note that this “signal” is basically the same when the PDH detector’s light blocked:
1.2 vs 1.5 µV rms! So evidently we have an rf pickup signal somewhere in the PDH photodetector circuit, remaining, even after improving the shielding and power supply filtering, and removing all identified ground loops. Some next steps involve making phase-sensitive measurements of this “signal” with and without light. At this point, with 2.6 ppm apparent RAM demonstrated, and considering that some analog gains were not allocated ideally, it seems highly likely this system will give locking accuracy at the 1E-6 level, and stability far beyond. Future serious effort is required to avoid non-optical pickup of the modulation rf in the circuits: steel panels 1 mm thick affect the background rf signal, which is present in only one of the two photodetectors. The optical sampling also needs care, as we find the weak light arriving via the optical fiber’s cladding contributes a variable apparent RAM. The fact that Fig 1b shows the RAM-induced dc offset has changed by 1 ppm only in 5 hours is very encouraging for our being able to have our region of stable RAM to be at zero in the near future. A separate publication will explore the total absence of the instability effects usually encountered in Multi-Degree-of-Freedom servos, when operated at high gain. The essential point is that our circuit compares the error to the exact same phase reference which is used to create the Feed-Forward correction signal. Modest variations of the alignment of our two channels with the I and Q phases of the optical signal are basically not important, as our correction space of two amplitudes separated by 90° can completely represent any RAM error to be corrected.

VII. CONCLUSIONS

With a RAM cancellation method of reaching shot-noise sensitivity level, we demonstrate here 2.6 E-6 RAM-induced offset levels for hours, but with a reserve precision sufficient to still be free of RAM-induced problems when the bandwidth has been narrowed to some tens of milliHz, by using extended averaging times, co-addition of multiple scans, or FFT-based direct spectral resolution. Basically this paper announces the RAM-Buster approach long needed to achieve the shotnoise-limited spectroscopy as had been anticipated in our 1983 PDH paper with Drever.

REFERENCES

Carrier Phase and Pseudorange Disagreement as Revealed by Precise Point Positioning Solutions

Demetrios Matsakis, U.S. Naval Observatory (USNO)

Demetrios Matsakis
U.S. Naval Observatory (USNO)
Washington, DC USA
demetrios.matsakis@usno.navy.mil

Zhiheng Jiang
Division of Time, Frequency, and Gravimetry
International Bureau of Weight and Measures (BIPM)
Paris, France

Wenjun Wu
National Time Service Center of China (NTSC)
Linton, China

Abstract In GNSS data reduction, carrier phase (phase) and pseudorange (code) data are complementary. An illustrative example of their interplay is provided, and then it is shown that frequency biases in phase data can be estimated by examination of the difference between the code and phase residuals in Precise Point Positioning (PPP) solutions. Apparent frequency biases, in some cases approaching 0.2 ns/day have been found, although many are an order of magnitude less. These frequency biases could be due to small design imperfections in the GPS receivers. We have also noted that PPP processing is sensitive to the relative weights given the pseudorange (code) and the phase in the sense that down-weighting the code by a factor of 10,000 is preferable to down-weighting by a factor of 10 billion; we think we understand the reason for this.

INTRODUCTION

In GNSS carrier phase solutions, the high precision of phase data results in their being typically weighted >=10,000 times more than the code data. Therefore, the phase data dominate in the determination of most parameters including orbits, atmosphere, Earth Orientation, site position, and the clock frequencies. Clock times however, cannot be determined by the phase data because of the unknown ambiguities and therefore the code data determine the average values of the integrated clock frequencies – in essence providing the constant of integration for integrating frequency to time, which is also equivalent to setting the average ambiguity. Phase and code data are therefore complementary in use although not entirely independent of each other.

The interplay between these two kinds of complementary data is the theme of this paper. The phase and code are measured differently inside the receiver, and we report a method of checking fielded receivers independently of each other.

We note with pleasure that although the technique described here may be new in some ways, similar analyses have been presented earlier by Marc Weiss [1, 2], who himself credits still earlier work.

AN EXAMPLE OF THE INTERPLAY

An experiment was reported at the 2014 ION-PNT [3] wherein PPP solutions were generated from data in which the code and phase of just one satellite (PRN1) were manually offset by 10 ns from their measured values by editing the RINEX files. In this instance, the positions and troposphere values were not affected, nor were the phase residuals. Rather, 96-98% of the 10 ns was absorbed into the code residuals of PRN1, while the receiver clock errors varied by 220-390 ps and the ambiguity errors were of similar magnitude but opposite sign.

The explanation is that the ambiguities served to align PRN1’s data with the other satellites so that clock frequencies would be unperturbed. The overall clock time would be set by an average of all satellites, so that PRN1’s 10 ns offset perturbed the answer by roughly 1/32 of its value, and the remaining 31/32 were absorbed into PRN1’s code residuals.

A COMMON CLOCK/ANTENNA EXAMPLE

Figure 1 shows the difference between PPP solutions for the time of two receivers of the same make observed in common clock/common antenna mode. Using the NRCan PPP analysis package, each day’s values were extracted from the middle day of Kalman filter solution based on averaging the results of the forward and backward passes of a 7-day “round-trip” solution. This approach is known to reduce day-boundary issues considerably, because the time (average of integrated frequency) is based on seven days of code-frequency differencing instead of just one. Since the two receivers shared both clock and antenna, their frequency difference would be expected to be zero and most modelling or solution errors would be expected to cancel as well. The fact that many errors neatly cancel enhances the
prominence of the sawtooth pattern. Note that although the frequency offset between the two receivers exists over a day, the time average of each day is roughly the same as the previous day’s average. The explanation is that the sawtooth is the result of the frequency of the phase data being recorded differently in the two receivers (and therefore erroneously in at least one of them). Since the daily time averages did not vary, we infer that the code data had no frequency difference as measured.

Figures 2-6 support this explanation; they were generated from the raw RINEX (raw data) files without use of the NRCan PPP package. Data for each signal and epoch were extracted from the RINEX files of each receiver, and the difference between the L1 and L2 carrier phase signals were computed. Also, the L1 and L2 differences were averaged as weighted by ionosphere-removal process to create “L3” values for each satellite track. The L1, L2, and L3 values for each satellite track were then individually and independently fit for offsets and slopes. The offsets are related to the ambiguities and biases; they are not of interest here. The fitted slopes correspond to a frequency offset between the phase data of the two receivers. Although considerable noise is present, the frequency offsets are definitely not zero. A firmware change on MJD 56010 greatly reduced the difference, but did not entirely eliminate it. This indicates that the problem is in the receiver design and not due to the data reduction process or satellite signals.

THE METHODOLOGY FOR THE STAND-ALONE ANALYSIS

In our methodology, multitag PPP solutions using the NRCan PPP package [4] were generated from a variety of geodetic GPS receivers whose data were analyzed completely independently. The phase residuals from the PPP solutions were subtracted from the code residuals of each multiday solution. The data for each complete satellite track were then fit for an offset and a rate. Ideally this would be done separately for each satellite track, however the goodness of the fit is vastly improved if just one offset and slope parameter are fit for all residuals. Either way, the offsets would be related to the ambiguities and biases, and discarded. The rates were retained for study, and they represent the frequency offsets of the phase data. Ten–day and fifty-day averages over all satellites are presented in this paper; un-averaged data are noisier although they could potentially be processed so as to yield other kinds of information [5].

This technique is independent of effects that would affect the code and the phase data equally, such as the orbits, clocks, troposphere, site positions, and Earth orientation. It is independent of the ionosphere to the extent that the dual-frequency pre-processing removed its effects. It is not independent of multipath, second-order ionosphere, or the phase wind as the satellite rotates in orbit [6], but

neither code nor phase multipath would be expected to vary linearly over a satellite track, on the average and phase-wind is removed within the PPP package. It would also not be independent of environmental effects, such as temperature which typically but not always affects the code more than the phase. For sites in the Americas, the temperatures and second-order ionosphere effects would be expected to always be largest over the last six hours of any UTC-day (from 18:00 to 24:00). However, there should be little effect because in the 7-day and longer analyses reported here there would be almost as many tracks terminating at a temperature maximum as starting at one.

In the case of second-order ionosphere effects, the total error during the severe ionosphere storm of October 30, 2003 was estimated to be of order a hundred ps in the slant line-of-sight [7], while the effect on the clocks was of order 10 ps [8].

RESULTS WITH 7-DAY SOLUTIONS

Figures 7-20 show the daily rate averages over time, for selected receivers. Some temporal variations are apparent. The above-mentioned unit that changed its behavior due to a firmware upgrade (receiver Y in Figures 3-6) is unit 38. Figure 21 shows the time-averaged satellite slope of each unit, grouped by manufacturer. The one-sigma limits for each receiver are shown as an envelope about the points, and computed from the scatter in the fit residuals. No brand was immune to the effect, and units of the same make showed variations in performance.

One future approach for verification would be to look for solution-boundary jumps in long-term solutions such as the monthly solutions generated by the BIPM.

Since the details of receiver design are proprietary, our speculation as to the means of improvement is limited to general statements such as the need to improve the phase lock parameters. Given the current situation, it is possible that receiver biases can be adequately compensated by parameterization of the code-phase frequency bias within the PPP solution, or in a similar post-fit procedure, but this has not been explored.

FINDINGS WITH 34-DAY SOLUTIONS

In order to study this effect further, we studied reductions of all the data contributed to the BIPM participating labs, for the months of October through December, 2014. This section is to be considered preliminary, until a full understanding is achieved.

Initially, using the default procedures used by the BIPM for PPP, we found very small slopes in the code-phase residuals of the weighted forward and backward solutions, outputted in the PPP package with the identifier “BWD”
(Figure 22). However, we found the slopes would appear by setting the weight given the code to the USNO default (Figure 23). In the NRCan package, the weights are given by the inverse square of the Pseudorange Sigma (PSIG, which the USNO set to 5 while the BIPM set to 1) and the Carrier Phase Sigma (which was set to .01 by both institutions). In Figure 23, one laboratory showed no slope but did display a large constant offset. This large offset is the “memory” of an ambiguity jump that occurred in the filter’s forward pass; such effects limit the power of this technique.

We explain the dependence upon the relative code weights using the example of Figures 24-26, which show the 86,653 ambiguities over the 1625 satellite passes observed by the receiver NIST in the December 2014 solution. Because the troposphere, site vertical, and clock/ambiguity parameters are correlated, our PPP solutions allow the ambiguities to float. In the backward pass, the initial values are determined by the forward pass. Figure 26 shows that the ambiguities can vary by tens of picoseconds over a tenth of a day, and also that many points do not contribute to setting the ambiguity difference between tracks (although the code contributes to the clock and ambiguity values of all points). It is clear that even the largest observed slope of 200 ps/day could be absorbed within the ambiguity variations shown, and therefore the code data could correct for the receiver’s phase bias if the data’s time-range was large enough to provide an adequate lever-arm.

In order to find other possible explanations for the receiver’s apparent phase frequency bias, we also considered the residuals as a function of satellite direction. The receiver NIST is located in a highly asymmetric geographic location, with mountains to the west and much flatter topography to the east. An asymmetric unmodelled troposphere between the east and west directions would cause rising satellite’s ambiguities to be set so as to bring about agreement with phase data from satellites that are setting over the mountains. This would lead to a frequency variation over time. Figure 27 shows that there is a nonzero and constant code-phase difference between east and west for NIST. However, Figure 28 shows that NIST and IP02 (in Portugal) have the same east-west asymmetry in magnitude and sign. However, their slopes are of opposite sense.

By comparing the timing data to Two Way Satellite Time Transfer (TWSTT, also termed TWSTFT), it is shown that the 34-day solutions giving the code higher weight make a better match in cases where there is an apparent receiver phase bias (Figures 29-31). This is what would be expected if the code is correcting the phase frequency bias.

Although the highly underweighted code solutions are not optimal for generating clock differences over 34-day periods, they still could provide a means of checking for frequency bias in the receiver’s phase data, which if present could contaminate PPP solutions on daily or weekly periods. Figure 32 shows the results on most of the receiver data contributed to the BIPM for the months of October through December, 2014.

We note that PPP solutions with integer ambiguities might be more sensitive to receiver phase variations. Also, analyses based entirely on the RINEX files, without any PPP package but incorporating the phase-wind corrections, should accomplish the same thing.

CONCLUSIONS

While this is still a work in progress, we have found some receivers contributing data to the BIPM have a frequency bias in their phase. Although the BIPM’s current 34-day data reduction scheme is not very sensitive to them, problems would be found in analyses covering shorter time periods. The effect of weighting code data has been explored.

A similar paper will be given at the April 2015 ION-PNT meeting, and any progress since this submission will be reflected in the proceedings of that conference. Those proceedings have no page limit, and the full set of figures will appear there.

DISCLAIMER

USNO, BIPM, and NTSC as a matter of policy do not endorse any commercial product. Any information that might enable manufacture identification is provided for scientific clarity only. We further caution the reader that the performances reported herein may not be characteristic of any receiver currently marketed, and could perhaps be dependent upon their configuration or on the ancillary equipment. Another possibility is that our software contains a bug as implemented, and this will be tested through the use of other software [9-13].

ACKNOWLEDGMENTS

We thank Stephen Mitchell for generating the USNO PPP solutions and for many helpful discussions, along with Christine Hackman, Francois Lahaye, Ed Powers, Judah Levine, Victor Slabinski, and Jian Yao.

REFERENCES


719
Applications Meeting, December 2012, Reston, Va


---

Figure 1. PPP clock difference between two geodetic receivers

Figure 2. Difference in slopes of satellite tracks at the L3 frequency (2.54*L-1.54*L2). Each point represents the slope of a completed satellite track at its midpoint.
Figure 21. Average slope of code-phase, over all complete satellite tracks, for each receiver studied. All units between two vertical markers have a common manufacturer. One-sigma limits are indicated by the continuous curves; therefore the large variations in units 24 and 25 are not significant. Unit 38 is receiver “Y” in figures 2-5, and is based only upon data since the firmware upgrade.

Figure 23. Code-Phase residuals using the same processing as the previous figure, except that the weight of the code was decreased by a factor of one million. Note that three laboratories now display a systematic frequency offset. The large constant offset of one laboratory is the memory of an ambiguity jump in the forward pass.

Figure 26. A very small portion of the previous figure. The solution proceeds in the backwards direction, so the initial ambiguities initially vary considerably as the approach maturity. The final ambiguities are not applied to the entire satellite track in these solutions.
Figure 28. Code-phase residuals for two receivers that display opposite slopes, and the very similar difference between their easterly and westerly averages. The upper two curves were shifted for display.

Figure 29. Timing difference for PTB-NIST measured three different ways. The green curve is TWSTT data, with diurnals apparent. The other two ways are PPP solutions, and the one that gives best fit to the TWSTT data is the one in which the code is given the higher weight.

Figure 32. Slope of residuals in solutions for October, November, and December 2014. Receivers between vertical markers are of the same time. Each receiver has a unique abscissa-value. If all three months provided reasonable data there will be three points for that receiver. The formal errors are comparable to the dot size. A spread between points for the same receiver could indicate a change of that receiver’s properties.
Long-Term Uncertainty in Time Transfer Using GPS and TWSTFT Techniques

Victor Zhang, Thomas Parker, Jian Yao
Time and Frequency Division, National Institute of Standards and Technology (NIST), Boulder, Colorado, U.S.A.

Abstract—The techniques of GPS time and frequency transfer (code based and carrier phase) and TWSTFT are widely used in remote clock comparison and in the computation of TAI and UTC. Many timing laboratories in the world utilize both techniques (GPS and TWSTFT transfer links) to compare each other’s clocks. A time link must be calibrated to assure the time transfer accuracy. In many cases, calibration campaigns have been very infrequent due to the expense and lack of suitable equipment. In lieu of repeated calibrations, some information regarding the long-term stability of these links can be obtained through comparisons between the two links (a so-called double difference). Without frequent calibrations it is impossible to tell where the instabilities originate, but information regarding the magnitude of the instabilities can be obtained from double difference data. We have been investigating the combined variations of GPS and TWSTFT links for a number of laboratory pairs, including both long and short baselines. Our results show that the relative change between GPS and TWSTFT transfer links can be as large as 6 to 7 ns over a few years and that all of the laboratory pairs that have been investigated show similar magnitudes in the double difference data. Currently the longest set of good double difference data is about 7 years. The study results point out the need for frequent calibration campaigns if accuracies at the nanosecond level are required.

Keywords—time and frequency transfer; GPS carrier-phase time and frequency transfer; Precise Point Positioning; Revived Rinex-Shift Algorithm; two-way satellite time and frequency transfer; time transfer link calibration; Type A and Type B time and frequency transfer uncertainty.

I. INTRODUCTION

Time and Frequency transfer is used to compare remote clocks or frequency standards. It is also an integral part of the generation of International Atomic Time (TAI) and Coordinate Universal Time (UTC). Timing laboratories around the world use time and frequency transfer to contribute data from their clocks and primary frequency standards to the computation. The International Bureau of Weights and Measures (BIPM) computes TAI and UTC. The monthly Circular T publication [1] reports TAI – TAI(k) and UTC – UTC(k), where TAI(k) and UTC(k) are a laboratory k’s real-time realization of TAI and UTC.

According to the Circular T 326 published on March 10, 2015, 69 of the 71 contributing laboratories used Global Positioning System (GPS) code and carrier-phase time and frequency transfer [2, 3, and 4] and two-way satellite time and frequency transfer (TWSTFT) [5] to transfer their clock data to the computation. Many laboratories in Asia, Europe and the United States employ both of the techniques or transfer links for remote clock comparisons and the TAI/UTC computation. The uncertainty of each link is a combination of the Type A and Type B uncertainties. The Type A uncertainty is mainly introduced by the stability of the time and frequency transfer technique used, and the Type B uncertainty is a measure of the time transfer accuracy, which is dominated by the uncertainty of link delay calibration. The typical Type A uncertainty for the GPS carrier-phase links (LinkGPSCP) and the TWSTFT links (LinkTW) is 0.3 ns. We will focus on the LinkGPSCP for GPS time and frequency transfer in this paper. The Type B uncertainty of a link depends on several aspects, such as how the link was calibrated. In recent years, many successful link calibration campaigns have reported the calibration uncertainty at about 1 ns using traveling dual-frequency GPS receivers with the GPS carrier-phase solutions and using mobile TWSTFT stations [6, 7]. The Type B uncertainty for these recent calibrated links is from 1 to 1.2 ns in Circular T 326.

The Type B uncertainty of a time transfer link is also associated with effects other than calibration uncertainty. After a calibration, any change in the link, such as the delay change due to equipment aging or malfunction, will change the calibration result and therefore increase the uncertainty. Therefore, it is necessary to have frequent calibrations in order to keep the Type B uncertainty of a time transfer link as close to the calibration uncertainty as possible. However, some laboratories’ link calibrations have been very infrequent (no calibration for two and more years) due to the expense and lack of suitable equipment. In lieu of repeated calibrations, some information regarding the long-term stability of LinkGPSCP and LinkTW between two laboratories can be obtained through comparisons between the two links (a so-called double difference). Because the LinkGPSCP and LinkTW between two laboratories compares the same pair of remote clocks, the double difference removes the clock difference, and reveals the combined relative change between the two links. Although it is impossible to tell where the changes or instabilities originate, information regarding the magnitude of the instabilities can be obtained from double difference data.
In this paper, we use the double difference technique to study the long-term uncertainty in \textit{LinkGPSCP} and \textit{LinkTW} among the National Institute of Standards and Technology (NIST) in Boulder, Colorado, the LNE-SYRTE, Observatoire de Paris (OP) in France, the Physikalisch-Technische Bundesanstalt (PTB) in Germany, and the U.S. Naval Observatory (USNO) in Washington, DC. The four timing laboratories are chosen because all of them have had stable \textit{LinkGPSCP} and \textit{LinkTW} data for more than four years. The four laboratories also enable the study to cover the time transfer links between the timing laboratories in Europe and the USA as well as the links within both Europe and the USA. Section II shows how we prepared the \textit{LinkGPSCP} and \textit{LinkTW} data for the study. We then present the study results in Section III and summarize the study in Section IV.

II. THE GPS CARRIER-PHASE, TWSTFT AND DOUBLE DIFFERENCE DATA

We use the BIPM TAIPPP [8] solutions as the \textit{LinkGPSCP} in remote clock comparisons. The TAIPPP solution uses the precise point positioning (PPP) technique to obtain difference between a GPS receiver’s reference clock (REF) and rapid product of the International GNSS Service Time (IGRT). The BIPM started the TAIPPP process in April 2008. The monthly TAIPPP solution for a laboratory contains the REF-IGRT difference over a 35-day or 40-day period. We can apply the delay correction involved in the GPS carrier-phase measurements to each of the TAIPPP solutions and then difference the two laboratories’ delay-corrected TAIPPP solutions of the same time stamp to obtain the time difference of the \textit{LinkGPSCP} between the two laboratories. The TAIPPP solutions contain a data boundary discontinuity due to noise of the pseudo-range measurements. To study if the 35-day or 40-day data boundary discontinuity affects the long-term uncertainty of the \textit{LinkGPSCP}, we compared the TAIPPP results to the Revised Rinex-Shift PPP (RRSPPP) results [9, 10]. The RRSPPP is an algorithm developed at NIST to minimize data boundary discontinuity and to handle data anomalies, which are achieved by continuously processing multi-day data batches with the successive data batch advanced to one day later, and producing a PPP carrier-phase solution at the midpoint of each multi-day data batch. Fig. 1 shows the double difference of TAIPPP - RRSPPP for comparing UTC(NIST) and UTC(PTB) over a period of more than six years. The TAIPPP agrees with RRSPPP to within ±0.5 ns most of the time, indicating that data boundary discontinuity from the TAIPPP solutions will not deviate for more than 1 ns in the long-term stability of \textit{LinkGPSCP} study.

NIST, OP, PTB and USNO all participate in the transatlantic TWSTFT. OP and PTB also take part in the Europe-to-Europe TWSTFT. We do not have a direct TWSTFT link between NIST and USNO. The TWSTFT between NIST and USNO are obtained from the difference of [UTC(NIST) – UTC(PTB)] - [UTC(USNO) – UTC(PTB)]. The regular TWSTFT measurements are made during even hours, 12 times a day. Each laboratory’s TWSTFT measurements are reported in a daily file in the format according to the International Telecommunication Union (ITU) recommendation, ITU-R TF.1153 [11]. In addition to the measurements, the file contains information of link calibrations, delays of the local reference signal and TWSTFT equipment. When we compute the TWSTFT difference of two remote clocks, we difference the two TWSTFT measurements with corrections of link calibration and delays of reference signal and equipment of each TWSTFT station.

Fig. 1. Double difference of the daily averaged TAIPPP – RRSPPP for the UTC(NIST) – UTC(PTB) comparison. Data period is from November 2008 to January 2015.

The reference signals for both \textit{LinkGPSCP} and \textit{LinkTW} at NIST, PTB and USNO are derived from UTC(NIST), UTC(PTB), and UTC(USNO), respectively. Each of the links’ delay correction is available for comparisons between two laboratories. However, the reference signals for OP’s GPS carrier-phase and TWSTFT links are directly from a hydrogen Maser clock, and sometimes from two different hydrogen Maser clocks. The delay corrections for converting the reference Maser clock to UTC(OP) are included in the TWSTFT data files, but we don’t have the information for OP’s TAIPPP data. To cancel the OP’s reference clock (OPH) in the double difference, we first compute the daily averaged
TWSTFT and TAIPPP data, then difference the daily averaged TWSTFT and TAIPPP data when both used the same reference clock, and finally remove the estimated time steps due to the different delays in the TWSTFT and GPS carrier-phase measurements. Fig. 2 shows the UTC(NIST) and OPH comparison result obtained from the procedures described above. The result contains many larger than 5 ns outliers from unknown causes. The outliers do not obscure the long-term trend between the $Link_{GPSCP}$ and $Link_{TW}$. Thus, we will disregard the outliers and use the cleaned-up double difference in the analysis.

In the next section, we analyze the $Link_{TW} - Link_{GPSCP}$ double difference for comparisons of NIST/OP, NIST/PTB, USNO/OP, and USNO/PTB over the transatlantic baseline and NIST/USNO and PTB/OP over the United States and Europe baselines. The double difference of the NIST/PTB comparison covers the period of MJDs from 54553 (March 28, 2008) to 57051 (January 29, 2015). The NIST/OP comparison has the second longest double difference stretch (MJDs 54874 – 57051, February 12, 2009 – January 29, 2015). The data used in the PTB/OP comparison is over MJDs 55104 – 57051 (September 30, 2009 – January 29, 2015). The USNO TWSTFT facility finished renovation at the end of 2010. The double differences involving USNO start on MJD 55562 (January 1, 2011) and end on MJD 57051 (January 29, 2015).

III. THE LONG-TERM STABILITY OF GPS CARRIER-PHASE AND TWSTFT LINKS

The double difference of comparisons among NIST, OP, PTB and USNO are grouped in Fig. 3 through Fig. 5. The figures show the $Link_{GPSCP}$ and $Link_{TW}$ can differ by more than 1 ns relative to each other over a one-year period. From MJD around 56261 (December 2012) to MJD 57051 (January 2015), the double differences for the UTC(NIST) – UTC(PTB) and UTC(NIST) – UTC(USNO) in Fig. 3 show about a 6 ns decrease, while the double differences for the UTC(USNO) – OPH and UTC(PTB) – OPH comparisons in Fig. 4 show about a 3 ns increase. On the other hand, the double differences for the UTC(NIST) – OPH and UTC(USNO) – UTC(PTB) comparisons do not have big changes except for the about 2 ns change around MJD 56291 in the UTC(NIST) – OPH comparison and the about 3 ns change after MJD 56940 (October 2014) in the UTC(USNO) – UTC(PTB) comparisons. There is no evidence that the relative changes between $Link_{GPSCP}$ and $Link_{TW}$ are baseline related.

The changes in Fig. 3 could come from a decrease in the NIST $Link_{TW} - Link_{GPSCP}$. It is also possible the downward change is caused by an increase of $Link_{TW} - Link_{GPSCP}$ from USNO and PTB and that both laboratories change in the same direction by similar amount. This possibility can be seen in Fig. 4 for the upward change in the double differences for the UTC(USNO) – OPH and the UTC(PTB) – OPH comparisons. However, we can also argue the upward changes are caused by a decrease in OP’s $Link_{TW} - Link_{GPSCP}$. There is no obvious upward or downward change in Fig. 5 for the double differences of UTC(NIST) – OPH and UTC(USNO) -
UTC(PTB) comparisons. The result does not necessarily mean there is no relative change in the LinkTW and LinkGPSCP between NIST and OP or between USNO and PTB. The LinkTW - LinkGPSCP changes for comparisons between NIST and OP or between USNO and PTB can be canceled if the changes are in the same direction and at about the same magnitude.

With the double differences, we are able to see that LinkTW and LinkGPSCP are changing relative to each other over the period of study, but unable to answer the question of which of the links of a remote clock comparison is the main source of the change. A calibration is needed to identify which link has changed. Fig. 6 and Table I show the record of the USNO mobile TWSTFT calibration of UTC(NIST) – UTC(USNO) [12]. USNO has been doing the calibration twice per year since 2011 to keep the two time standards in the United States as close as possible. For the calibrations in the past four years, the NIST/USNO TWSTFT via PTB agreed with the calibrations within ±1 ns, but the NIST/USNO TAIPPP is off the calibrations by 6.2 ± 2 ns at these calibration points. In Fig. 3b the 3 ns drop in the last year of data appears to be mostly from the GPS carrier-phase link.

IV. CONCLUSIONS

We used the double difference to study the long-term time transfer uncertainty using GPS carrier-phase and TWSTFT techniques for comparisons among NIST, OP, PTB and USNO. The GPS carrier-phase link and TWSTFT link for each pair of the six comparisons (NIST/OP, NIST/PTB, NIST/USNO, PTB/OP, USNO/OP and USNO/PTB) all show changes relative to each other over time. There is no evidence the changes are related to the baseline of the comparisons. The changes can be more than 1 ns over a one-year period and reach about 6 ns over a longer period of time. The change contributed by individual laboratory could be canceled or added in the double difference if the changes were in the same direction with similar magnitude, or in the opposite direction. Only link calibrations can check if a time transfer link, either using GPS carrier-phase or using TWSTFT, has changed with respect to the last calibration result. Although calibrations using traveling GPS receivers and mobile TWSTFT can achieve calibration uncertainty of 1 ns, time transfer uncertainty is equal to the calibration uncertainty only at the time of calibration and it increases as time goes by. For time transfer using GPS carrier-phase and TWSTFT, we need a link calibration at least once a year in order to achieve time transfer uncertainty at the 1 or 2 ns level.

REFERENCES


Quality Factors of Quartz Crystal Resonators Operating at 4 Kelvins

Serge Galliou*, Philippe Abbé*, Maxim Goryachev†, Michael. E. Tobar†, and Roger Bourquin*

*Time and Frequency Department, FEMTO-ST Institute (UMR 6174, CNRS, ENSMM, UFC, UTBM), Besançon, France. serge.galliou@femto-st.fr
†ARC Centre of Excellence for Engineered Quantum Systems, School of Physics, the University of Western Australia, Crawley, WA, Australia. maxim.goryachev@uwa.edu.au, michael.tobar@uwa.edu.au

Abstract—Quartz crystal resonators can exhibit huge quality factors close to 1 billion at liquid helium temperature. Nevertheless, they must satisfy a set of conditions to meet this high level of resonance. With the help of experimentation, the main identified conditions are considered in this paper, such as the material quality, the energy trapping through the vibration modes and resonator diameters, and the electrode effect.

I. INTRODUCTION

Recently, quality factors greater than 1 billion have been measured on bulk acoustic wave quartz crystal resonators working at liquid-helium temperature [1]. This feature offers the opportunity to use such low-loss resonators as acoustic cavities in quantum hybrid systems or similar physical experiments [2]. Beyond this specific use, such high-Q resonators are obviously also attractive for applications involving frequency sources. Nevertheless, all the tested resonators do not exhibit Q-factors greater than 1 billion at 4 K. Indeed, Q-factors of some of them can even be limited to just a few tens of millions. This depends on various factors or conditions that are discussed in this paper. The tested resonators are state-of-the-art devices, typically SC-cuts initially optimized to work at 5 or 10 MHz at room temperature. Although some generalities are available for other material than quartz, the latter is the reference material in this paper and, in addition, without any specific mention, given data referred to the SC cut, i.e. the well-known doubly rotated cut where both (quasi-pure) shear modes and a (quasi-pure) expansion mode can be electrically excited.

Firstly, a short review reminds that the devices under test are operating in the Landau-Rumer regime for which the usual relationship $Q \times f = const.$ does not hold anymore. Then the analysis is mainly based on a set of experimental data extracted from resonators tested within [3.5 K 12 K] over a wide range of overtones of $A$, $B$, and $C$ modes, up to 300 MHz, and with different resonator assemblies, the resonator being an electroless part or not. The material quality is also analyzed, and finally, nonlinearities are considered with regard to the issue of the energy stored inside the device.

II. SHORT REVIEW

The propagation of an acoustic wave in a resonator is coupled with losses that are usually described by the device quality factor $Q$, i.e. the inverse of the losses $1/Q$. These losses can be sorted into intrinsic losses in an ideal medium including phonon-phonon interaction, thermoelastic effect, and engineering losses like energy losses inside the holders, scattering due to the surface roughness or absorption due to impurities inside the medium. In any case the resulting losses $1/Q$ are the sum of all these individual losses $1/Q_i$:

$$1/Q = \sum 1/Q_i$$

Actually it should be mentioned that thermoelastic losses do not exist for shear modes like the $B$ and $C$ modes in a quartz SC cut because there is no successive expansion and compression areas. By definition such areas exist for the longitudinal mode, the $A$ mode, but, nevertheless, thermoelastic losses are no longer significant for frequencies typically greater than a few megahertz like in this paper.

In the Time and Frequency community it is quite usual to set that the product $Q \times f$ is a constant for a given material vibrating on a given mode. This constant is about $1 \times 10^{13}$ for quartz. Although this common feature is available at room temperature, it does not hold anymore at cryogenic temperatures, at least when losses from phonon-phonon interaction dominate. Indeed, because of the increase of the thermal phonon lifetime, mechanisms of phonon-phonon interactions at cryogenic temperatures are different from those at 300 Kelvins. When the temperature goes down below 10 Kelvins, the thermal phonon lifetime $\tau$ increases and achieves the condition $2\pi f \tau > 1$: the absorption coefficient $\alpha(f)$ of the propagating wave is changed, and in turn, the $Q$-factor.

At room temperature, the absorption coefficient $\alpha(f)$ is proportional to $f^2$. As a consequence, because the $Q$-factor is proportional to $f$ and inversely proportional to $\alpha(f)$, it turns that the product $Q \times f$ does not depend on $f$. This is the Akheiser regime [3].
At liquid-helium temperature, the absorption coefficient $\alpha(f)$ becomes proportional to $T^4 \times f$. Thus, $Q$ does not depend on $f$ anymore and behaves versus temperature according to $1/T^4$, according to the Landau-Rumer theory [4].

It can be mentioned that, actually, when recording data to calculate the Q-factor against temperature within [3 K, 10 K] - the experimental set-up as well as the procedure to calculate Q-factors have been described in [5] [6]- one can observe that the behavior is more complicate than the simple $1/T^4$ law, or more exactly this law does not hold all over the range, as shown in Fig. 1. First of all, values measured at 4 K should be compared with those at 300 K; one may remind that high-quality SC-cut quartz crystal resonators, typically designed to work at 5 MHz (respectively 10 MHz) on the $3^{rd}$ overtone (OT) of their $C$-mode, exhibit quality factor of about $2 \times 10^6$ (respectively $1 \times 10^6$) at room temperature. At 4 K, values are typically more than ten times greater than those measured at 300 K (see Fig. 1). Moreover, it can be noticed that the Q-values behave as $1/T^n$ with $n \approx 4$ for both shear modes but $n \approx 6$ for the extensional mode, the $A$ mode, for $6K < T < 10K$. For $T < 6K$, $n$ decreases dramatically to be close to 0.3, i.e. 1/3, which could be attributed to a two-level-dependency (TLS) effect [7].

In addition, as shown in Fig. 1, the $A$ mode exhibits higher Q-values than the $B$ mode, and $B$-mode values are greater than those of the $C$-mode. This can be explained by the fact that a more or less large part of the vibration energy is lost inside the lens holders ($1/Q_1$ due to losses into the holders). Indeed, results of Fig. 1 are from trapped energy resonators, and it can be demonstrated that the $C$ mode is less trapped than the $B$ mode which is itself less trapped than the $A$ mode [8] [5].

III. EXPERIMENTAL DATA

A. Quality factors around 4K versus frequency

Various types of SC-cut resonator technologies have been tested: some of them are electrode-deposited resonators whereas others are non-metalized resonators, i.e. BVA-type resonators. The latter are 24 mm diameter or 13 mm diameter resonators. 5 MHz as well as 10 MHz resonators have been used, i.e. resonators optimized to initially work at 5 or 10 MHz at room temperature, on the $C$-mode $3^{rd}$ OT. Results shown in Fig. 2 are from four different types of structures. The so-called "5 MHz BVA OSA" is a 24 mm diameter 5 MHz BVA-type resonator from Oscilloquartz SA, Neuchatel, Switzerland; "5 MHz BVA Ind." and "10 MHz BVA Ind." are 13 mm diameter resonators - 1 mm and 0.5 mm tick respectively - from BVA Industrie, France (this manufacturer has presently stopped this activity), oscillating at 5 and 10 MHz respectively; "5 MHz Metalized" is an electrode-deposited resonator from FEMTO-ST, France. Q-values of just a selection of overtones are reported for each of the four resonators. All the OTs do not exhibit Q-value as high as expected. Nevertheless, when taking apart the best results of each resonator, a general trend can be extracted from, in three asymptotic behaviors. When starting from low OT frequencies towards higher OT frequencies, Q-values increase first, then reach the maximum values of a given type of resonator, and finally go down for the higher frequencies [1]. Among these three different behaviors, just the middle one can be attributed to a Landau-Rumer regime dominated by a three phonon interaction. Actually, lower OTs are less trapped than higher ones, and as a consequence their corresponding Q-factors are more reduced as the frequency is low. On the other side, when the frequency increases, the surface roughness should be considered and scattering effects occur: Q becomes proportional to $1/f$.

B. Comparison between electrodeless resonator and metalized resonator

In Fig. 2 it appears that losses inside an electrode-deposited resonator seem to be more important than in an electrodeless resonator. Nevertheless, this comparison is done with two different resonators. In order to verify whether or not the electrode deposition can impact the Q-factor, a BVA-type resonator has been first measured, then disassembled and gold
electrodes have been deposited on the resonator part that has been measured again, alone this time. Results are given in Fig. 3.

Fig. 3 clearly demonstrates that electrodes impact the losses. Actually, those results are not so surprising and rather intuitive. Gold is known to be a rather soft material, that is to say an acoustic absorbent material. So this difference in terms of Q-factor between electrodeless and electroded resonators still exists at room temperature but is not detectable for the corresponding “low” Q-values. Thus this difference of acoustic losses is just revealed at cryogenic temperatures. It can be noticed that the energy being proportional to the square of the displacement, the Q-drop does not depend of the mode, i.e. is similar for all the three modes.

C. Comparison between swept and unswept quartz crystal resonator

Scattering can occur when impurities and/or defects exist inside the crystal where the vibration is propagating. Tests have been performed with a resonator made from an unswept quartz crystal. The corresponding results have then been compared with those obtained from a resonator of the same design but made from a swept crystal. A quality factor $Q$ proportional to $1/f^3$ is the signature of such a scattering [9]. According to Fig. 4, beyond this behavior, the Q-factor of an unswept resonator seems to be always lower than that of a similar one made in a swept material, even at low frequency. Nevertheless, the comparison of results from Fig. 4 should be read with care because the reference resonator is particularly remarkable. Such results should be confirmed by additional measurements on a set of devices.

D. Nonlinearities

Because of the (extremely) high Q-factors, the issue of the dissipated power inside the resonator should be discussed. Measurements described above have been made by exciting the resonator with a feeding power as low as possible. Otherwise, thermal effects can be observed. Indeed, the frequency is shifted all along the measuring time that can take a few minutes (the sweeping time should be very long) when Q-factors are so high. In addition, nonlinearities can induce de
faults on Q-factor measurements. Fig. 5 shows a set of records of the resonator impedance phase for different excitation powers.

IV. CONCLUSION

As a matter of fact, it is shown that Q-values differ from one design to another, according to engineering options, as expected or not. In addition, it is reminded that Q-values depend on the vibration mode because of the energy trapping, and of course on the material quality. Measurements at cryogenic temperatures can reveal properties imperceptible at room temperatures. This is the case for swept and unswept quartz resonators that can exhibit approximately the same Q-values at room temperature. The same conclusion is available for electrode-deposited resonators that can reach equivalent features than electrodeless devices at room temperatures but not at 4 K.

In terms of applications, quartz resonators working at cryogenic temperatures could be good candidates for ultra-stable frequency sources. Indeed, one can suppose that high quality factors are correlated with low noise. Nevertheless, temperature compensated cuts should be identified first. Furthermore, more fundamental experiments can be considered such as the quest of the ground state [2], gravitational wave detection [10], or investigations on the isotropy of space [11].

ACKNOWLEDGMENT

The authors would like to thank the Conseil Régional de Franche Comté for its financial support. Special thanks are given to Oscilloquartz SA, Neuchatel, Switzerland for providing resonators, and especially Isabelle Lozach, Jean-Pierre Aubry, Luc Schneller. MG is thankful to the Australian Research Council under grant CE110001013.

REFERENCES

Quality Factor of Bulk Acoustic Wave Resonators at Cryogenic Temper-


"Losses in high quality quartz crystal resonators at cryogenic temper-

of Rayleigh phonon scattering through excitation of extremely high
overtones in low-loss cryogenic acoustic cavities for hybrid quantum

[10] M. Goryachev and M. E. Tobar, "Gravitational wave detection with
90, 102005, 2014.

M. Goryachev and M. E. Tobar, "Testing the isotropy of space using
rotating quartz oscillators," http://arxiv.org/abs/1412.2142v1
Bias Corrections in Primary Frequency Standards

Parker, T. E., Heavner, T. H. and Jefferts, S. R.
Time and Frequency Division
NIST
Boulder, CO USA
tparker@boulder.nist.gov

Abstract—Primary frequency standards serve the function of calibrating the rate of International Atomic Time, TAI, and therefore play a critical role in the accuracy of the world’s time. The Working Group on Primary and Secondary Frequency Standards, WGPSFS, is an advisory body to the Time Department of the Bureau International des Poids et Mesures and to the Consultative Committee for Time and Frequency on matters related to primary and secondary frequency standards that are used to determine the rate of TAI. A current issue being considered by the WGPSFS is establishing guidelines for deciding when and how to make corrections for newly discovered frequency biases in primary frequency standards. This paper is intended to generate discussions on this topic in an audience wider than just the WGPSFS.

Keywords—primary frequency standards; bias corrections; TAI

I. INTRODUCTION

Primary frequency standards, PFS, and secondary frequency standards, SFS, serve the function of calibrating the rate (frequency) of International Atomic Time, TAI, and therefore play a critical role in the accuracy of the world’s time. In a PFS all known frequency biases must be evaluated and, if necessary, corrected. The Working Group on Primary and Secondary Frequency Standards, WGPSFS, is an advisory body to the Time Department of the Bureau International des Poids et Mesures, BIPM, and to the Consultative Committee for Time and Frequency, CCTF, on matters related to primary and secondary frequency standards that are used to determine the rate of TAI. As the uncertainties of PFS decrease with improved technology, frequency biases that were previously insignificant may become more important. A current issue being considered by the WGPSFS is establishing guidelines for deciding when and how to make corrections for newly discovered, or newly relevant, frequency biases in primary frequency standards. This paper is intended to generate a discussion on this topic in an audience wider than just the WGPSFS.

II. SOME HISTORY

A brief historical retrospective is presented here using a few biases that have been included as corrections to PFS that report to BIPM. These include biases that were once corrected but no longer are, biases that were unrecognized until long after the definition of the second and have now been included in all PFS bias tables, as well as corrections that are outside of the definition of the SI (International System of Units) second but are applied for the generation of TAI.

The Millman effect is an example of a physical phenomenon that was once used to explain observed frequency biases in some PFS, but was later shown not to be true. The correction was a generalization of an effect that had long been recognized in atomic beam physics. The generalization was initially accepted and a “correction” applied in spite of the fact that no experimental evidence of the effect existed. Wineland and Hellwig [1] later showed that, in fact, the physics of the effect was incorrectly described and that the frequency shift was not allowed on the clock transitions in Cs. Consequently, the frequency bias is no longer considered in PFS.

On the other hand, the blackbody correction is a good example where a previously unknown frequency bias of significant magnitude was proposed theoretically [2], measurements were made to confirm it, and it was formally recommended by the CCTF in 1996. This bias is unusual in that it actually required a clarification to the formal definition of the second [3]. More details of the blackbody correction are given in Section IV.

When the accuracy of clocks and frequency standards improved to the level where shifts due to relativity needed to be included, there was no consensus on how this was to be accomplished. For example, for some period TAI was incorrectly considered a form of proper time rather than coordinate time. While the gravitational redshift was experimentally verified in the late 1950s, it was not until 1991 that the IAU adopted the specific metric for use in comparing frequency standards which is in use today [4]. The gravitational redshift correction is, in fact, not part of the definition of the second, but is part of the implementation of TAI. It does not depend on the design or operation of the clock, but instead on where it is located.

III. CURRENT SITUATION

The first cesium fountain PFS to report regularly to the BIPM started operation in 1999. Since early 2008, fifteen fountains have reported to the BIPM, of which about eleven report on a fairly regular basis. The situation currently exists where a significant bias correction is being applied to some...
A list of typical biases from recent PFS reports is presented in Table 1. These biases are divided into three categories. Category 1 includes four biases for which frequency corrections, along with appropriate uncertainties, are currently applied to all fountain PFS. Category 2 includes bias corrections and uncertainties that are applied to some fountain PFS but not others. Finally, category 3 includes small biases that are handled by simply adding an additional uncertainty without any corrections being made. These are all very small biases that have negligible impact on the total uncertainty and that do vary among the fountains.

The physics of the biases in Category 1 is well understood and everyone agrees that these biases should be evaluated in each PFS and that appropriate frequency corrections and uncertainties should be applied.

In Category 2 the situation is different. The magnitudes of some biases in Category 2, such as microwave leakage for example, are unique to individual standards and the decisions whether to make corrections or not are made by the operators of the individual standards. The microwave lensing shift [5] in Category 2 is the bias that triggered the current discussion. As PFS have improved, the fractional frequency total uncertainty of TAI in any given month can now be as low as about 2x10^{-16}. The magnitude of the proposed bias is about 0.7 to 0.9x10^{-16} in some PFS, and therefore could potentially pull the rate of TAI by more than 30% of its uncertainty. There is no experimental verification of the microwave lensing bias and the theoretical analysis is not universally accepted [6, 7]. Thus, not all laboratories agree that the correction should be made. The details of the microwave lensing shift are not the issue in this paper, but it is the more general topic of how and when any significant new bias should be uniformly evaluated in, and applied to (if necessary), all PFS.

For now, all of the very small biases in Category 3 are not of any concern.

IV. WHAT SHOULD BE DONE?

To address the issue of new biases several questions need to be asked and answered. If this were a purely academic situation, the issue could be left to resolve itself in the literature. However, this is not simply an academic situation since the accuracy of the rate of TAI is at stake. So the first question is: should the WGPSFS step in and provide some recommended guidelines for uniformly evaluating, and if necessary, introducing significant new biases into the list of biases for which corrections should be applied? Or should the WGPSFS do nothing and let individual laboratories make their own decisions? It is the opinion of the authors that the WGPSFS should be involved.

If it is decided that the WGPSFS should develop recommended guidelines, then there are several more questions to be answered. How large does a bias have to be relative to the total uncertainty of TAI before it becomes a concern? Is it in the range of 10%, 50% or 100%? Does the magnitude of the bias relative to the uncertainty of an individual PFS make any difference as to whether the bias should be corrected for in that standard?

If a new bias is deemed to be of concern, what criteria does it have to meet to be considered valid for evaluation by all PFS? Is experimental confirmation necessary? Can it be of purely theoretical origin? If so, are more than one independent theoretical derivations needed? This may have to be decided on a case by case basis, but some guidelines would be helpful. Clearly, experimental verification is highly desirable, but this can sometimes be very difficult to provide. If the circumstance regarding the bias remains unclear, the best alternative may be

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cesium Fountain Bias List</td>
</tr>
</tbody>
</table>

### Category 1

Bias corrections, with uncertainties, made on all fountain PFS.
- Second order Zeeman effect
- Blackbody shift
- Atom density (spin exchange, cold collisions)
- Gravitational red shift

### Category 2

Bias corrections, with uncertainties, made on some fountain PFS.
- Microwave lensing
- Distributed cavity phase shift
- Cavity pulling
- Microwave leakage

### Category 3

Biases covered by increased uncertainty (no corrections made).
- Rabi, Ramsey pulling
- Microwave spectral purity
- Majorana transitions
- AC Zeeman (heaters)
- Fluorescence light shift (AC Stark shift)
- DC Stark shift
- Background gas collisions
- Bloch-Siegert shift
- Second order Doppler
- Electronics
to recommend that an additional uncertainty be added to the uncertainty of TAI without making any PFS corrections.

The situation with the blackbody correction is a good example of what can happen. The theory was first presented in 1982 [2] but the correction was not applied to all PFS until mid-1995 and only later was direct experimental verification obtained [8]. The blackbody fractional frequency bias is nominally $2 \times 10^{-14}$ in room temperature standards, yet total uncertainties of some PFS at the time ranged from 1 to $3 \times 10^{-14}$. There was a period of at least a year prior to mid-1995 in which this significant correction was made on some regularly reporting PFS and not others, even though all PFS had about the same bias. During this period, the uncertainty of TAI was increased above that calculated by the standard procedure to $2 \times 10^{-14}$ in order to handle this inconsistency. Clearly, this is the type of situation that the PFS community would like to avoid.

In the blackbody situation there was a period of time in which the scatter in the data was not consistent with the stated PFS uncertainties. This is not a circumstance unique to PFS. It is not uncommon in many areas of science to have inconsistent data in which the scatter in the data is too large to be consistent with the stated uncertainties. Unlike the blackbody case, in which the cause was known, in many circumstances there is no explanation for the discrepancy. In any precision measurement there are almost always unknown things occurring, but fortunately they are generally too small to be of concern. But this is not always true. There is no generally accepted way of handling large inconsistencies and in many cases uncertainties are simply increased so that they are consistent with the observed scatter in the data.

Fortunately this is not a problem with the current fountain PFS data. Using all fountain data reported to the BIPM from early 2008 to the present a Birge ratio ranging from 0.9 to 1.1 over time is obtained indicating that the data is generally consistent with the stated total uncertainties. The relative microwave lensing bias is not large enough at this time to have a significant impact on the Birge ratio. However, if a large enough bias is not uniformly corrected among the PFS, the situation could arise where the PFS community might again have to consider arbitrarily increasing uncertainties to make the scatter consistent with the uncertainties.

V. Summary

The PFS community is facing a situation where a potentially significant frequency bias is not being uniformly applied to all PFS. This paper does not address the validity of the bias, but the more general question of whether the WGPSFS should develop some recommended guidelines for addressing this type of situation. If it is decided that guidelines should be developed, then the WGPSFS will have the task to implement them. It will have to determine how large a bias needs be before it is recommended that it be evaluated in all PFS, and also what criteria must be met regarding the validity of the bias. It is hoped that this paper will generate a broad discussion among the frequency standard community that can be used by the WGPSFS to shape its recommendations to the CCTF.

Everything said here also applies to secondary frequency standards (microwave and optical) and eventually to optical PFS when or if there is a redefinition of the second.

Acknowledgment

The authors thank Neil Ashby and Mike Lombardi for useful comments.

References

Transmission of a Frequency Channel Through a Long-Haul Optical Fiber Communications Link

Curtis R. Menyuk
Computer Science and Electrical Engineering Dept.
University of Maryland Baltimore County
Baltimore, MD 21250, USA
menyuk@umbc.edu

Abstract—Finding ways to communicate precise frequency and time through an optical fiber network has become a critical issue. In dark optical fibers, Rayleigh scattering may impose a limit on duplex transport that can be reduced by appropriately modulating the signal. New limits appear in communication networks in which neighboring communication channels can impair the performance of a frequency communication channel. We review the physical phenomena that can potentially impair a frequency communication channel. These include linear impairments — amplified spontaneous emission (ASE) noise, chromatic dispersion, and polarization mode dispersion. These also include nonlinear scattering impairments — Rayleigh scattering, Brillouin scattering, and Raman scattering. Finally, these include impairments due to the Kerr nonlinearity — self-phase modulation, cross-phase modulation, and four-wave mixing. We show that ASE noise imposes a lower limit on the frequency signal of ~1 nW, while both Brillouin scattering and self-phase modulation impose an upper limit of ~1 mW. Finally, we heuristically examine the effect of cross-phase modulation and show that it can lead to a fractional frequency uncertainty ~6.5×10^{-16} after 1 s.

Keywords—optical fibers, frequency signal transmission, impairments, Rayleigh scattering, cross-phase modulation

I. INTRODUCTION

The invention of technology for locking frequency combs led to a breakthrough in our ability to measure precise optical frequencies [1], which has in turn led to a steady improvement in optical frequency sources. The demonstrated residual uncertainty in frequency combs is better than 1 part in 10^{19} [2], while the residual uncertainty in Al⁺ clocks is better than 1 part in 10^{17} [3]. This remarkable improvement in frequency sources has opened up a host of potential new applications and appears likely to lead to a redefinition of the second [4].

More recently, there has been great progress in the transfer of frequencies over long distances. Starting from early experiments at the National Institute of Standards and Technology (NIST) [5] and progressing to more recent work in Germany, an 1840-km two-way link has been demonstrated that has a modified Allan deviation of better than 10^{-18} with 100 s of averaging [6]. However, this sytem requires dark fiber and special Brillouin amplifiers. Lopez et al. [7] demonstrated two-way transmission through a commercial system in which they bypassed the commercial amplifiers and achieved a fractional frequency instability of 5×10^{-15} after 1 s in 10 Hz. Any high-precision frequency transfer system must do two-way transfer to compensate for environmental changes in the effective fiber length. The NIST work demonstrated that duplex transport through two neighboring fibers led to the Allan deviation saturating at a high level after 1 s of averaging [5]. For this reason, later work used two-way transport through a single optical fiber [6], [7]. Two-way transport through a single fiber is incompatible with commercial optical fiber data transmission systems since these systems use isolators at the amplifiers. The reason for the failure of the Allan deviation to continue averaging down after one second has still not been completely determined. However, there is strong evidence that it is due to Rayleigh scattering [8]–[10], and recent work indicates that it is possible to substantially reduce its impact by modulating the frequency signal before it is transmitted [11].

A frequency communication channel that is transmitted in an optical fiber as part of an optical fiber communication system will be subject to a variety of impairments. These include linear impairments — amplified spontaneous emission noise (ASE) from the amplifiers, chromatic dispersion, and polarization mode dispersion. These include nonlinear impairments due to light scattering — Rayleigh scattering, Brillouin scattering, and Raman scattering. Finally, these include impairments due to the Kerr nonlinearity [12], [13]; these are self-phase modulation, cross-phase modulation, and four-wave mixing.

We will demonstrate that the dispersive impairments have little influence on a frequency communication channel because they depend on the bandwidth of the channel, and the channel would occupy a narrow bandwidth. Similarly, Raman scattering is negligible because it only becomes significant for a broadband channel. Brillouin scattering sets an upper limit on the power that can be transmitted in the frequency channel, which is on the order of 1–2 mW [13, p. 61, p. 325]. Since the data communication channels in an optical fiber communication system typically have powers on the order of milliwatts, this power limit is much higher than would be appropriate for a frequency channel. Rayleigh scattering is a narrowband effect; in communication systems, it leads to linear attenuation [13, p. 59].

Turning to the Kerr effect, self-phase modulation also sets a limit on the optical power in the frequency channel, which we would expect to be on the order of milliwatts. Four-wave mixing can only be important when it is phase-matched and that can be avoided by operating the frequency channel at wavelengths that are removed by several times its bandwidth.

This work has been supported by Raytheon Corporation and the US Army Research Laboratory.
from the zero dispersion wavelength. That only leaves cross-phase modulation (XPM), which is the most important process that impairs the frequency channel through an interaction with neighboring data channels. XPM will effectively act as a source of random, multiplicative noise.

In the remainder of this paper, we first briefly review the work on Rayleigh scattering in [8]–[11]. We next review the physical impairments that appear in an optical fiber communication system and justify our statements that all the linear impairments and most of the nonlinear impairments can be neglected. We then discuss in more detail the impact of cross-phase modulation on a frequency channel.

II. RAYLEIGH SCATTERING AND ITS IMPACT

Here, we briefly review work that has been done to elucidate the impact of Rayleigh scattering on transmission of a narrow-band frequency signal through an optical fiber.

Adles et al. [14] demonstrated that Rayleigh scattering is responsible for limiting the $Q$ that can be obtained in optoelectronic oscillators. They found that there was no advantage to increasing the fiber length in these devices beyond 6 km.

Shortly thereafter, Okusaga et al. [8] demonstrated that the Rayleigh gain as a function of frequency, $G_R(\omega)$, defined as the ratio of the backscattered optical-intensity noise spectrum to the input optical-intensity noise spectrum, is consistent with an analytical expression given by Boyd [15, p. 464] for stimulated Rayleigh scattering,

$$G_R(\omega) = A_R \left[ \frac{4\omega / \Gamma_R}{1 + (2\omega / \Gamma_R)^2} \right], \quad (1)$$

where $A_R$ is a constant and $\Gamma_R$ is the Rayleigh bandwidth. This result is consistent with replacing $\lambda / 4\pi$, where $\lambda$ is the light wavelength with the core radius $a$, which lowers the frequency at which the spectrum peaks by a factor of 2000, which in turn is consistent with assuming that the process is driven by the transverse gradient of the light intensity, either electrostrictively or electroabsorptively. In Fig. 1, we show the input and backscattered noise spectra, and we show the Rayleigh gain for a 6 km length of fiber.

Subsequent work shed doubt on this hypothesis. If the process is stimulated, we would expect the process to be dominated by either electroabsorption or electrostriction. Through heterodyne measurements, which allowed us to distinguish the upshifted from the downshifted components, Okusaga et al. [9] verified that the electroabsorptive and electrostrictive components had the same magnitude, which is consistent with a spontaneous process. We show these spectra in Fig. 2. Additionally, the backscattered intensity grows linearly with both fiber length and with intensity, which is once again the expected behavior in a spontaneous process. Thus, it was concluded that the process is spontaneous [9].

Most recently, using low-noise light sources, it was found that there is a frequency range in which the phase noise grows superlinearly with length up to 20 km and then grows linearly [10]. At the time that this is being written, it appears most likely that this superlinear growth is due to a combination of the phase noise in the optical source and a delay of the backscattered signal [16], [17], which is reflected from the frozen-in density fluctuations that are also responsible for attenuation of light in communication fibers [13, p. 59].

Regardless of its origin, the work to date has made it apparent that the scattered spectrum above 500 Hz can be substantially reduced by appropriately modulating the input
Techniques, rather than a single modulation frequency. More work is being done to suppress the spurs using broadband modulation frequencies between 1 kHz and 10 kHz. The best suppression is observed with large modulation depths and with modulation depth is 10 MHz. The best suppression is harmonics. In Fig. 3, the modulation frequency is 10 kHz, and in the region below the modulation frequency and its harmonics. The power in the phase modulated spectrum decreases away from the frequency channel as long as the frequency signal is stronger than around 10 nW. ASE noise is negligible due to the narrow bandwidth of the frequency channel. The noise power due to 10 amplifiers equals 1 nW. The ASE noise is negligible due to the narrow bandwidth of the frequency channel as long as the frequency signal is stronger than around 10 nW.

### A. Linear Impairments

**ASE Noise:** The noise power that is added by a single amplifier is given by

$$S_{\text{noise}} = (G-1)n_{\text{sp}}h f B,$$

where $G$ is the amplifier gain, $n_{\text{sp}}$ is the spontaneous emission factor, $h$ is Planck’s constant, $f$ is the signal frequency, and $B$ is the bandwidth. As an example, we may consider an optical fiber transmission system that is 800-km long with amplifiers that are spaced 80 km apart, with an operating wavelength of 1.5 $\mu$m, and a loss of 0.2 dB/km. These parameters imply a loss of 16 dB over 80 km, so that $G-1 = 39$. We then find that the noise power due to 10 amplifiers equals 1 nW. The ASE noise is negligible due to the narrow bandwidth of the frequency channel as long as the frequency signal is stronger than around 10 nW.

**Chromatic Dispersion:** In a communication system, chromatic dispersion leads to spread of a signal outside its bit slot. If we assume that chromatic dispersion leads to slippage of the zero crossings in a modulated signal, then we can estimate the slippage of the highest frequencies relative to the lowest. The time spread due to dispersion is given by $\tau_{\text{CD}} = \beta_2 BL$ [13, p. 40], where $\beta_2 = 17$ ps/nm-km in standard single-mode fiber (SMF) and $L$ is the propagation length. Taking a bandwidth of 10 MHz, we find that $\tau_{\text{CD}} = 1$ ps at 800 km, which is negligible relative to the modulation time $\tau_{\text{mod}} = 100$ ns.

**Polarization Mode Dispersion:** This effect is due to the group velocity difference between two orthogonal polarizations at the same frequency. There will be two orthogonal polarizations, called principal states, at which this difference is a maximum. Because the birefringence is randomly varying, it leads to a random walk in the time spread. The standard deviation of this spread, $\tau_{\text{PMD}}$, is given by

$$\tau_{\text{PMD}} = \sqrt{n \Delta n / c} \sqrt{BL},$$

where $c$ is the speed of light, $h$ is the fiber correlation length, and $L$ is the propagation length. Writing $\tau_{\text{PMD}} = D_p L^{1/2}$, we find that typical values of $D_p$ are 0.1–1.0 ps/km$^{1/2}$. Taking $D_p =$...
1.0 ps/km$^{1/2}$ and $L = 800$ km, we find $\tau_{\text{PMD}} = 29$ ps, which is once again much smaller than the modulation time.

B. Nonlinear Scattering Impairments

Rayleigh Scattering: When optical fibers are fabricated, the density fluctuations in their molten state are frozen into the glass. These frozen-in fluctuations lead to loss, which is on the order of 0.12–0.17 dB/km at 1.5 $\mu$m and is the principal source of attenuation in optical fibers. Thermal fluctuations are a separate source of Rayleigh scattering and are universally present in any medium, including glasses [15]. However, the bandwidth of the Rayleigh scattering is inversely proportional to the life-time of the fluctuations. Since the frozen-in fluctuations have a nearly infinite lifetime, their bandwidth is quite narrow. Hence, they can have a strong impact at low frequencies close to the carrier. At the present time, these frozen-in fluctuations appear to be the source of the phenomena that we have described in Sec. II [16], [17].

Brillouin Scattering: Brillouin scattering is a process in which light couples to acoustic waves and produces light at a lower frequency. One can also describe this process as one in which an incoming photon decays into a lower-frequency photon and a phonon [13, pp. 60–62]. In order for this process to be phase-matched, it must be the case that $\omega_{\text{pump}} = \omega_{\text{Stokes}} + \omega_{\text{acoustic}}$ and $\beta_{\text{pump}} = \beta_{\text{Stokes}} + \beta_{\text{acoustic}}$, where these quantities refer respectively to the frequencies and wavenumbers of the pump, Stokes, and acoustic waves. Due to the low velocity of sound relative to light, the process is only phase-matched in the backward direction. The stimulated process grows from thermal noise and the rate of growth is proportional to the input light intensity. If the growth rate exceeds the loss rate due to linear attenuation, then the scattered light will grow exponentially and will ultimately destroy the communication signal. This effect sets a maximum power $P_{\text{threshold}}$ that can be transmitted within the gain bandwidth of the Brillouin process. In standard optical fibers, this value is around 1 mW [13, p. 61]. This power is a serious limitation for data signals in optical fibers; however, the gain bandwidth of the process, which is approximately 100 MHz, is narrow compared to a typical data channel, which is on the order of 10 GHz in a long-haul communication system. As a result the power that is transmitted in a data channel can be higher than 1 mW. By contrast, the bandwidth of a frequency channel will fit entirely into the Brillouin bandwidth, so that 1 mW — its exact value depends on the system details — is a hard upper limit. However, as long as the power in the frequency channel is held well below this limit, the Brillouin effect will not cause trouble. Since the lower limit is 1 nW, this constraint is not difficult to satisfy.

Raman Scattering: This process is a broadband process (~20 THz) whose power threshold is 500 mW, which is far above the practical limit in communication systems.

C. Kerr Effect Impairments

The Kerr effect leads to a nonlinear change in the refractive index of the optical fiber. We may write the complex envelope of the field intensity of a communication signal as $u(z,t)$ where $z$ is distance along the fiber and $t$ is retarded time. The complex envelope $u(z,t)$ is proportional to the complex electric field. If it is normalized so that $|u(z,t)|^2$ equals the power, then we find that $u(z,t)$ is governed by the nonlinear Schrödinger equation [12, p. 44],

$$\frac{\partial u}{\partial z} = \frac{\alpha}{2} u - i \frac{\beta_2}{2} \frac{\partial^2 u}{\partial t^2} + i \gamma |u|^2 u,$$

where $\alpha$ is the attenuation coefficient, $\beta_2$ is the chromatic dispersion coefficient, and $\gamma$ is the Kerr coefficient. For SMF-28 fiber, we may use the value $\gamma = 1.3 \times 10^{-3}$ W$^{-1}$·m$^{-1}$ [12, p. 45].

We are assuming that a narrowband frequency channel will be placed in between two wavelength-division-multiplexed (WDM) data channels, as we show schematically in Fig. 5. The standard ITU grid for an optical fiber communication link sets the WDM channels at multiples of 12.5 GHz apart [19]. If the basic information transmission rate is 9.953 Gbaud/s, as is the case for a WAN 10-G ethernet or an OC192-STM64 channel [20], then there is spectral space for 25.6% of overhead that can be used for physical layer overhead and error correction. The amount of free spectrum between two WDM channels will depend on the transmission protocol. In the case of the interface that the ITU recommends for optical transport networks, every frame of 4080 bytes contains 256 bytes for forward error correction (FEC) and 14 bytes of additional overhead for a total of 2790 overhead bytes, which constitutes 6.6% of the total transmission. In this case, each channel would occupy 10.6 GHz, leaving 1.9 GHz free between channels [21]. The amount of free spectrum is expected to be smaller in some systems, for example, in submarine systems that use super-FEC [22]. So, we have to examine the impact of neighboring channels on a frequency channel as a function of their separation.

The total transmitted signal $u(z,t)$ can be divided into the data signal $u_d(z,t)$ and the frequency signal $u_f(z,t)$, so that $u(z,t) = u_d(z,t) + u_f(z,t)$. The power in the frequency channel should be sufficiently small so that self-phase modulation is negligible. We will calculate the limit that this requirement imposes on the power in the frequency channel when we discuss self-phase modulation. Since dispersive scale lengths are small compared to nonlinear scale lengths in modern-day optical fiber communications systems, it follows that the nonlinear effect of the frequency channel on the data channels is negligible. In this case, the evolution of the data signal is given by

$$\frac{\partial u_d}{\partial z} = \frac{\alpha}{2} u_d - i \frac{\beta_2}{2} \frac{\partial^2 u_d}{\partial t^2} + i \gamma |u_d|^2 u_d,$$

in which the effect of the frequency channel is ignored. The evolution of the frequency signal is then given by
\[ \frac{\partial u_f}{\partial z} = -\frac{\alpha}{2} u_f + i\gamma |u_f|^2 u_f + 2i\gamma |u_d|^2 u_f + i\nu u_d^* u_f. \] (7)

The three nonlinear terms on the right-hand side of (7) correspond respectively to self-phase modulation, cross-phase modulation, and four-wave mixing. We are neglecting chromatic dispersion in this equation, in keeping with our earlier calculation that demonstrated that it was negligible. However, it should be noted that we are effectively assuming a carrier frequency at the central frequency of the frequency channel. Dispersion appears implicitly in (7) since the WDM channels have a different group velocity from the frequency channel and will pass through it.

We now examine in more detail the effects of self-phase modulation (SPM), four-wave mixing (FWM), and cross-phase modulation.

**Self-Phase Modulation:** This effect is governed by

\[ \frac{\partial u_f}{\partial z} = -\frac{\alpha}{2} u_f + i\gamma |u_f|^2 u_f. \] (8)

We must operate at powers that are sufficiently low to make this nonlinearity negligible in order to avoid nonlinear distortion. We note that nonlinearity will only be effective over a limited length between the amplifiers because of linear attenuation [11, p. 98]. The effective nonlinear scale length is given by

\[ L_{\text{eff}} = \frac{1}{|\alpha|} \left[ 1 - \exp(-\alpha L_{\text{amp}}) \right], \] (9)

where \( L_{\text{amp}} \) is the amplifier spacing. We have \( \alpha = 0.2 \) dB/km, which translates to \( L_{\text{eff}} = 20 \) km when \( L_{\text{amp}} = 80 \) km. The nonlinear phase shift due to self-phase modulation is given by \( u_i = \gamma P_f N L_{\text{eff}}, \) where \( P_f \) is the power in the frequency channel and \( N \) is the number of amplifiers in the link. We then find that the power at which nonlinearity leads to a 1 radian change in the phase is given by \( P_{\text{nl}} = 1/(\gamma N L_{\text{eff}}). \) For our example, in which we set \( N = 10, \) we find that \( P_{\text{nl}} = 3.8 \) mW. In practice, it is necessary to operate far below this limit to avoid problems with Brillouin scattering.

**Four-Wave Mixing:** This effect is governed by

\[ \frac{\partial u_f}{\partial z} = i\nu u_d^* u_f. \] (10)

For two given WDM channels with angular frequencies \( \omega_m \) and \( \omega_a \) and corresponding wavenumbers \( \beta(\omega_m) \) and \( \beta(\omega_a), \) the contribution from (10) will only be phase-matched if we can simultaneously satisfy the conditions \( \omega_m + \omega_a = 2\omega_f \) and \( \beta(\omega_m) + \beta(\omega_a) = 2\beta(\omega_f), \) which is not possible except at the zero dispersion wavelength of the optical fiber. Placing the central frequency of the wavelength channel greater than about 5 times its bandwidth from the zero dispersion wavelength should be sufficient to avoid any difficulties.

**Cross-Phase Modulation:** This effect is governed by

\[ \frac{\partial u_f}{\partial z} = -\frac{\alpha}{2} u_f + 2i\gamma |u_d|^2 u_f. \] (11)

This term can be expected to dominate the distortion of a frequency channel in a real communication network. In the next section, we discuss a heuristic model for estimating the impact of the WDM data channels on the frequency channel.

IV. LIMITATIONS IMPOSED BY CROSS-PHASE MODULATION

From (11), we see that once the evolution of the data channels is given, the evolution of the frequency channel becomes linear. Since a WDM channel moves through the frequency channel at a rate that is given by the group velocity difference of the two channels, the effect of each WDM channel will appear on the frequency signal as a source of multiplicative noise.

We must take into account the following physical parameters for each WDM channel, \( u_m(z,t). \) The effect of the WDM channels will be additive.

1. The power variance of the \( m \)-th WDM channel. We note that only the variance will affect the frequency channel. The mean only affects the overall phase and does not lead to distortion.

2. The correlation function \( R_m(z-z'). \) This function gives the fiber length over which the effect on channel \( m \) at a single point in time of the frequency channel can be considered constant.

3. The time correlation function \( R_f(z, t-t') \) in the frequency channel as a function of distance along the fiber. Time and phase jitter of the frequency signal will only effectively occur over uncorrelated intervals.

We have considered a simple on-off-keyed (OOK) system that is not pre-dispersed. We assume that the frequency channel is primarily affected by its two neighbors, and the effect of each of the neighbors will be the same and will double the variance that is due to a single neighboring channel. Assuming that the data channel has a 10 GHz bandwidth, we find that the dispersive scale length over which it is expected to significantly disperse in the time domain is 80 km. Since we are assuming that the amplifiers are spaced 80 km apart and the effective length over which the nonlinearity interact is only 20 km, we may assume that the dispersion does not affect \( |\mu_d|^2 \) in the interval between each amplifier and then randomizes \( |\mu_d|^2 \) as the data signal propagates from one amplifier to the next, so that the effect of each 80-km interval on the phase of the frequency variance can be added. Starting from the expression

\[ \frac{\partial \phi}{\partial z} = 2\gamma |u_d|^2, \] (12)

where \( \phi \) is the phase in the frequency channel, we find that the change in the \( n \)-th interval is given by

\[ \Delta \phi(n,t) = 2\gamma P_d(n,t)L_{\text{eff}}, \] (13)

where \( P_d(n,t) = |u_d(n,t)|^2 \) is assumed to be constant at any point in time \( t \) in the \( n \)-th interval. We find for the variance,

\[ \sigma_d^2(n) = \left\langle (\Delta \phi(n,t))^2 \right\rangle = 8\gamma^2 L_{\text{eff}}^2 \sigma_d^2(n) \]

where \( \sigma_d^2(n) = \left\langle (P_d(n,t))^2 \right\rangle - \langle P_d(n,t) \rangle^2 \).  

\[ = 8\gamma^2 L_{\text{eff}}^2 \left\langle (P_d(n,t))^2 \right\rangle - \langle P_d(n,t) \rangle^2. \] (14)
An OOK system will have either a power of 2\(P_{av}\) or 0 in each bit slot in the first 80-km interval, where \(P_{av}\) is the average power. We then find
\[
\sigma_p^2(l) = P_{av}^2. 
\] (15)

In the second interval, each bit will have spread to twice its original duration and the total power in each bit slot will be a mix of two bits. More generally, the power in each bit slot in the \(n\)-th interval will be a mix of \(n\) bits, and the final result for the sum over \(N\) intervals — if \(N\) is not too large — is
\[
\sigma_p^2(N) = 8\gamma^2 L_{eff} P_{av} \sum_{n=1}^{N} (1/n). 
\] (16)

For the example that we are considering here, we have \(L_{eff} = 20\) km and \(N = 10\), corresponding to a propagation length of 800 km. If we consider a system with \(P_{av} = 1\) mW, we find that the variance of \(\sigma_p\), given by (16), is equal to 0.016.

This phase variance will translate into jitter of the optical zero crossings and ultimately into frequency jitter. However, time points in the frequency signal that are nearby will experience the same jitter since the WDM channel affects them in the same way. To determine the impact on the measured frequency, we must determine the time separation that is required for time points in the frequency signal to experience uncorrelated phase jitter. As previously noted, the dispersive scale is large enough so that any time point in the frequency signal is interacting with only one bit slot of the neighboring WDM channels during each amplifier interval. Thus, it is reasonable to take 100 ps, which is the time duration of a bit slot as this correlation time.

\[\tau_{corr} = 100\text{ ps} \]

As previously noted, the dispersive scale is large enough so that any time point in the frequency signal is interacting with only one bit slot of the neighboring WDM channels during each amplifier interval. Thus, it is reasonable to take 100 ps, which is the time duration of a bit slot as this correlation time. The frequency variance is given by the phase variance divided by the square of \(\tau_{corr}\). Hence, we conclude \(\sigma_{\phi} = 1.3\) GHz. That translates into a fractional uncertainty of \(6.5\times10^{-6}\). If we assume that this fractional uncertainty averages down proportional to the time interval, we find a fractional uncertainty of \(6.5\times10^{-16}\) after 1 s and a fractional uncertainty of \(6.5\times10^{-18}\) after 100 s.

Of course, a far more careful study must be made to verify these conclusions.

**ACKNOWLEDGMENT**

It is a pleasure to acknowledge useful discussions with A. Kowalewicz and P Sykes. Joint work at ARL and the Technion was carried out with J. Cahill, G. Carter, M. Fleyer, M. Horowitz, E. Levy, O. Okusaga, and W. Zhou.

**REFERENCES**

Preliminary time transfer through optical fiber at NIM

LIANG Kun, ZHANG Aimin, YANG Zhiqiang, CHEN Weiliang, WANG Weibo
Division of Time and Frequency Metrology
National Institute of Metrology (NIM)
Beijing, China
liangk@nim.ac.cn

BAI Long, FU Guitao
Beijing Satellite Navigation Center (BSNC)
Beijing, China

Abstract—TWOTFT has been studied and implemented according to the similar principles to those of TWSTFT. We initially constructed the experiment system for TWOTFT and done some experiments such as time and frequency transfer through the laboratory optical fiber and the real optical fiber links including the real link with 109 km length at NIM. We can get the time stability of less than 6 ps/s and 0.9 ps/100s, and the standard uncertainty of less than 200 ps for time transfer.

Keywords—time transfer; two way; optical fiber

I. INTRODUCTION

Since 2012, we have finished building the new timekeeping system at the Changping campus of National Institute of Metrology (NIM) and changed the realization of UTC (NIM) to this campus from the Hepingli campus[1]. There is still one H-maser and two cesium clocks located at the Hepingli campus, based on which one time scale UTC (NIM) Hepingli is generated, and we provide most of the calibration and traceability services directly reference to the time scale at this campus. So we need one precise time and frequency transfer method to link our two campuses which are about 40 km away on road. GPS time transfer has been laid out on this baseline between our two campuses, however, with the development of time transfer using optical fiber, we may try to link our two campuses by the method with the lower uncertainty. As well, in Beijing, within 100 km distance, there are several important metrology institutes who have the requirements for precise time and frequency transfer referenced to NIM. So it make us for the development of time and frequency transfer through optical fiber.

TWSTFT (Two Way Satellite Time and Frequency Transfer) is a famous precise time transfer method to compare two time scales with a very long distance, together with GNSS (Global Navigation Satellite System) time and frequency transfer. BIPM (Bureau international des poids et mesures) evaluates the method with the uncertainty A of better than 0.5 ns. According to the similar principles, we could use the optical fiber instead of satellite and free space for time and frequency transfer and we call it two way optical fiber time and frequency transfer (TWOTFT) as BIPM did. PTB (Physikalisch-Technische Bundesanstalt) has finished the first successful implementation on the baseline between PTB and the Institut für Quantenoptik (IQ) at Leibniz Universität Hannover with the length of 73 km, and the uncertainty was below 100 ps[2], and O. Lopez, etc. also did the similar experiment in optical fiber network carrying the internet data with the length 540 km and acquired the uncertainty of 250 ps[3].

II. PRINCIPLES AND EXPERIMENT SET-UPS

In TWOTFT, we use the intermediate-frequency signal modems and the O/E converters and E/O converters for the experiments, and thus compared with TWSTFT we can omit some delays such as sagnac delay, satellite path delay through the satellite transponder, ionospheric delay and other delays due to the satellite and free space.

A. Principles of TWOTFT

The concept of TWOTFT is very similar to that of TWSTFT. In TWSTFT, the difference of the time and frequency references between the two laboratories is determined as follows[4].

\[
TS(i) - TS(2) = +0.5TI(1) - 0.5TI(2) + 0.5[SPT(1) - SPT(2)] - 0.5[SCD(1) - SCU(1)] + 0.5[SCD(2) - SCU(2)] + 0.5[SPU(1) - SPD(1)] - 0.5[SPU(2) - SPD(2)] + 0.5[TX(1) - RX(1)] - 0.5[TX(2) - RX(2)]
\]

(1)

Where,
TS(i) is Local time and frequency reference
TI(i) is Time interval reading
SPT(1) is Satellite path delay through the transponder
SCD(1) is Sagnac delay for a signal propagating from the GEO satellite to station i
SCU(i) is Sagnac delay for a signal propagating from station i to the GEO satellite
SPD(1) is Signal path downlink delay
SPU(i) is Signal path uplink delay
TX(i) is Signal delay in the transmit path of the TWSTFT station i
RX(i) is Signal delay in the receive path of TWSTFT station i

In TWOTFT, We just need to consider the path delay through the optical fiber link. So in TWOTFT, we can get the time transfer result as follows.

Supported by Chinese NSFC program 11303024.

978-1-4799-8866-2/15/$31.00 ©2015 IEEE 742
\[ TS(1) - TS(2) = 0.5T_i(1) - 0.5T_i(2) + 0.5SP(1) - 0.5SP(2) + 0.5[TX(i) - RX(i)] - 0.5[TX(2) - RX(2)] \] (2)

Where,
SP(i) is Signal path delay
TX(i) is Signal delay in the transmit path of the TWOTFT station \( i \)
RX(i) is Signal delay in the receive path of TWOTFT station \( i \)

**B. Experiment Set-ups**

In TWOTFT, as shown in Fig. 1, we use one local modem referenced to the local time scale to generate the microwave signal modulated by the local time information, after the conversion of E/O, the signal is transferred through the optical fiber instead of satellite and free space, and then at remote site after the conversion of O/E, the other modem is used to demodulate the local time information referenced to the remote time scale from the local microwave signal, at the same time we do vice versa, and at last combining the two time information demodulated by the two modems separately located at the two sites will give access to compare the two time and frequency references.

**C. Time Link Calibration**

The calibration should be implemented first before time transfer. For TWOTFT, the total delay of the time transfer equipment should include three parts of delays that must be calibrated, and these are the external reference delay, the internal reference delay and the internal delay of the equipment.

The external reference delay and the internal reference delay could be measured using the time interval counter, such as SRS SR620. In addition, for the TWOTFT link, we also need to involve the difference between the internal delays of the local and remote equipments, that is to say, we do not have to know the exact delay for each equipment located in each site. Thus we might do the common clock difference experiment to cancel out the effect of the time and frequency references for getting the difference of the internal delays between the two equipments with the local and remote equipments side by side before we transfer the remote equipment to the remote site.

**III. NUMERICAL RESULTS**

We have two rolls of laboratory optical fiber with the length of about 50 km, and two real optical fiber links available including the link with the length of about 55 km between Hepingli campus and Changping campus and the link between Changping campus and Beijing Satellite Navigation Center (BSNC) with the length of about 32 km.

**A. CCD with Laboratory Optical Fiber Roll**

In terms of this kind of principles, some experiments have been implemented at NIM. First, to test the performance of the time transfer system itself, we compare the same time and frequency reference, which is the common clock difference (CCD) experiment. The link length for time transfer has been increased step by step. The first step is about 50 km, then 100 km.

From Fig. 2, we can find that there is the lowest time stability in 100 seconds possibly because of the timing parameter of the modem time interval counter of the time constant of the phase lock loop of the modem that is repeated in the experiments on the real optical links as demonstrated in Fig. 6 and 8. In Fig. 4, we see that time stability is worse and not so regular with the increased attenuation derived from the longer length fiber and the unknown reason coming from the optical fiber itself. From Fig. 3 and 5, it shows that the measurements could be dominated by the white phase noise in the considerable extent and it should be clear if there is more data in the experiments hereafter.

**Fig. 1.** TWOTFT principle

**Fig. 2.** TDEV of CCD on the lab optical fiber baseline of 50 km
Furthermore, more specifically, we use the real optical fiber link outside the laboratory including between Hepingli campus and Changping campus ring-connected with the length of about 109 km and between Changping campus and BSNC ring-connected with the length of about 63 km to implement the transfer with the real optical fiber link partly underground and partly over ground.

Next step, we will implement the experiment in the 63 km optical fiber with the two fibers between our Changping campus and BSNC campuses ring-connected.
From the experiments using laboratory fiber and real optical fiber links, we can get the time stability of less than 6 ps/s and 0.9 ps/100s as follows in Fig. 2-9. Besides, it is showed that the frequency stability better than 2e-17 at one day could be achieved for TWOTFT link, although frequency transfer is not the main concern in TWOTFT, anyway, here on the real optical fiber links there is much complex constitution of noise elements especially after sampling interval of one thousand seconds than in the laboratory optical fiber as a result of the outside environment influence and the connection of the real optical fiber links.

In the results of four days shown in Fig. 10, we can see a certain sinusoid fluctuation cycle of half a day with the peak to peak about 100 ps that cannot be understood at present. Moreover, the changes of temperature and humidity of the laboratory have been surveyed, anyway there is no strong proof for the relation for this cycle fluctuation as show in Fig. 11.

C. Time Transfer

To verify the effect of time transfer by TWOTFT based on the real optical fiber link, we finished the time link calibration to give the time transfer results. Moreover, we compare the time transfer results of TWOTFT in the direct connection that we can think as the 0 m baseline between UTC(NIM) and UTC(NIM1) that is the backup of UTC(NIM) with GPS time transfer link and also time difference was measured directly by SR620. GPS PPP link results have been removed by some offset for the better view on account of the deficiency of calibration of GPS PPP link. From Fig. 12, we can see that the results by TWOTFT could be more smooth and precise than GPS PPP and the results of TWOTFT can agree with those by SR620 directly within the range 200 ps at most on the accuracy.

We finally implemented the time transfer experiment in the baseline between Hepingli campus and Changping campus at the length of 109 km. We can get the consistency among TWOTFT, GPS and SR620 time transfer links as Fig. 13 shows. From Fig. 13, we can see that the average difference between TWOTFT measurements and the measurements by SR620 is 0.57 ns for one day, which is within the standard uncertainty of SR620 measurement. There is no absolute measurements for PPP results because none of the related delays have been corrected. In any case, compared to PPP results for precision, we can conclude that TWOTFT is much smoother that mean there is much less jitter and much stable measurements in the long baseline.
normally divided into two types. One is the uncertainty type A that is almost from the stochastic elements and the other is the uncertainty type B that is almost from the systematic and empirical elements.

From the formula of the TWOTFT, we can see that there are several uncertainty sources, such as, stability of real measurements, the calibration of the difference of the equipments and the path delays, measurements of the reference delays. The uncertainty budget is as follows in Table I.

<table>
<thead>
<tr>
<th>Uncertainty Source</th>
<th>Type</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stability of measurements</td>
<td>A</td>
<td>17 ps</td>
</tr>
<tr>
<td>External reference delay measurement for site A</td>
<td>B</td>
<td>10 ps</td>
</tr>
<tr>
<td>External reference delay measurement for site B</td>
<td>B</td>
<td>10 ps</td>
</tr>
<tr>
<td>Calibration of the difference of the equipments and the path delays</td>
<td>B</td>
<td>150 ps</td>
</tr>
<tr>
<td>Internal reference delay measurement for site A</td>
<td>B</td>
<td>24 ps</td>
</tr>
<tr>
<td>Internal reference delay measurement for site B</td>
<td>B</td>
<td>21 ps</td>
</tr>
<tr>
<td>Total standard combined uncertainty</td>
<td></td>
<td>155 ps</td>
</tr>
</tbody>
</table>

In evaluation, we take the standard deviation of the CCD experiments as the uncertainty of the measurement stability as it is normally used in the evaluation of other time transfer methods. For the uncertainty from the external reference delay measurements, the standard deviation of the measurements by the external time interval counter(TIC, type SRS SR620) has been used. And from the measurements of internal reference delays by the same TIC other than the internal TIC of the modem owing to the less stability of the internal TIC than the external TIC, we found that this delay is not so stationary and a little bit noise with the level of about tens of picoseconds, so that it should be involved in the uncertainty budget.

The path delay, in TWSTFT, has been deemed as being cancelled out because the significant hypothesis that the back and forth paths are equal. Whereas TWOTFT is expected to be more precise than TWSTFT, and the difference between the back and forth paths might be considered. We investigated the differences of the internal delays of the equipments among the different optical paths by doing CCD experiments in the different optical links as shown in Fig. 14. We can see that if we change the different optical links, the CCD results, i.e., the calibration of the equipments and the path delays, may be various as described in Fig.14. The uncertainty reflecting this change should be considered. Here, supposed we use 100 km laboratory optical fiber to do CCD for getting the calibration of the equipment and the path delays, we can get the difference of the calibration as 85 ps from that of real optical link between Hepingli campus and Changping campus at the length of 109 km as shown in Fig. 14, and if we use 50 km laboratory optical fiber to do CCD, we can get more difference from that of 109 km real optical link. So we conservatively estimate the uncertainty from this as 150 ps. Anyhow, this uncertainty would be analyzed and evaluated more thoroughly in the near future.

![CCD results](image)

Fig. 14. Change of calibration correction with optical link length

V. CONCLUSIONS AND PROSPECTS

In the baseline of the real optical fiber link with the length of up to 109 km outside the laboratory, we can get the time stability of less than 6 ps/s and 0.9 ps/100s, and the standard uncertainty of less than 200 ps for time transfer. Next, we would try to investigate the influence of cycle jitter of half day, the calibration of the difference of the equipments and the path delays and all the other possible uncertainty sources to improve the time transfer uncertainty and at the same time we will lay out the TWOTFT equipment separately in two remote sites for the real time transfer application.

ACKNOWLEDGMENT

The authors thank for the financial support of NSFC.

REFERENCES

Actively and Passively Compensated RF Frequency Disseminations on Branching Fiber Network

Bo Wang1, 2, Xi Zhu1, 3, Yu Bai1, 3, Chao Gao1, 2, and Lijun Wang1, 2, 3, 4

1Joint Institute for Measurement Science, Tsinghua University, Beijing 100084, China
2State Key Lab of Precision Measurement Technology and Instruments, Department of Precision Instruments, Tsinghua University, Beijing 100084, China
3Department of Physics, Tsinghua University, Beijing 100084, China
4National Institute of Metrology, Beijing 100013, China

Email: bo.wang@tsinghua.edu.cn

Abstract—We present two RF frequency dissemination methods (active and passive phase fluctuation compensation) for branching network. For the active one, the phase noise compensation function placed at the client site. One transmitting module hence can be linked with multiple client sites. For the passive one, without any phase control on the disseminated rf signals or usages of active feedback loop, the highly stable reference radio frequency signal can be delivered to several remote sites simultaneously and independently.

Keywords—frequency dissemination; phase fluctuation compensation

I. INTRODUCTION

Over the past decade, ultra-stable optical and RF frequency disseminations via fiber link have been demonstrated by many groups. To improve its accessibility, multi-access frequency dissemination methods have been proposed and demonstrated by several groups. How to apply these techniques to satisfy the frequency synchronization requirements of different large-scale Scientific and Engineering facilities is the next challenge to many us. Currently, the European FT-Neat consortium is constructing the time and frequency synchronization network through the telecommunication fiber network; the Beijing regional time and frequency synchronization network is under construction, 5 remote institutes (5 H-Masers) have been linked by the network; the square kilometer array telescope (SKA) requires reference frequency synchronization for thousands of dish antenna. These practical applications required a frequency dissemination methods suitable for the branching network.

In this paper, we demonstrate two RF frequency dissemination methods (active and passive phase fluctuation compensation) for branching network. For the active one, the phase noise compensation function placed at the client site. One transmitting module hence can be linked with multiple client sites. As a performance test, using two separate 50 km fiber spools, we recover the 100 MHz disseminated reference frequencies at two remote sites, separately. Relative frequency stabilities between two recovered frequency signals of 2.8×10⁻¹⁴/s and 2.5×10⁻¹⁷/10⁵s are obtained.[1] For the passive one, without any phase control on the disseminated rf signals or usages of active feedback loop, the highly stable reference radio frequency signal can be delivered to several remote sites simultaneously and independently. Relative frequency stability of 6×10⁻¹⁵/s and 7×10⁻¹⁷/10⁴s is obtained for 10km dissemination. Detailed experimental results will be shown during the conference.

II. ACTIVELY COMPENSATED RF FREQUENCY DISSEMINATION

For the frequency dissemination system used to construct a time and frequency synchronization network, there are some basic requirements: 1, the system can be used in the star-topological structure. For SKA, one central station (where H-maser is placed) should be linked with hundreds of dish telescopes; 2, using the dissemination system, the recovered frequency signals’ specification (stability and phase noise) should be same with that of the frequency reference at center station (for example, the H-maser). Currently, for most of the actively compensated frequency dissemination systems, the phase noise compensation function is placed at the transmitting site. This would generate extraordinary space requirement, and cause unnecessary complexities and difficulties for future expansion.

Considering these requirements, we design an actively compensated RF frequency dissemination system, and the phase noise compensation function is placed at the receiving site. Its detailed schematic diagram and the phase noise compensation principle have been reported in Ref [1]. Based on it, we have designed a prototype module. Figure.1 shows the photo of transmitting module (TX). The TX is designed to distribute 100MHz reference signal. Its function is very
simple-just modulating and broadcasting. The TX has eleven cards-1 main control card and 10 transmitting cards (only 2 cards are inserted in the figure). The main control card receives the 100 MHz reference frequency from H-Maser and outputs identical modulated laser signals. The RF modulated light signal is amplified and broadcasted via each transmitting cards.

Figure 1. The transmitting module of the actively compensated frequency dissemination system.

Figure 2 shows the photo of receiving module (RX). The RF modulated light signal is inserted to the FC-APC connector (yellow arrow). The phase fluctuation induced during fibre transmission is compensated via the lock-in loop in RX. There are 3 frequency output ports (blue arrows), and output recovered 100 MHz and 1 GHz frequency signals, respectively. We have measured the frequency dissemination stability on a 50 km fiber spool. Relative frequency stabilities between two recovered frequency signals of $2.8 \times 10^{-14}/s$ and $2.5 \times 10^{-17}/10^5s$ are obtained. [1] It can be used to disseminate the frequency signal of H-maser clock.

Fig. 2. The Receiving module of the actively compensated frequency dissemination system.

III. PASSIVELY COMPENSATED RF FREQUENCY DISSEMINATION

Figure 3 shows the schematic diagram of the passive frequency dissemination for branching network experiment. As a laboratorial demonstration, the branching network consists of one local site, three remote sites (C, D, E), and three fiber spools with lengths of 2km, 3km and 5 km, respectively. For the convenience of relative frequency stability measurement, the entire system is placed in the same lab. Detailed experiment results have been shown in Ref. [2]. Using the passive phase noise cancelation method, the highly stable reference frequency signal can be delivered to remote sites simultaneously and independently. For 10 km distance dissemination, relative frequency stability of $6 \times 10^{-15}/s$ and $7 \times 10^{-17}/10^4s$ are obtained. This simple and scalable scheme can be used to construct the frequency synchronization network of DSN and SKA.

Fig. 3. Schematic diagram of the fiber based radio frequency dissemination scheme for branching networks with the passive phase noise cancellation method.

This work was supported by the National Key Scientific Instrument and Equipment Development Project (No.2013YQ09094303) and the Beijing Higher Education Young Elite Teacher Project (No. YETP0088).


Abstract—Phase noise performance of oscillators using various forced oscillation using self-injection locked phase locked loop (SILPLL) configurations has been reported. Comparisons are made to identify the key parameters that have the most impact on phase noise.

Keywords—self injection locked phase locked loop; self-injection locking; self-phase locked loop; fiber optic link; residual phase noise

I. INTRODUCTION

Self-injection locked phase lock loop (SILPLL) utilizing fiber optic feedback is a novel technique for phase noise reduction. It is achieved by simultaneously combining SIL [1]-[2] and SPLL [3]-[4]. The authors have provided detailed analysis and successfully demonstrated the advantages of this technique with different oscillators at X-band [5]-[7]. In this paper, the authors propose an optimized fiber optic feedback configuration for lowest possible phase noise, and the experimental results are reported for the first time. A comparison of different SILPLL topologies are also provided in this paper to provide more insights into this technique, followed by discussions and prediction of further phase noise reduction.

A block diagram that depicts the working principle of SILPLL is shown in Fig. 1. A VCO is used as a seed source to drive the multiple functions in the system. A portion of the VCO output is sent to a Mach-Zehnder modulator (MZM) to modulate the light from a laser. The modulated lights is split into two; both pass through long fiber delays, and are amplified after being converted to electrical signal by photo-detectors (PDs). One of them is sent to a PCB board depicted as ‘Mixer+LPFA’ block in Fig. 1. The PCB board integrates a double balanced mixer as PLL phase detector and an Op-Amp as PLL low pass filter amplifier (LPFA). The phase of the delayed signal is compared against that of the VCO to generate an error signal for the control of the VCO frequency. The SILPLL circuitry is enclosed using a purple color box in Fig. 1. Another delayed signal is directly feedback to the VCO through an adjustable attenuator, completing the SIL function with various injected power ratio. The attenuator is used to control the injection strength expressed as $\rho = \sqrt{P_i/P_o}$ with $P_i$ being the injecting signal power and $P_o$ being the VCO power.

The organization of this paper as follows: Section II provides an overview of current SILPLL performance with various circuit topologies and describes the pros and cons of various forced oscillation topologies. Experimental performance of optimized SILPLL is provided in Section III. Finally, Section IV concludes the paper with discussion and prediction for further phase noise reduction using low residual phase noise amplifiers.

II. PERFORMANCE OF DIFFERENT SELF-ILPLL TOPOLOGIES

A. SILPLL with OEO

A standard OEO with 100 meter fiber delay is placed in the ‘VCO’ block in Fig. 1 to serve as VCO in SPLL portion. Phase noise of this OEO free running at 9.6 GHz is depicted as the black curve in Fig. 2. The measured phase noise is -69 dBc/Hz at 1 kHz offset, and the noise floor level is -140 dBc/Hz at far away offsets. This high noise floor level is directly related to the noise to signal ratio of the fiber optic link. The fiber optic link loss of 47 dB greatly attenuates the signal while laser RIN of -150 dB/Hz and shot noise related to the photo-detection process introduce excessive noise to the system, resulting in a high noise to signal ratio. SILPLL is employed to improve the phase noise performance of this OEO, but the delay configuration differs slightly from that of Fig. 1. One single fiber delay of 5 km is used, and the delayed signal is then coupled to both the SIL and SPLL function. In other words, SIL and SPLL share the same delay to reduce component cost and optical power loss. The phase noise of this OEO
employing SILPLL with 5km delay is depicted as the red curve in Fig. 2. A phase noise of -96 dBc/Hz at 1 kHz is achieved by employing SILPLL, which demonstrates the effectiveness of this technique by reducing phase noise by 27 dB. However, the high noise floor level of the OEO limits the overall phase noise performance, especially in the far away offset region. In order to reduce the noise floor level, a stable DRO is used in place of the OEO, which is discussed in section B.

C. SILPLL with DRO using Dual Drive MZM

Even though EAM-LD is attractive in many ways, it has a higher noise figure than an MZM based, where performance of LD and modulator could individually be improved. To achieve the least possible phase noise, MZM is used again to construct the DSILPLL for the same DRO. In this case, a dual drive MZM (DD-MZM) configuration is adopted. The benefit of DD-MZM is that the driving voltage can be reduced by half as opposed to a single drive MZM. As has been shown in [5], the absolute close-in to carrier phase noise of SILPLL system is determined by the noise to signal ratio after the PD. Amplifying the PD’s output ensures a proper power level for functional SILPLL, but it does not reduce the ratio. One viable method to minimize the noise to signal ratio is to increase the signal level at the PD by increasing the MZM drive RF signal. Since the output noise from the PD is usually dominated by laser relative intensity noise (RIN) and shot noise, amplifying the MZM driving signal has little effect on the output noise, and subsequently the ratio is reduced by a higher signal power. Therefore, a custom designed push-pull amplifier with single input and differential output is used to boost up the MZM driving signal.

Measured phase noise for DSILPLL employing DD-MZM is provided in Fig. 4. Once again, we can see that longer delay provides better phase noise reduction. The achieved phase noise for delay combination of 4 km and 8 km (blue color curve) is -105 dBc/Hz at 1 kHz offset and -127 dBc/Hz at 10 kHz offset. In the experiment with DD-MZM, the differential amplifier provided 5 dB gain in each path, therefore a minimum of 5 dB reduction in phase noise is expected due to an increased signal level after PD. From the phase noise comparison between single drive and dual drive shown in Fig. 4, it could be deduced that the performance of DD-MZM is better than single drive MZM by 3 dB at 10 kHz offset and 7 dB at 1 kHz offset, which agrees with our assertions of improved gain in dual drive versus single drive MZM.

B. SILPLL with DRO using Electro-Absorption Modulator

The DRO used in the experiment is from Synergy Microwave Corp. (DRO100). The free running phase noise of this DRO is shown as the black curve in Fig. 3, with a phase noise level of -83 dBc/Hz at 1 kHz offset and a noise floor level of -160 dBc/Hz. In the SILPLL configuration, a small form factor single package of EAM-LD is employed that integrates an electro-absorption modulator (EAM) and a laser diode (LD). The overall component cost and size are reduced, when the monolithically integrated EAM-LD replaces independent laser source and the MZM in the previous case. The use of EAM-LD also reduces the overall system power consumption. In addition, EAM is cost effective and is capable for monolithic integrated production. For the delay configuration, two different fiber lengths are used; the two delayed signals are converted to electrical signals independently by two PDs and then they are combined electrically before being sent to the SIL and SPLL portions. Once again, the dual delays are shared by SIL and SPLL and this configuration is termed as Dual loop SILPLL (DSILPLL). Phase noise performance for DSILPLL using EAM is shown in Fig. 3. As it is rendered in Fig. 3, a longer delay provides a better phase noise performance. The achieved phase noise is -100 dBc/Hz at 1 kHz offset for 5km+8km, which corresponds to a 17 dB reduction in phase noise over DRO.

III. SILPLL WITH OPTIMIZED FEEDBACK CONFIGURATION

In the previous DSILPLL experiments, the optical delays are shared by both SIL and SPLL functions. While it is desirable to use long delays in SPLL portion for better frequency discriminator sensitivity, the side-modes associated with the long delays are more closely spaced introducing higher spurious signal level. By using a shorter independent
delay in SIL portion, the spurious signal level could be reduced due to multi-loop operation. In addition, a dedicated SIL delay provides higher injection strength for an increased injection locking bandwidth, which results in a better phase noise reduction in far-away offset region. The block diagram for DSILPLL with independent SIL delay is very similar to Fig. 1 except that two fiber delays are used in the SPLL portion, and the delayed signals are combined optically before being converted to electrical signal.

The measured phase noise for DSILPLL with independent SIL delay using DD-MZM is shown in Fig. 5. Different delay topologies are used to explore the impact of the dedicated SIL delay. From the measured results, an independent SIL delay of 3km combined with DSPLL using 1 km and 5 km delays (blue color curve) provides the best phase noise of -127 dBc/Hz at 10 kHz offset, corresponding to a 14 dB reduction. For comparison, phase noise of DSILPLL with shared delays of 4 km and 8 km (magenta color curve) are also provided. It seems that independent SIL delay does not provide the additional phase noise reduction in far-away offset region. This is due to the fiber optic link noise floor being higher than the DRO noise floor, and a loop bandwidth of SILPLL being too large. A large loop bandwidth allows the noise from the fiber optic link passes into the VCO without being attenuated at far away offset frequencies. For the best phase noise results in both close-in and far away offsets, the attenuator in the SIL path must be carefully adjusted to yield a proper loop bandwidth that ensures significant noise reduction in close-in region while maintaining low noise level in far away offsets.

![Fig. 5 Phase noise comparison of OEO](image)

**IV. DISCUSSION AND CONCLUSION**

From the measured phase noise performance with various SILPLL configurations, it seems that there is a phase noise plateau about -125 dBc/Hz from 6 kHz to 30 kHz that keeps the phase noise from getting lower. Since the noise floor levels of both the MZM link and EAM link are lower than -125 dBc/Hz and the location of the plateau is in the close-in offset region, this plateau is most likely due to the amplifier flicker noise. The amplifiers used in all these experiments are GaAs based p-HEMT, which typically exhibits high flicker noise. On the other hand, SiGe based HBT has excellent flicker characteristic. If HBT amplifiers that exhibit a flicker constant of -120 dB/Hz² [8] are used to construct the SILPLL system, further phase noise reduction is predicted as shown in Fig. 6, with phase noise of -125 dBc/Hz and -140 dBc/Hz at 1 kHz and 10 kHz respectively.

![Fig. 6 Simulated phase noise of SILPLL using a lower residual noise for SiGe HBT amplifiers](image)

In conclusion, this paper has demonstrated that SILPLL using optical feedback is a novel technique that is an effective method to reduce phase noise in different oscillators at X-band. Due to the inherently broad bandwidth of optical components, this technique is capable of reducing phase noise and provides high purity signals at frequencies of millimeter wave and beyond.

**REFERENCES**


Abstract—We describe the development and frequency instability measurements of a highly miniaturized, buffer gas cooled, trapped-ion atomic clock. The clock utilizes the 12.6 GHz hyperfine transition of the $^{171}$Yb$^+$ ion. A custom-built 3 cm$^3$ vacuum package containing the ion trap is integrated with other key elements of the atomic frequency standard, including a photo multiplier tube, miniaturized laser sources at 369 nm and 935 nm, a local oscillator, and control electronics. With the clock physics package assembled on a 10 cm x 15 cm breadboard, the long-term fractional frequency instability was measured to be $6 \times 10^{-14}$ at 25 days of integration. Later, the clock physics package was further miniaturized, and the frequency instability was measured to be $2 \times 10^{-11}/\tau^{1/2}$ at integration times up to 10,000 s.

Keywords—atomic clock; trapped ions;

I. INTRODUCTION

Miniature atomic clocks are thought to have a broad range of potential applications including precision navigation, mobile high-speed communications, secure anti-jamming communications, and remote sensing. Vapor-cell based miniature atomic clocks, such as the CSAC (chip-scale atomic clock) [1, 2], have been developed and commercialized over the past decade. Miniature vapor cell clocks, however, suffer from long-term frequency drift mechanisms and therefore can only be used for short-term applications with periodic calibrations. For achieving improved long-term and short-term performance in miniature atomic clocks, ion-trapping technology has several advantages. Because of the robust ion-trapping mechanism (~ eV trap depth), trapped ions are well isolated from the environment and insensitive to vibrations and accelerations. The background gas pressure can be as high as $10^{-5}$ torr, and therefore no active pumps are required to maintain the vacuum. In fact, buffer-gas cooled trapped-ion systems usually use a few micro-torr of buffer gas, such as He or Ne, to maintain the temperature of the trapped ions around 1000 K. This temperature condition is sufficient for making a microwave atomic clock and does not require the complexity of optical cooling. There have been several high-performance microwave ion clocks reported using trapped Hg and Yb ions [3, 4]. Hence, we believe that microwave ion clocks are well suited for miniaturization. In the past few years, we have focused on developing various technologies for a miniature microwave Yb ion frequency standard, including: miniature, sealed ion-trap vacuum packages; miniature, low-power laser sources; low-power ion trapping [5]; a miniature Yb source [6]; integrated optics; and low-power microwave source and...
In this paper, we present two tests of the performance of a 3 cm³ vacuum package with it being used in two different clock implementations.

For the miniaturized clock, we chose $^{171}$Yb$^+$ for its laser accessible optical transition at 369 nm and its nuclear spin of $1/2$. The ground state then contains the $F = 0$ and $F = 1$ hyperfine levels, where the $F = 0$ level contains only a single Zeeman state (Fig. 1), simplifying the optical pumping. At the start of a clock cycle the ions are prepared in the $F = 0, m_F = 0$ ground state. A 12.6 GHz microwave field is pulsed on and off to drive a π-pulse from the $F = 0, m_F = 0$ state to the $F = 1, m_F = 0$ state. To determine how many ions made the transition, the 369 nm and the 935 nm light fields illuminate the ions, and the 297 nm ion fluorescence from the $^3D_{3/2}^{3/2}$ state to the $^2S_{1/2}$ ground state transition is collected on a detector. The detector signal is conditioned in loop control electronics to generate a control signal to lock the frequency of a local oscillator to the ion hyperfine transition. The state detection process also serves to optically pump the ions back to the $F = 0$ ground state.

II. MINIATURE VACUUM PACKAGE AND ION TRAP

The $^{171}$Yb ions are trapped inside of a 3 cm³ vacuum package (Fig. 2). This package is described in [7], and we provide a short summary here for clarity. The vacuum package is constructed from a titanium body with electrical and optical feedthroughs welded onto the body. The package has three sapphire optical feedthroughs—one is on opposite sides of the package to allow lasers to pass through the center of the ion trap, and the other is used to collect fluorescence from the trapped ions. Titanium tubes are filled with a small amount of Yb metal and welded onto the package to provide a source of neutral Yb for loading the ion trap. There are two Yb oven appendages on opposite sides of the vacuum package, one filled with isotopically purified $^{171}$Yb and the other with natural abundance Yb. The ion trap is a linear quadrupole rf Paul trap, and it is positioned in the center of the vacuum package. A copper tube is brazed onto the package to serve as a vacuum pump-out port. Also inside the vacuum package is a non-evaporable getter. Once appropriate vacuum conditions are achieved through a high temperature bake, the vacuum package is back-filled with $2 \times 10^{-6}$ torr of helium buffer gas, and the copper tube is pinched off to form a cold weld seal. The process permanently seals the vacuum package, and the getter passively maintains the vacuum.

The vacuum package is assembled with a magnetic shield and photo multiplier tube (PMT) (Fig. 3). Within the magnetic shield is a pair of C-field coils. A collection lens directs light to the PMT through several fluorescence filters that reject light at the two laser wavelengths of 369 nm and 935 nm but pass the 297 nm fluorescence. Around the Yb appendages a small oven is placed to heat the Yb to a temperature of $-450$ °C to create a vapor of neutral Yb. Ions are created inside the ion trap by illuminating the Yb-coated ion trap electrodes with 405 nm laser light. This creates photoelectrons, which then ionize the Yb through electron impact. The trapping field is created by applying a 190 V$_{pp}$ potential difference at 3 MHz to diagonal pairs of the four quadrupole electrodes. A -5 V DC potential applied to the quadrupole trap electrodes relative to the vacuum package walls is maintained to provide the longitudinal confinement of the ions.
III. LONG-TERM FREQUENCY INSTABILITY MEASUREMENTS

For the first clock measurements described in this paper, the vacuum package assembly is mounted on a 10 cm × 15 cm breadboard. For these first measurements, 20 µW of 369 nm laser light is brought to the breadboard via an optical fiber. The 369 nm laser source is a large frequency doubled commercial laser system (Toptica TA-SHG pro). The 369 nm light passes through an electrostatically actuated MEMS mechanical shutter. Also mounted on the breadboard is a 935 nm vertical cavity surface emitting laser (VCSEL), and 0.5 mW of this light is combined with the 369 nm light and directed into the vacuum package.

To operate the atomic clock, a custom-built electronics board provides microprocessor control over most functions of the physics package. The electronics board contains a 120-mW oven controlled crystal oscillator (OCXO) (CTS Valpey Corp.) operating at 10 MHz to act as a local oscillator for the clock. The stability of the OCXO is shown in Fig. 9.) A 12.6 GHz microwave signal, phase locked to the OCXO, drives the hyperfine transition in the ¹⁷¹Yb ions. The electronics board also provides current to the 935 nm VCSEL and stabilizes its temperature. The wavelength of the 935 nm VCSEL is stabilized by maximizing the fluorescence signal of the ions. Critical timing control of the clock feedback loop is provided by a field-programmable gate array (FPGA). The FPGA controls the switching of the MEMS shutter and the power from the microwave source. For the data shown in Figs. 4 and 5, the microwave interrogation time is 200 ms while the optical pumping time is 300 ms, such that a single clock cycle has a duration of 500 ms. In between clock cycles, the frequency of the microwave source is hopped to either side of the Yb hyperfine resonance with a step of 4.8 Hz, and the fluorescence signals from every two clock cycles are subtracted to provide feedback to the OCXO. The commercial 369 nm laser is controlled by its own electronics, and its wavelength is locked to a commercial wavemeter.

This breadboard clock system was delivered to the National Institute of Standards and Technology (NIST) for an independent test of its frequency stability performance. The NIST clock measurement system collected data every 12 minutes. We collected a total of 49 days of good clock data at NIST, and continuous sets of data lasting 6 days, 26 days, 3 days, 4 days, and 10 days were concatenated for Figs. 4 and 5. Dead time between data sets ranged from a few hours to a few days, and the dead time had negligible effect on the results. During this clock test, whenever there was a break in the operation of the clock, ions were reloaded into the ion trap.

As can be seen in Fig. 5, the short-term stability is degraded after day 32 of the composite data set. After day 32, the new 935 nm VCSEL came out of lock because the laser wavelength drifted to longer wavelengths over time. To bring the laser back on wavelength, the current needed to be reduced, thereby reducing the 935 nm power, resulting in a reduced fluorescence signal from the ions and a degraded short-term stability. However, the long-term stability was not affected, and using the THEO statistic [8], we demonstrated a long-term stability of $6 \times 10^{-14}$. The frequency reproducibility between data sets was $2 \times 10^{-13}$. With the 49 day measurement, the clock did not reach a flicker floor nor indicated its onset [9]. An important conclusion is that the stability of a trapped ion system utilizing a highly miniaturized package continues to improve over a 49-day run.

Fig. 4. Allan deviation measurement of the breadboard ¹⁷¹Yb⁺ clock data. Several phase steps in the data were removed prior to calculating the Allan deviation. The black line shows the Allan deviation for the first 32 days of the composite data set. The last two points of each data set are calculated using the THEO statistic.

After the first 6 days of data collection, the 935 nm VCSEL was replaced because the previous VCSEL was continuously drifting to longer wavelengths and no longer had the ability to be tuned onto the ⁵D_{3/2} to ⁵D[3/2]₁/₂ optical transition. After the VCSEL was replaced, we achieved continuous autonomous operation of the clock for 31 days. However, the data was broken up due to the 369 nm laser coming unlocked and a four day section where the clock was running at a distinctly different frequency due to excessive 369 nm laser frequency and amplitude noise. During the 31 day run, the ion trap was not reloaded, demonstrating a 1/e lifetime of the trapped ions of approximately 3 weeks.

As can be seen in Fig. 5, the short-term stability is degraded after day 32 of the composite data set. After day 32, the new 935 nm VCSEL came out of lock because the laser wavelength drifted to longer wavelengths over time. To bring the laser back on wavelength, the current needed to be reduced, thereby reducing the 935 nm power, resulting in a reduced fluorescence signal from the ions and a degraded short-term stability. However, the long-term stability was not affected, and using the THEO statistic [8], we demonstrated a long-term stability of $6 \times 10^{-14}$. The frequency reproducibility between data sets was $2 \times 10^{-13}$. With the 49 day measurement, the clock did not reach a flicker floor nor indicated its onset [9]. An important conclusion is that the stability of a trapped ion system utilizing a highly miniaturized package continues to improve over a 49-day run.

Fig. 5. The fractional frequency deviation is plotted as a function of time for the composite data set. The short-term stability of the clock can be seen to degrade at about day 32. This was due to the reduced power output of the 935 nm laser.
IV. INTEGRATED CLOCK PHYSICS PACKAGE

Subsequent to the measurements at NIST, we developed a much more integrated clock using the same vacuum package assembly that was used for the NIST measurements (and is shown in Figs. 2 and 3). The new system contains an integrated optical package, a miniaturized 369 nm laser, and a physics package interface board. In addition, a new control electronics board was made.

The integrated optical package includes the 935 nm VCSEL and its collimating lens, two half wave plates (one for 935 nm light and the other for 369 nm light), an angle polished FC fiber connector input for the 369 nm light, an assembly to hold the MEMS shutter, and a dichroic beam combiner. The 935 nm VCSEL used in the integrated optical package was newly fabricated and no longer showed the wavelength drift problem observed in the previous measurements. The most complex aspect of the integrated optical package was the MEMS shutter assembly. Since the MEMS shutter moves only 60 μm, the light must pass through a focus to be effectively shuttered by the small shutter blade. In addition, the shutter must be packaged with a 50 μm pinhole to provide an extinction ratio of 40 dB and 94% transmission when open. Alignment of the fiber to the MEMS shutter is achieved by fine positioning of the FC fiber holder. The two waveplates are mounted in holders, which allow adjustment of their angle. Overlap of the 369 nm and 935 nm beams is accomplished by active alignment of the dichroic beam combiner along with an adjustment of the 935 nm VCSEL assembly in and out of the plane of Fig. 6. Through a 5-axis adjustment scheme the two beams are overlapped, and once the correct position for the dichroic beam combiner is established, it is fixed in position with an epoxy.

The physics package interface board is designed to be mounted directly on the vacuum package. It provides the required rf trapping voltages [5], current to the C-field coil, and a microwave synthesizer to produce the 12.6 GHz microwave radiation. The synthesizer utilizes the 10 MHz signal from the OCXO on the control electronics board.

The miniaturized 369 nm laser is a direct diode source. An external cavity diode laser (ECDL) is made using a diode laser from Nichia Corp. that is selected with a wavelength of < 371 nm.
nm. Then, the laser is packaged by Ondax, Inc. with a collimating lens and a volume holographic grating, forming the ECDL (see Fig. 7(a)). The packaged size of the laser is ~1 cm³. This laser is mounted in a custom-built fiber coupling package to deliver light to the integrated optical package (see Fig. 7(b)). The fiber coupling package is 10 cm³ in size.

With the Ondax packaging, the laser wavelength is tuned with current and temperature—there is no direct control of the cavity length. By tuning the wavelength of the laser with current, we can achieve a mode-hop-free-tuning range of 8 GHz. As is typical with ECDLs, the laser has several modes that it hops through as the current is tuned, but the exact wavelength range of continuous tuning within a mode varies with temperature. This laser is quite stable over a time scale of several hours. However, over days and weeks the wavelengths of the modes drift, and new temperature and current settings need to be found so that the laser operates at the 369 nm optical transition. Long-term measurements indicate that to operate at the required wavelength, one must slowly increase the laser temperature over time. Because of the slowly drifting mode structure, the longest we could ever run the clock without the laser mode-hopping was 7 days. When the laser mode-hops, we need to stop clock operation for several hours while new laser operating parameters are found.

While operating the clock, both the 369 nm and 935 nm lasers need to be stabilized. Instead of using commercial wavemeters for laser frequency stabilization, we have utilized temperature-stabilized, miniature etalons (6 × 6 × 12 mm³) for short-term (up to several hours) laser frequency stabilization. The long-term laser stability can be in-principle maintained by servoing the etalon temperatures and using the ion signal to maximize the fluorescence. We have demonstrated locking of both the 369 nm and 935 nm lasers to etalons. In fact, referring to the data shown in Fig. 9, the 935 nm laser is locked to an etalon. The 369 nm laser was still stabilized by a wavemeter due to a minor technical conflict that prevents us from locking the 369 nm laser while recording the clock data.

The changes to the updated control electronics board most pertinent to the data shown in Fig. 9 were changes to both the hardware and firmware that allowed an increase in the total time of the clock cycle and an increased frequency resolution in the microwave synthesizer. The previous board allowed only a 655 ms clock cycle and a minimum frequency step of 1.6 Hz for the 12.6 GHz microwave synthesizer. In this new board, a clock cycle time of 20 s and a minimum microwave frequency step of 0.1 Hz are possible. Longer interrogation time allows us to take advantage of the coherence time of the trapped ions. In our case, the limiting factor is the low-power OCXO. We found the optimal clock cycle time was 1 s with 700 ms of microwave interrogation and 300 ms of optical pumping. At microwave interrogation times longer than 700 ms, the instability of the OCXO would broaden the linewidth of the Yb hyperfine transition. Fig. 9 shows the advantage of going from a 200 ms microwave interrogation time to 700 ms. To truly take advantage of the long coherence time of the buffer-gas-cooled trapped ions, a local oscillator with better stability is needed.

V. CONCLUSION

To our knowledge, this is the smallest trapped-ion frequency standard published to date. The volume of all of the components of the physics package including the etalons for laser lock is ~100 cm³, and the power consumption of the clock is <1 W. The titanium vacuum package technology has proven to be very robust. At the time of measurements of Fig. 9 the vacuum package had been sealed and passively pumped by the getter for two years. Since the first two months after sealing the vacuum package, we have seen no significant changes in the trapped ion lifetime and the number of trapped ions after reloading the trap.

An area of significant potential improvement is in the fluorescence signal size from the trapped ions. For the data in Fig. 9, the clock is operating at the photon shot noise limit. During operation of the clock, we observe that 70-80% of the ions are in the long-lived $^2F_{7/2}$-state and are not contributing to the fluorescence signal. While the clock still operates well, it would be advantageous to eliminate this significant $F$-state trapping. An additional laser could be used to move population out of the $F$-state, but this approach is not desirable for a small, low-power atomic clock. Additionally, nitrogen buffer gas has been shown to quench the $F$-state [10]. However, nitrogen is pumped by the getter in the vacuum package. It may be possible to use a hydrocarbon buffer gas since it is not pumped by the getter, and we have initial results indicating that methane may work well to quench the $F$-state. Signal improvements may also be achieved by collecting fluorescence at 369 nm. There are 200 times more photons per ion emitted at 369 nm compared to 297 nm. Thus far, we have achieved the best signal-to-noise ratio by collecting 297 nm fluorescence because of the high scattered light background from the 369 nm laser. We believe that in the miniaturized physics package, with careful control of the 369 nm light scatter, detection at 369 nm is possible, which would greatly enhance the signal size.

While we have achieved a great deal of miniaturization, additional improvements in the component technologies of the physics package of the clock could lead to perhaps another factor-of-ten reduction in its size. We have already
demonstrated a vacuum package with a volume < 1 cm³, and this result is to be reported in a later publication. The component that requires the most development at this point is the 369 nm laser. The described direct diode source consumes several hundred milliwatts of power and has poor long-term stability. Improvements in the 369 nm laser technology would enable a new class of low-power, long-term stable trapped-ion atomic clock. With improved SNR by use of 369 nm fluorescence detection, acquisition of a better-frequency-stability miniature local oscillator becomes more important. Overall, the ¹⁷¹Yb ion clock shows excellent stability in a small package and shows promise as a highly miniaturized frequency standard.

REFERENCES
Towards a high-performance microwave frequency standard based on $^{113}\text{Cd}^+$ ions

Jianwei Zhang, Kai Miao, Lijun Wang  
Department of Precision Instrument, Tsinghua University  
State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University  
Joint Institute for Measurement Science (JMI), Tsinghua University  
Beijing, China  
zhangjw@tsinghua.edu.cn

Xiaolin Sun, Lijun Wang  
Department of Physics, Tsinghua University  
Joint Institute for Measurement Science (JMI), Tsinghua University  
Beijing, China

Abstract—Microwave frequency standards based on ions are promising as the next generation atomic clocks. This paper reports the recent progress of the microwave frequency standard based on laser-cooled $^{113}\text{Cd}^+$ ions developed at Tsinghua University. Based on the current experimental setup, the ground-state hyperfine splitting of $^{113}\text{Cd}^+$ is measured with higher precision and a better short-term frequency stability is obtained than before. To improve the performance of the Cd$^+$ clock further, a new experimental setup with dual-traps is designed, and bi-ions operation is proposed.

Keywords—microwave frequency standard; cadmium ions; ground-state hyperfine splitting; frequency stability

I. INTRODUCTION

Trapped ions are widely applied in the frequency metrology due to the well isolation from the external fields. The frequency inaccuracy is achieved to $8.6 \times 10^{-18}$ for the optical clock based on single $\text{Al}^+$ [1], and to $2 \times 10^{-17}$ for the optical clock based on $\text{Sr}^+$ [2]. A microwave frequency based on several laser-cooled $\text{Hg}^+$ ions has been demonstrated, and the frequency stability reaches to $3.3 \times 10^{-13} \tau^{-1/2}$ and the frequency uncertainty to $1 \times 10^{-14}$ [3]. Meanwhile, a microwave frequency based on buffer-gas-cooled $\text{Hg}^+$ is investigated [4, 5], and a stability of $5 \times 10^{-13} \tau^{-1/2}$ has been demonstrated at Jet Propulsion Laboratory (JPL) [5]. And a very compact $\text{Hg}^+$ clock is developed at JPL, which is going to be applied in the deep space navigation [6].

The microwave frequency standard based on laser-cooled $^{113}\text{Cd}^+$ ions has a promising performance of frequency stability and uncertainty [7]. A group at the Communications Research Laboratory, Japan, investigated $^{113}\text{Cd}^+$ ions in a parabolic ion trap and a linear ion trap, respectively, and measured the ground-state hyperfine splitting to 15 199 862 858(2) Hz [8, 9]. Later, a group at JPL measured the hyperfine splitting of $^{113}\text{Cd}^+$ again, and obtained 15 199 862 855.0(12) Hz [10]. The obtained performance of Cd$^+$ clock is not as good as what we expected. Based on the previous works, we upgraded the magnetic shield to reduce the influence of the fluctuation of the ambient magnetic field, stabilized the power of the probe laser to prevent the technical noises of lasers from clock signal, and installed one more Photomultiplier tube (PMT) to increase the collection efficiency of the fluorescence signal. The signal-to-noise-ratio (SNR) is improved than before, and a series of new measurements are carried out.

In order to evaluate the performance of the Cd$^+$ clock, a fiber link with active phase compensation [13] developed to transfer the standard frequency signal of an active hydrogen maser (H-maser) from National Institute of Metrology, China (NIM) to JMI, which is referenced to the UTC (NIM).

We locked the local oscillator, an oven controlled crystal oscillator (OCXO), to the $^{113}\text{Cd}^+$ ions, and compared it with the reproduced standard frequency of UTC (NIM) transferred from NIM to Tsinghua University. The frequency stability of the Cd$^+$ clock is measured, and we obtained $6 \times 10^{-13} \tau^{-1/2}$, which is obviously better than the previous results, $1.7 \times 10^{-12} \tau^{-1/2}$. Continuous comparison between the Cd$^+$ clock and the standard reference gives us a chance to measure the ground-state hyperfine splitting of $^{113}\text{Cd}^+$ with higher precision than before. The measurement results indicates that the measurement precision of the clock transition is improved from $5.3 \times 10^{-13}$ to $6 \times 10^{-14}$ [14].

II. MEASUREMENT OF THE FREQUENCY STABILITY AND CLOCK TRANSITION

The obtained performance of Cd$^+$ clock is not as good as what we expected. Based on the previous works, we upgraded the magnetic shield to reduce the influence of the fluctuation of the ambient magnetic field, stabilized the power of the probe laser to prevent the technical noises of lasers from clock signal, and installed one more Photomultiplier tube (PMT) to increase the collection efficiency of the fluorescence signal. The signal-to-noise-ratio (SNR) is improved than before, and a series of new measurements are carried out.

In order to evaluate the performance of the Cd$^+$ clock, a fiber link with active phase compensation [13] developed to transfer the standard frequency signal of an active hydrogen maser (H-maser) from National Institute of Metrology, China (NIM) to JMI, which is referenced to the UTC (NIM).}

III. NEW EXPERIMENTAL SETUP

Although the frequency stability and frequency uncertainty of Cd$^+$ clock are improved than before, there are still spaces to...
After enough ions are loaded and cooled, the dc voltage on the middle segments to confine ions in the loading zone. The middle segments of the electrodes work as a gating electrodes. During ions loading, a dc voltage is applied on the middle segments again to confine ions in the working zone. The laser beam is aligned along the axis of the trap to cool, pump and detect ions.

In the ion traps, $^{24}\text{Mg}^+$ ions are going to be trapped simultaneously with $^{113}\text{Cd}^+$ ions to sympathetically cool Cd$^+$ ions during interrogation. According to our previous investigation [15], the fluorescence decay during detection can be reduced greatly if the temperature of $^{113}\text{Cd}^+$ ions can always be kept at a low value. This will give us better SNR to improve short-term frequency stability. The transition line of $^2\text{S}_{1/2} \rightarrow ^2\text{P}_{3/2}$ of $^{24}\text{Mg}^+$ is about 280 nm and the natural line width is about 40 MHz. It is convenient to Doppler cool Mg$^+$ ion via this transition, and the laser is not difficult to access.

The light shift induced by 280 nm laser could be one of the biggest issues for the bi-ions clock. Because the 280 nm laser resonates with Mg$^+$ ions only, the Mg$^+$ ions can be pushed to one end of the bi-ions cloud if the temperature of the ion cloud is low enough, and the incident angle and position of the laser beam is properly controlled. Thus, Cd$^+$ and Mg$^+$ can be separated in space. According to our calculation, the light shift can be controlled to a low value if the Cd$^+$ ions are not in the 280 nm laser field directly.

For the pulse-mode operation of the ion clock, the dead time due to the unavoidable initialization and detection time causes the Dick effect to degrade the short-term frequency stability. In the new designed physical package, two identical ion traps are implemented to realize the interleaving interrogation of the local oscillator.

The interrogation cycles of the dual-traps clock are shown in Fig. 2. The same continuous microwave pulses are applied in the two trap simultaneously to interact with ions, and the period of the microwave pulses is $T$. For both of the traps, the pump lasers, detection lasers and detectors are also controlled by the similar sequences, and the period is $2T$. As shown in Fig. 2, the pump laser is shined for trap-1 to initialize ions before the first microwave pulse, and the detection laser is shined for trap-2 to detect ions' states after the first microwave pulse. Around the second microwave pulse, the pump laser is shined for trap-2 before the pulse, and the detection laser is shined for trap-1 after the pulse. At the third microwave pulse, the same sequences are repeated as at the first microwave pulse.
for both traps.

With the dual-traps, the local oscillator is always steered by one of the two traps. Thus, the Dick effect due to the dead time in the interrogation cycles is avoided, and the Dick-effect-limited frequency stability can be greatly reduced. For the new experimental setup with an OCXO as the local oscillator, the Dick-effect-limited frequency stability is expected to be \(1.2 \times 10^{-14} \tau^{-1/2}\). To our knowledge, this interleaving method is first proposed by Dick et. al. [17]. Later, it is also demonstrated in cesium fountains by Biedermann et. al. [18]. With the new setup, the frequency stability of the \(\text{C}^+\) clock will decrease as \(1/\tau\) over a long sampling time until that the stability reaches the quantum-noise limit [19], which was also demonstrated in Ref. [18].

IV. DISCUSSION AND CONCLUSION

The physical package of the new experimental setup has been designed, and most of the parts are fabricated. In the next few weeks, the assembling of the ion traps and vacuum tight test of the vacuum package will be finished. The 280 nm laser for \(\text{Mg}^+\) ions is also ready. So, the preliminary results of the dual-traps clock with bi-ions are expected in the late of this year. The temperature of the \(\text{C}^+\) ions can be keep low thanks to the sympathetic cooling of \(\text{Mg}^+\) ions, which improves the SNR of the clock signal greatly. A SNR of 1000 is expected. Thus, the quantum noise limit of the frequency stability is estimated to be \(2 \times 10^{-14} \tau^{-1/2}\). Meanwhile, The Dick-effect-limited frequency stability is expected to be \(1.2 \times 10^{-14} \tau^{-1/2}\) thanks to the interleaving operation. With the new setup, a high-performance \(\text{C}^+\) clock with high performance is expected.

ACKNOWLEDGMENT

We thank X. Zhu, C. Gao and B. Wang for the reproduced standard frequency signal of UTC(NIM) transferred from NIM to Tsinghua University, and thank K. Liang for the help about UTC(NIM).

REFERENCES

ELSTAB - electronically stabilized time and frequency distribution over optical fiber - an overview

P. Krehlik and Ł. Śliwczyński
Department of Electronics
AGH University of Science and Technology
Krakow, Poland
krehlik@agh.edu.pl

Abstract—This paper presents an overview of the ELSTAB technology developed at AGH University of Science and Technology for time and RF frequency distribution over optical fibers. The general idea of the stabilization of the fiber link delay is described, and its application for joint time and frequency distribution is presented. Next some extensions of the basic solution, proposed for long-distance and multipoint links are introduced. Finally, two field-deployed installations of the ELSTAB system in Poland are characterized.

Keywords— fiber optic; time transfer; frequency transfer; delay stabilization

I. INTRODUCTION

An access to precise time signals and frequency references is important in many areas of human activities, involving commercial applications (e.g. telecommunication, navigation, metrology, developing the timescales) as well as scientific experiments and development of novel frequency standards. Aside of widespread satellite methods of the time and frequency transfer, solutions based on optical fibers become the rapidly emerging alternative, offering much better accuracy. However, as any transmission medium, the optical fibers display some fluctuations of the propagation delay, caused mainly by the influence of the temperature on the refractive index of silica glass. For typical telecommunication-grade fibers the thermal coefficient of propagation delay is about 38 ps/(km·K) [1], which leads to tens of nanoseconds of seasonal fluctuations even for moderate transmission distance. Various ideas for reducing the propagation instability are based on redirecting the signal reaching the remote end of the link backward to the local side, and arranging a feedback system which can compensate the fluctuations of the phase (or delay) of propagating signals. (The two fundamental ideas were presented in [2] and [3].) Generally, the solutions stabilizing the phase of transmitted sinusoidal signal are well suited for high-quality frequency distribution, but can not be extended for time distribution. On the other hand, systems stabilizing the group delay of the signal reaching the remote end can be used both for frequency and timing signals.

II. BASICS OF THE ELSTAB SOLUTION

The underlying idea of the ELSTAB solution is to implement the compensation of the fiber delay fluctuations in the electronic domain, by using a pair of precisely matched variable delay lines placed both in the forward and backward paths of the delay-locked-loop (DLL) structure - see Fig. 1.

The frequency signal (10 or 100 MHz) propagates through the forward-direction variable delay line, electro-optic (E/O) converter (intensity-modulated laser transmitter) and optical circulator to the optical port of the local module. Next, through the fiber link, the signal reaches the remote module, where it is redirected backward to the fiber and then to the local module. Here, via the backward path (O/E and the second variable delay line), it enters the phase detector. The phase detector

This work was partially supported by EMRP (SIB-02 NEAT-FT project), NCN (DEC-2011/03/B/ST7/01833 project), and NCBiR (PBS1/A3/13/2012 project).

![Simplified block diagram of the ELSTAB system.](image_url)

Fig. 1. Simplified block diagram of the ELSTAB system.

![Allan deviation of stabilized and not-stabilized frequency distribution over 60 km-long urban fiber link, calculated from two-weeks measurement.](image_url)

Fig. 2. Allan deviation of stabilized and not-stabilized frequency distribution over 60 km-long urban fiber link, calculated from two-weeks measurement.
senses the phase difference between the input and feedback signals, and cancels it by driving the variable delay lines. Thus the round-trip delay in the DLL system is kept constant, unaffected by the fluctuations of the fiber delay. Assuming that the fiber delay variations are same in both directions, and that the tuning characteristic of both delay lines are identical, it may be noticed that also the delay from the input to the remote output is constant. (For more detailed analysis and discussion - see [4].) The demand of excellent matching of the two delay lines was fulfilled by designing a pair of delay lines as single CMOS integrated circuit.

The efficiency of this stabilization system may be observed in Fig. 2, where the overlapping Allan deviation (ADEV) is plotted for simple local-to-remote transmission (with no stabilization), and for active closed loop stabilization. For our stabilized system the ADEV is monotonically decreasing with averaging time, falling to $1.2 \times 10^{-7}$ for averaging time of $10^5$ s. In case of non-stabilized transfer the fiber temperature fluctuations degrade the ADEV by about two orders of magnitude.

III. JOINT TIME AND FREQUENCY DISTRIBUTION

To add the time transfer we developed the dedicated embedder, in which the time signal (in form of pulse-per-second – 1PPS) is incorporated as a certain violation (phase modulation) made on the periodic signal of frequency. The two de-embedders (remote and local ones) extract the time signal by detecting the violations, and restore the original periodicity of the frequency signal. Because the time signal passes the same path as the frequency signal, its delay is also stabilized by the DLL system. To obtain the possibility of calibration of the time transfer delay, two auxiliary 1PPS outputs are added in the local module. The exact value of the delay of time signal outgoing the remote module may be determined by measuring the round-trip delay, dividing it by two and applying some corrections resulting from the chromatic dispersion of the fiber, Sagnac effect, and hardware delays inside the local and remote modules. The detailed analysis of the calibration procedure and uncertainty budget is out of the scope of this paper, but may be find in [5] and [6]. The short conclusion from this works is that for up to 500 km-long links the standard uncertainty of time calibration is below 25 ps. The stability of time distribution, illustrated in form of a time deviation (TDEV) is plotted in Fig. 4. The value below 1 ps was observed for wide range of averaging periods.

IV. LONG-HAUL EXTENSION

The basic configuration of the ELSTAB distribution system comprises the local and remote modules connected via singlemode optical fiber, and is capable of time and frequency distribution at distances up to approximately 100 km. This limitation of the reach is twofold: one factor is a power budget of the system, which limits the maximum fiber attenuation to 25 dB, and the second factor is the tuning range of the delay lines (being 100 ns), which might be insufficient for all-over-year delay compensation for longer links.

To overcome the attenuation-caused limit we developed a single-path bidirectional amplifiers (SPBA), based on erbium-doped fiber – see Fig. 5. Thanks to bidirectional operation with the same optical path for forward and backward directions, the propagation delay (and its possible fluctuations) is same for both direction, thus insertion of such amplifier do not destroyed the symmetry of the optical path. However, a long haul link with many SPBAs is vulnerable to backsattered and reflected signals propagating in the fibers, and precise analysis and optimization should be performed before particular installation [7, 8].

The extension of the compensation range was realized by a hybrid combination of the electronic delay lines, and additional switched optical delays, realized in form of short pieces of spooled fiber. In this solution the continuous compensation of the fiber delay fluctuations is realized by electronic variable delays, as in standard solution. Switching of optical delays is activated only occasionally, when the tuning of electronic delays reaches its limits. During the switching and relocking period the output frequency signal is generated continuously in

!

Fig. 3. Joint time and frequency distribution scheme.

Fig. 4. Time deviation of stabilized and not stabilized frequency distribution over 60 km-long urban fiber link, calculated from two-weeks measurement.

Fig. 5. Single-path bidirectional amplifier.
a PLL system temporarily driven into a holdover mode. Continuity of 1 PPS transmission is guaranteed because the switching and relocking procedure fits between two consecutive PPS pulses. Measured phase (delay) discontinuity caused by optical delay switching is not higher than a few picoseconds. In current design the switched optical delays cover the range of 1150 ns (in 50 ns steps), which is enough for the link length up to 1000 km [9].

V. MULTIPONT DISTRIBUTION

Another possible extension of the basic ELSTAB setup is a multipoint distribution of the time and frequency signals, realized by tapping of the main link. It should be noticed however, that the compensation of the fiber delay fluctuations is effective only at the end of the fiber, so a simple tapping of the forward-propagating signal is not sufficient to obtain stabilized signals. Fortunately it was observed that when tapping both forward and backward signals at any particular point of the fiber link, the tapped signals undergo delay fluctuations of the exactly same magnitude but opposite sign. Thus, using once again a pair of matched variable delay lines in a proper arrangement, the stabilized tapping module may be realized [10, 11]. A block diagram of this module is depicted in Fig. 6. It may be noticed that the tapping module comprises similar building blocks as local and remote modules, which make the system flexible in configuration.

The stability of time and frequency signals at tapping module output (see Fig. 7) is slightly worse than at remote output, because the chain of electronics blocs involved is noticeably higher. From the other hand, the signals quality at the remote (main) output is unaffected, as the additional electronics of the tapping module do not influence a core system. The absolute calibration of the delay of the 1PPS outgoing the tapping module is also available, with help of two auxiliary outputs (AuxA and AuxB), used after particular installation in calibration procedure [11].

The idea of the tapping node would also be extended in the way that the tapping point could initialize a side branch, which provide the stabilized signals to locations apart from the main link [10].

VI. OPERATIONAL INSTALLATIONS IN POLAND

Thanks to the OPTIME project supported by Polish National Centre for Research and Development (NCBiR) we started a cooperation with Polish timekeeping laboratories and fiber network operators, oriented to build an operational time and frequency distribution network. The first step was a link connecting UTC(PL) and UTC(AOS) labs over 420 km-long fiber. This link was lunched in January 2012 and is operating continuously with only a few gaps caused by the fiber breaks. In July 2013 the BIPM-developed GNSS time transfer calibrator, named METODE, was installed in UTC(PL) and UTC(AOS) labs side by side with the optical link, to compare the time calibration. It occurred that the calibration of time transfer performed in our system and the value obtained with the METODE differs only by 0.56(±0.2) ns, which is below the

![Fig. 6. Block diagram of the tapping module.](image)

![Fig. 7. ADEV of 10 MHz signal at the tapping module output, compared with ADEV at remote output.](image)

![Fig. 8. ELSTAB links installed in Poland.](image)
uncertainty specified for the METODE calibrator (1.5 ns). This was probably the first comparison of two completely different time transfer schemes at this level of consistency [12].

In December 2014 we launched a second link, providing the signals from UTC(AOS) to National Laboratory of Atomic, Molecular and Optical Physics (KL FAMO) in Toruń. The next planned step is to install the tapping module in Poznań, which will be used to feed the local time and frequency repository, which is being built within the OPTIME project.

VII. SUMMARY

In this paper we presented a short overview of the ELSTAB technology, with thorough references to more detailed analysis and experimental results, spread in our previous papers.

The general characteristics of our solution is that by usage of a relatively simple solution based on electronic variable delay lines the fluctuations of the fiber propagation delay may be compensated with picoseconds accuracy. This allow to obtain frequency distribution stability at order of $1 \times 10^{-17}$ (for $10^7$ s averaging period). The timing signal (1PPS) stability may be characterized by TDEV well below 1 ps for wide range of averaging periods.

The unique feature of the ELSTAB is time calibration facility, allowing to determine the delay of the output 1PPS with respect to UTC(k) with uncertainty in range of 10 ... 30 ps.

Using various extensions of the basic system one could build a long-haul and multipoint distribution network. From practical point of view it is also important that the local and remote modules (and tapping module as well) are relatively small electronic devices consuming below 20 W of power each, needing no any maintenance during continuous operation.

REFERENCES

Comparison of forward- and backward-propagating optical-fiber-induced noise for application to optical fiber frequency transfer

James P. Cahill\textsuperscript{*+}, Olukayode Okusaga\textsuperscript{+}, Weimin Zhou\textsuperscript{*}, Curtis R. Menyuk\textsuperscript{*}, and Gary M. Carter\textsuperscript{*}

\textsuperscript{*}University of Maryland: Baltimore County, Baltimore, MD, USA
\textsuperscript{+}U.S. Army Research Laboratory, Adelphi, MD, USA
E-mail: james.p.cahill.ctr@us.army.mil

Abstract—Current schemes for the photonic transfer of radio frequencies rely on bidirectional active feedback in which the optical signal must propagate in both directions through a single optical fiber. This requirement is not readily compatible with existing optical fiber networks; hence, it is important to develop alternate means of suppressing optical-fiber-induced noise. Previously, we experimentally characterized an optical-fiber-length-dependent noise source that contributes to the phase noise of the radio frequency signal, and we demonstrated that it can be mitigated by frequency dithering the laser. However, we have not developed an adequate model to describe the noise source. In this work, we compared the experimentally-measured forward- and backward-propagating optical-intensity noise spectra. We found that the power of the backward-propagating noise is over 40 dB higher than the power of the forward-propagating noise. We also found that the forward-propagating noise scales faster with respect to optical-fiber length than does the backward-propagating noise. These results will aid the development of a complete theory to describe the optical-fiber-length-dependent noise source.

Keywords—Radio frequency (RF) transfer, RF-photonics, Rayleigh scattering

I. INTRODUCTION

Frequency transfer over optical fibers is a promising technique to enable the transmission of highly stable frequency sources between metrology laboratories and to other users across the world. It is well-known that noise occurring in the optical fiber limits the frequency stability that can be transmitted via optical fibers. Previous schemes have mitigated this noise by using active feedback \cite{1,2} that requires light to be sent bi-directionally through the same optical fiber \cite{3}. This requirement is not readily compatible with the majority of installed fiber-optic networks. Hence, it is important to identify the physical mechanisms that lead to the optical-fiber-induced noise and to use this knowledge to explore alternate methods of mitigating this noise. In our past work, we have shown experimentally that optical-intensity noise that is induced in the optical fiber can be converted into phase noise of the transmitted frequency reference \cite{4}. While we initially concluded that the physical mechanism causing the optical-intensity noise was a third-order nonlinear scattering mechanism—guided entropy mode Rayleigh scattering—in the linear regime \cite{5}, recent experimental results have shown that this physical picture is inadequate to describe the noise source \cite{6}. As we search for a theory that completely explains our results, it is helpful to further characterize the optical-fiber-induced intensity noise.

In this work, we experimentally compared the intensity noise of the forward- and backward-propagating light. We found that the shape of the intensity noise spectra of the forward- and backward-propagating light is similar, but the power is 40 or more dB higher in the backward direction. These results are consistent with the Rayleigh-scattering-induced interferometric optical-intensity noise described in \cite{7,8}. We also found that the ratio of the backward- to the forward-propagating optical-intensity noise decreases with increasing optical fiber length, which indicates that the optical-intensity noise of the forward-propagating light grows faster with the optical fiber length. More theoretical work is required to definitively establish the physical mechanism that causes the observed optical-intensity noise.

II. EXPERIMENTAL SETUPS

A. Forward-propagating light

Fig. 1. Experimental apparatus to measure the forward-propagating optical intensity noise.

The measurement system that we used to measure the optical intensity noise of the forward-propagating light is shown in Fig. 1. It is based on the cross-correlation, relative-intensity-noise measurement setup described in \cite{9,10}. The light for the experiment was generated by a narrow linewidth, grating-stabilized distributed feedback laser (Teraxion NLL) with a wavelength of 1550 nm. The light passed through a variable optical attenuator (VOA) and then through the optical fiber under test, where the scattering occurred. We will refer to the optical power measured directly after the VOA as the
input optical power. At the end of the optical fiber, a 50% coupler split the light into two identical arms. In each arm, the light was detected by a photodiode and amplified by a custom-designed low noise, AC-coupled amplifier (LNA). The LNAs have a voltage gain of 100 and an operating range of 1 Hz – 1 MHz. A vector signal analyzer sampled the voltage output of each amplifier and calculated the cross-correlation spectrum. Averaging the cross-correlation spectrum of the two arms mitigated uncorrelated noise such as amplifier noise and photodiode shot and thermal noise.

The power spectrum measured by the vector signal analyzer is proportional to the power spectrum of the intensity noise of the optical signal that illuminates the photodetector. The relation between the power spectral density of the forward-propagating optical-intensity noise, $S^{f}_{IN}$, with the power spectral density of the voltage measured by the vector spectrum analyzer, $S_{VSA}$, can be written as

\[
S^{f}_{IN}(f) = \frac{R_{VSA}}{\rho^2 G^2 R_{sh}^2} S_{VSA}(f), \tag{1}
\]

where $R_{VSA}$ is the impedance of the vector spectrum analyzer in ohms, $G$ is the voltage gain of the low noise amplifier, $\rho$ is the responsivity of the photodetector in amperes per watt, $R_{sh}$ is the effective shunt impedance in parallel with the photodetector in ohms, and $P_{out}$ is the optical power measured directly before the photodetector in watts. We note that $S^{f}_{IN}$ and $S_{VSA}$ have units of milliwatts per hertz.

### B. Backward-propagating light

![Fig. 2. Experimental apparatus to measure the backward-propagating optical intensity noise.](image)

The measurement system for the backward-propagating optical intensity noise is shown in Fig. 2. It uses a balanced detector optical homodyne receiver similar to the apparatus described in [11] to detect the optical intensity noise with increased sensitivity. The system used the same grating-stabilized DFB laser to generate the light. A 75%-25% coupler splits the light into two arms. One arm (the lower arm in Fig. 2) serves as a local oscillator, while the other (the upper arm in Fig. 2) guides the light through a variable optical attenuator (VOA) and into an optical circulator. The optical power measured after the VOA is denoted the input optical power, $P_{in}$.

The circulator is a directional device, which couples light that enters arm 1 into arm 2 and light that enters arm 2 into arm 3. The circulator isolates the output of arm 3 from the light that is input to arm 1. Hence, it directs the light from the VOA into the optical fiber under test, while preventing this light from traveling to the output of arm 3. Scattering occurs in the optical fiber, and the light that scatters backward travels back through arm 2 of the circulator and into arm 3, which is attached to a two-by-two 50% coupler. The backward scattered light combines with the local oscillator in the coupler, and the two outputs illuminate the diodes of a 50-Ω terminated balanced photodetector (balanced PD). In this detection scheme, the low power scattering signal can be measured with increased sensitivity due to the local oscillator. Moreover, the balanced PD uses common mode rejection to reduce the local-oscillator-induced noise terms that limit the noise floor of the measurement system. The vector signal analyzer samples the voltage across the balanced PD’s termination resistor and computes the power spectral density.

The power spectral density measured by the vector signal analyzer is proportional to the optical intensity noise of the backward-propagating light. The relation can be written as

\[
S^{b}_{IN}(f) = \frac{2R_{VSA}}{R_{lo} \rho P_{LO}} S_{VSA}(f), \tag{2}
\]

where $S^{b}_{IN}$ is the power spectral density of the optical-intensity noise of the backward-propagating light in milliwatts per hertz, and $P_{LO}$ is the optical power of the light in the local oscillator arm in watts.

### III. EXPERIMENTAL RESULTS

![Fig. 3. Optical intensity noise spectra of the forward- and backward-propagating light.](image)
Figure 4 displays the mean ratio of the backward- to the forward-propagating optical-intensity noise for several lengths of optical fiber. The forward-propagating intensity noise was measured with an input optical power of 6 dBm, and the backward-propagating intensity noise was measured with an input optical power of 0 dBm. Since the optical-input-power dependence of the backward-propagating optical-intensity noise is linear [6], we accounted for this discrepancy by multiplying the backward-propagating optical-intensity noise by a factor of 4. For all optical-fiber lengths measured, the ratio is greater than 40 dB. There is a monotonic negative trend in the ratio with optical-fiber length. This result indicates that the forward-propagating optical-intensity noise grows faster with the optical-fiber length than does the backward-propagating optical-intensity noise.

IV. DISCUSSION

In order to aid the development of a theoretical model, we enumerated in [6] the major characteristics of the optical-fiber-induced intensity noise. With the addition of the observations made in this paper, the list becomes:

1. The bandwidth of the optical-intensity noise is between 10 kHz and 100 kHz for the lasers used in [6], and the bandwidth depends on the laser phase noise.
2. The intensity noise spectrum is symmetric about the optical carrier—i.e., it has the same power spectral density at a given positive offset frequency as it does at the corresponding negative offset frequency.
3. The intensity noise power grows in direct proportion to the incident optical power.
4. The intensity noise power grows superlinearly, but not exponentially, with the fiber length in the offset frequency region 500 Hz – 10 kHz for lasers with phase noise comparable to the grating-stabilized DFB laser used in this paper. The intensity noise power grows in direct proportion to the optical fiber length for lasers with higher phase noise, such as the “broad linewidth” laser used in [6].
5. The laser intensity noise has little effect on the optical-intensity noise spectrum in the range of 50 Hz – 100 kHz.
6. For lasers with phase noise comparable to the grating-stabilized DFB laser used in this paper, the scaling of the optical-intensity noise power with optical-fiber length changes when the optical-fiber length exceeds 10 km.
7. The power of the backward-propagating optical-intensity noise is over 40 dB greater than the power of the forward-propagating optical-intensity noise for all optical-fiber lengths measured.
8. The ratio of the backward- to the forward-propagating optical-intensity noise has a monotonic, negative trend with respect to the optical-fiber length, which indicates that the forward-propagating optical-intensity noise scales differently with optical-fiber length than the backward-propagating optical-intensity noise.

In [6], we noted that the dependence on the laser phase noise and the superlinear growth of the optical-intensity noise within a limited frequency band are consistent with a frequency-dependent interaction of the laser phase noise, such as the Rayleigh-scattering-induced interferometric noise described in [7,8]. This process is also consistent with a many-orders-of-magnitude ratio of the backward- to the forward-propagating optical-intensity noise, since the backward-propagating optical-intensity noise is the result of a single reflection, whereas the forward-propagating optical-intensity noise is the result of two reflections. Notably, Rayleigh-scattering-induced interferometric noise has not been analyzed extensively in the limit of low-phase-noise lasers; so, more theoretical work is required to determine the predicted scaling with optical-fiber length of the Rayleigh-scattering-induced interferometric noise in this limit. Hence, while the experimental results in this paper are consistent with Rayleigh-scattering-induced interferometric noise, further theoretical study is required to definitively establish the physical mechanism or mechanisms that cause all of the observed characteristics of the optical-intensity noise.

V. CONCLUSION

As we have shown in previous work, optical-fiber-induced intensity noise can degrade a stable RF frequency reference that is transmitted through a photonic link. It is critical to determine the physical mechanism that causes the optical-fiber-induced intensity noise. In this work, we have demonstrated two additional characteristics of the optical-fiber-induced optical-intensity noise: (1) the backward-propagating intensity noise has a similar shape to but is over 40 dB higher than the forward-propagating intensity noise and (2) the forward-propagating intensity noise scales faster with the optical-fiber length than the backward-propagating intensity noise. These observations add to the list of characteristics that can be used to determine the physical mechanism that causes the optical-fiber-induced intensity noise.

REFERENCES


Abstract—Optical fiber links are beneficial not only for frequency metrology, but also for other fields of physical research, such as radioastronomy and geodesy. We realized the first optical fiber link from a National Institute of Metrology to a Radiotelescope used for Very Long Baseline Interferometry and for geodetic measurements, over a distance of 544 km. We performed a remote calibration of the hydrogen-maser used as a frequency reference at Medicina Radiotelescopes, in central Italy, against the Italian Cs fountain primary frequency standard. The comparison was limited by the statistical uncertainty of the hydrogen-masers. This experiment demonstrates that optical links can provide radioastronomical facilities with very accurate and stable frequency references, in perspective better than the currently used hydrogen-masers. This opens new perspectives in the ultimate limits of VLBI and a more precise space geodesy.

Keywords—coherent optical links; frequency dissemination; space geodesy; VLBI

I. INTRODUCTION

Phase-stabilized optical links have proved to be the most effective way to transfer ultrastable frequency signals over continental distances [1]–[4]. In fact, the stability of frequency dissemination via optical fiber is improved by 5 orders of magnitude with respect to that of satellite links, achieving the $10^{-19}$ level in hours of operation [5]. This represents an important technology for primary metrology, as it is the only way to compare remote optical clocks at their intrinsic level of uncertainty, now $10^{-18}$ [6], [7]; in addition, also remote primary standards can be effectively compared in hours instead of days. For this reason, many National Metrology Institutes are developing long-haul fiber backbones that can replace satellite links over continental distances on the time-base of some years. A more stable and more accurate frequency dissemination will be not only beneficial to primary metrology, but also to a variety of other applications, such as fundamental physics, geodesy and Very Long Baseline Interferometry (VLBI).

In this work, we investigate the potential impact of fiber-based frequency dissemination on VLBI techniques. VLBI is based on the simultaneous measurement of the same radio-source in the sky with different antennas, separated by many baselines $D_i$, each up to thousands of kilometers long. The final angular resolution of the array, obtained by correlating all data streams, is improved by the ratio $D_{\text{max}}/d$ with respect of that of a single dish with aperture $d$ [8]. The typical central observation frequencies span from 1 GHz to 26 GHz, with bandwidths from hundreds of megahertz to 1 GHz; thus, each antenna is equipped with a hydrogen-maser (HM) that serves both as a local oscillator for frequency down-conversion of the collected signal and for proper sampling and timing during the signal processing at each radiotelescope.

VLBI plays an important role also to obtain high-precision geodetic data, as it provides access to the best possible inertial reference system, made by quasars located at the edge of the observable Universe. Modern geodynamic VLBI measurement campaigns are based on successive observations of radio-sources all over the sky from many radiotelescopes spread all over the Earth. The differential delay resolution achievable by cross-correlating VLBI data is related to several parameters, among which is also the instability of the local HM.

Although HM frequency stability is adequate for present radioastronomy applications, the challenge of observations at higher frequency [9] and improved methods to model the tropospheric delay [10] raise the issue of the local oscillator instability. In geodesy, the goal of 1 mm positioning precision cannot be achieved with state-of-the-art HMs [11].

Optical links may offer some solutions: from one point of view, they enable the frequency distribution of optical atomic clocks, whose stability is three orders of magnitude better than that of a HM; in addition, they enable the dissemination of the same frequency standard at multiple antennas, that will allow a complete rejection of the clock instability.

Here, we present the first realization of a phase-coherent optical link and the ultrastable frequency dissemination from a National Metrology Institute to a VLBI site. The Italian National Metrology Institute (INRIM) realizes and maintains in Italy the definition of the SI second with the new nitrogen-cooled Cs fountain primary frequency standard ItCsF2, that is fully operative since 2013. Its Type-B uncertainty is $1.7 \times 10^{-16}$, while its short term stability in the high-density regime is $2 \times 10^{-13}$ [12], [13]. ItCsF2 has been used to calibrate TAI providing nine frequency evaluations during 2014, for a total measurement time of 165 days.

To perform the frequency dissemination on a national scale, INRIM has developed the LIFT project (the Italian Link for VLBI). INRIM has developed the LIFT project (the Italian Link for Time and Frequency) [3]. The present partners of this project are the European Laboratory for Non Linear Spectroscopy and the Institute of Optics in Florence, and the Institute of Radioastronomy of the National Institute of Astrophysics (INAF-IRA) in Bologna. INAF maintains in Italy three single-dish radiotelescopes, located in Medicina (Bologna), Noto (Sicily) and Cagliari (Sardinia). These antennas are part of the VLBI global network and of the European VLBI Network. INAF-IRA is member of the Joint Institute for VLBI in Europe and of the International VLBI Service for
Geodesy and Astronomy (IVS).
The map in Fig. 1 shows the present backbone for frequency dissemination in Italy. The 642 km link branch between INRIM and LENS is under operation since 2013. In 2014, we have realized a second branch, by splitting the radiation in equal parts in Bologna shelter, in central Italy. This arm connects INRIM to Medicina Radiotelescope (MR), at a distance of 544 km, and has been used to perform a remote calibration of the hydrogen-maser there located.

Our experiment demonstrates that remotely disseminated frequency standards can be a viable alternative to local HMs. This is a first step towards more extensive studies, where a remotely disseminated frequency will replace the local calibration of the hydrogen-maser there located.

In the following paragraphs we will describe the apparatus and the experimental results, and discuss the main criticalities and possible solutions.

II. THE EXPERIMENT

The setup of our experiment is shown in Fig. 2. An ultrastable laser is generated by frequency-locking a 1542 nm fiber laser to a high-finesse Fabry-Pérot optical cavity (FPC), with a residual instability of $8 \times 10^{-15}$ at 1 s. The frequency drift of this source is typically $10^{-15}$$/s$ [14]. To allow its use as an absolute frequency reference, we lock the laser to an HM ($\text{HM}_{\text{INRIM}}$) by using a fiber optical frequency comb. The control-loop is implemented via software, as the ultrastable laser and the optical comb are in two different laboratories. A control bandwidth of 0.05 Hz is chosen as a compromise between an adequate rejection of the HM noise and a sufficiently tight locking; the correction is applied once per second on a Acousto-Optic Modulator (AOM). $\text{HM}_{\text{INRIM}}$ is in turns continuously measured against the Italian primary frequency standard $\text{ItCsF}_2$. The ultrastable laser is sent to the optical comb via a 100-m long phase-stabilized fiber. Another phase-stabilized optical link, 544 km long, delivers the laser frequency from INRIM to MR. This link is part of the backbone going to Florence and described in [3]. In Bologna, a part of the radiation is extracted and delivered to MR along a multiplexed optical fiber, where a single channel of the ITU grid is dedicated to our experiments. The total loss for this link is 150 dB. 7 bidirectional Erbium-Doped-Fiber-Amplifiers (b-EDFAs) are used along the link, and another b-EDFA is placed in MR. It is operated slightly above threshold to avoid lasing effects induced by the presence of a mirror at a short distance. To perform the phase-stabilization, a part of the signal is back-reflected from MR to INRIM and here beaten with the original signal on photodiode PD1, following the typical scheme of Doppler-stabilized links [15]. In this way, we can detect the phase noise added onto the optical carrier by the double pass in the fiber and we compensate it with a phase-locked loop (PLL) acting on a Acousto-Optic Modulator (AOM1). The beatnote signal to noise ratio (SNR) changes with time from 27 dB to 32 dB in a 100 kHz bandwidth, hence it needs to be regenerated by a tracking filter to improve the SNR and filter out spurious signals. Cycles-slips on this PLL may happen on a statistical basis; their occurrence is strongly dependent on the SNR [16] and on time-varying Rayleigh-scattering events along the link. Thus, the beatnote is split into three equal parts and tracked by three independent voltage-controlled oscillators; in this way it is always possible to determine if the cycles-slips have occurred to the in-loop tracking filter. At MR, the optical signal is beaten with a narrow-linewidth diode laser on photodiode PD2; the diode laser is then phase-locked on the incoming radiation on a bandwidth of 50 kHz and used as a reference for a fiber optical frequency comb. The 40th harmonic of the repetition rate (250 MHz) is extracted and divided by 100; the resulting 100 MHz signal is directly compared to the output of the HM there located ($\text{HM}_{\text{MR}}$).

This setup is robust and currently capable to operate continuously for extended periods of time. The most critical issue are the large polarization changes of the signal travelling the optical link. They cause slow but continuous variations of the SNR on the PD2 beatnote, and periodic polarization adjustment is required to prevent PLL unlocks. An automatic polarization-adjustment stage has been developed and
will be implemented at MR in the next month, in view of another measurement campaign. To constantly monitor the proper operation of the whole apparatus and the occurrence of cycles-slips on any of the PLLs involved, each beatnote was continuously measured and all points differing from the lock-frequency by more than a specified threshold were discarded from the HMs comparison. We performed daily calibration runs to ensure a proper synchronization of the measurements in the two laboratories. The main source of delay between the measurements is due to the internal clocks of the PC used for data recording and can achieve several seconds; this issue could be easily mitigated in the future by connecting them to Network Time Protocol (NTP) servers. Since the synchronization was limited to 1 s during this measurement session, we also discarded the points adjacent to a cycle-slip. The typical instability of the HMs comparison is shown in Fig. 3. On the averaging times of few seconds, it is limited by HM_{MR}, whose specified instability is $8 \times 10^{-14}$ on 1 Hz bandwidth. The short-term instability of the delivered optical signal, $1 \times 10^{-14}$ on 1 Hz bandwidth, is negligible on this timescales. At measurement times longer than 5 s, the instability starts to be dominated by that of the delivered optical signal (ultrastable laser+HM_{INRIM}). The intrinsic instability of the HMs is achieved on a timescale of few hours.

We performed repeated frequency measurements of HM_{MR} vs HM_{INRIM} for two weeks. During the whole period the frequency of HM_{INRIM} was measured by the Italian primary frequency standard ItCsF2. The results of the absolute calibration of HM_{MR} are shown in Fig. 4. The points have been obtained by a non-weighted average of the frequency data, after the removal of points which were affected by cycles-slips on any of the PLLs. In the last two measurements, the laser was not actively locked to HM_{INRIM} and the HMs comparison was obtained by post-processing synchronous measurements of the laser frequency performed in the two laboratories. The uncertainty was in most cases limited by the combined instability of the measurements HM_{INRIM} vs ItCsF2 ($u_{INRIM}$) and HM_{INRIM} vs optical system ($u_{MR}$). In the last two measurements, when the laser was not actively locked to HM_{INRIM}, the uncertainty was given by

$$u = \sqrt{u_{INRIM}^2 + u_{MR}^2 + (d\tau)^2}$$

where $d\tau$ is a possible systematic uncertainty associated with a frequency drift $d$ in the ultrastable laser and a delay $\tau$ between the measurements in the two laboratories [17]. In those measurements, we observed a maximum laser drift of $5 \times 10^{-16}/s$ and $10^{-15}/s$ respectively, and set $\tau = 1$ s. This places an upper limit to the actual contribution of this term. The interpolated frequency drift for HM_{MR} is $(1.5 \pm 0.1) \times 10^{-15}/$day. The obtained results are in agreement with the data obtained by GPS measurements of HM_{MR}.

### III. Discussion and Conclusions

We realized the first optical link between a National Metrology Institute and a VLBI antenna, and performed the absolute characterization of the HM there located, at the level of its statistical uncertainty. These measurements are a preliminary step towards the direct use of the optically-delivered frequency reference in VLBI observations. From the operational point of view, among the main requirements of VLBI measurements is the capability of sustaining several days of uninterrupted operation. This is feasible with the present apparatus, as the main issue are the long-term polarization changes at the remote link end. This problem will be avoided in our next measurement campaign by using an automatic polarization-adjustment stage. From a more fundamental point of view, the presence of cycles-slips might induce a loss of phase-coherence on the delivered microwave. Nonetheless, it is important to stress that typical phase-jumps are at the level of $<1$ cycle/hour in the optical domain, which means that their contribution is $<5 \times 10^{-15}$. This is negligible with respect to the typical performance of a HM.

In conclusion, the results shown in this paper demonstrate that optical links are a suitable tool to perform high-quality frequency dissemination to VLBI antennas. In perspective, a fiber-based network of multiple antennas connected to a single clock can be envisaged; in addition, optical links can provide better frequency references than HMs. This is a prerequisite to investigate the ultimate performances of VLBI and opens the door to 1 mm precision in geodetic positioning via VLBI-based space geodesy.

### Acknowledgment

The authors thank the GARR Consortium for technical help with the fibers. This work was supported by the European Metrology Research Programme (EMRP) under SIB-02 NEAT-FT and by the Italian Ministry of Education and Research under the Progetti Premiali programme. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.
REFERENCES


772
A Round-Trip Fiber-Optic Time Transfer System Using Bidirectional TDM Transmission

Guiling Wu¹, Liang Hu¹, ², Hao Zhang¹, Jianping Chen¹
¹State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
²Dipartimento di Fisica e Astronomia & LENS, Università di Firenze, INFN Sezione di Firenze, via Sansone 1, I-50019 Sesto Fiorentino (FI), Italy
wuguiling@sjtu.edu.cn

Abstract—We propose a round-trip fiber-optic time transfer scheme over single optical fiber utilizing the same wavelength in both directions. It can suppress the impact of the Rayleigh backscattering and the dispersion-induced symmetric deviation over fiber link by using bidirectional time division multiplexing (TDM) mechanism. A 200 km time transfer over single optical fiber with identical optical wavelength in both directions is demonstrated. The measured stability in terms of TDEV are less than 40ps/s and 10ps/d, respectively. The uncertainty induced by uncalibrated fiber links up to 200km is less than 27 ps, which is mainly limited by the performance of used time interval counters.

Keywords—time transfer; optical fiber; round-trip; time division multiplexing.

I. INTRODUCTION

Fiber-optic time transfer has attracted widespread research interest [1-7] because of its advantages of low loss, high reliability, wide bandwidth and high stability, and satellite-based time transfer is difficult to eliminate the effects of fluctuations and interferences induced by atmosphere [8-9]. Up to now, fiber-based time transfer mainly concentrates on two-way [1-2] and round-trip [3-7] dissemination using bidirectional wavelength division multiplexing (WDM) transmission that can suppress the impact of the Rayleigh backscattering by using WDM filter. The bidirectional WDM-based scheme, however, will cause bidirectional propagation delay mismatch induced by the chromatic dispersion, which will degrade the precision and accuracy of time transfer. Although some calibration methods can be used to measure the propagation delay mismatch, the uncertainty and difficulty to calibrate the mismatch will increase with the increase of fiber link length since the effect of the measurement errors of the dispersion coefficient and wavelength will increase linearly with the increase of fiber length.

We have proposed a bidirectional time division multiplexing (TDM) based two-way time transfer scheme [1], which can suppress the effects of the Rayleigh backscattering and keeps the bidirectional symmetry of fiber propagation delay at the same time. In this paper, we propose a bidirectional TDM based round-trip time transfer scheme over signal fiber with same wavelength to disseminate time signals from reference clock directly to user with high-precision.

II. SYSTEM CONFIGURATION

Figure 1 illustrates the proposed round-trip time transfer system based on bidirectional TDM transmission schematically. The time signal (one-pulse-per-second, 1PPS) from the clock at site 1 is entered into the local modem. In the modem, the input time signal from clock is encoded into a time code by an encoder [7], and carried on the light with a wavelength of \( \lambda \) through an optical transmitter (E/O converter). The light from the optical transmitter is launched into the optical fiber link by switching on the optical switcher (OS) 1, which is switched off in non-transmission durations of time code. As the light carrying time code arrives at site 2, the optical signal is converted to electrical signal by an O/E converter, which is sent to decoder to extract the time signal in it. The recovered time signal is delayed by a Time Delay Adjuster (TDA) until the whole time code has been received. The delayed time signal is then encoded into a time code again and transmitted back to site 1 over the same fiber link with the same wavelength \( \lambda \) after turning on the OS 2, which is also switched off in non-transmission durations of time code. The time signal in the returned time code is extracted, and the round trip delay of fiber link is measured by TIC (time interval counter) 1. The time delay of TDA at site 2 is also measured by a TIC (TIC 2 in Fig.1).

Figure 1 (b) illustrates the timing sequence diagram for the bidirectional TDM based round-trip time transfer scheme. We have

\[
\begin{align*}
T_{LR} &= \tau_f^l + \tau_F^R + \tau_R^l \\
T_{RL} &= \tau_f^R + \tau_F^R + \tau_R^l \\
T_{LL} &= T_{LR} + T_{RL} + T_A
\end{align*}
\]  

where, \( T_{LR} \) (\( T_{RL} \)) is the one-way propagation delay from site 1(2) to site 2(1); \( T_{LL} \) is the round trip delay measured by TIC 1 at site 1; \( \tau_f^l\) (\( \tau_f^R \)) is the propagation delay of optical fiber link from site 1 (site 2) to site 2 (site 1); \( \tau_F^R\) (\( \tau_F^R \)) is the time delay from the input of encoder to the optical output port in site 1 (site 2); \( \tau_R^l\) (\( \tau_R^l \)) is the time delay from the optical...
input port to the output of decoder in site 1 (site 2); \( T_A \) is the time delay of the TDA measured by TIC 2 at site 2.

From Eq (1), one can obtain:

\[
T_{LR} = \frac{1}{2}[T_{LL} - T_A - T_{RL} + T_{LR}]
\]

\[
= \frac{1}{2}[T_{LL} - T_A + (\tau_{RL}^R - \tau_{RL}^B) + (\tau_{LR}^R - \tau_{LR}^B - \tau_{LR}^L)]
\]

Since the two directions employ identical wavelength in the same fiber, \( \tau_{RL}^R \) and \( \tau_{RL}^B \) can be considered equal in slowly varying circumstance as is the usual case. The asymmetry of sending and receiving delay between two sites, \( (\tau_{LR}^R - \tau_{LR}^B + \tau_{LR}^L - \tau_{LR}^R) \), can be calibrated through high-precision measurements in electric and/or optical domain since they only are related to the electric/optical cables and devices at local and remote ends. Therefore, the one-way propagation delay from site 1 to site 2 can be calculated by the measured \( T_{LL} \) and \( T_A \), and the time transfer from site 1 to site 2 can be implemented.

In the proposed scheme, the optical carriers at two sites are launched into fiber only during the transmission of time code, and are switched off during the non-transmission of time code. In such a way, although the two directions employ the same wavelength over a single fiber to achieve the maximum bidirectional symmetry of fiber propagation delay, the effect of single Rayleigh backscattering originating from local light sources can still be completely eliminated. Moreover, the effect of double Rayleigh backscattering from the transmitting light can also be suppressed partly since only the double backscattering generated in the duration of the transmitting time signal can overlap with it in time domain.

III. EXPERIMENTAL RESULTS AND DISCUSSION

An experimental system based on the proposed scheme in Fig. 1 is setup. In the system, the 1PPS from a Rb clock (Symmetricom, 8040C) is sent from site 1 to site 2. The dedicated codecs with synchronization function proposed in [7] are used to encode/decode the time signal into/from a time code. Two DWDM small form-factor pluggable (SFP) transceivers with the same wavelength are used to send and receive time codes with a bit rate of 1Mb/s. The optical switch used at each site has a switching time of several milliseconds. The TDA at site 2 is implemented in FPGA. Two SR620s
(Stanford Research System) are adopted to measure the round trip delay and the TDA delay, respectively. The one-way propagation delay from site 1 to site 2 is also directly measured by a SR620 (TIC 3 in Fig.1) for the evaluation of the system.

Figure 2 shows the adopted time code format. Its frame length is 200 bits corresponding to 200μs for 1Mb/s. The pulse-width coding rule is the same as the Inter-Range Instrumentation Group time code B (IRIG-B) [8], where binary “1”, “0”, and “P” are represented by the pulses with duration of 50%, 20% and 80%, respectively. The first 100 bits in the time frame are used to carry the reference mark, the time and control information. The reference mark consists of two consecutive “P” codes where the leading edge of the second “P” code indicates the on-time of 1 PPS. The following two 50 bits are used to carry the measured round trip delay at site 1, and the measured delay of the TDA, respectively.

Figure 3 shows the measured wavelength of the two adopted SFPs in an air-conditioned room in 24 hours using an optical wavelength meter with an accuracy of 0.3pm (YOKOGAWA, AQ6151). We can see that the wavelength difference of the two SFPs is always less than 1pm when the temperature fluctuates between 22°C and 25°C. The wavelength difference of 0.5pm can be reached when the temperature fluctuation is controlled in 1.5°C. That indicates a bidirectional propagation delay asymmetry of less than 8.5 ps for a 1000 km fiber link with a typical dispersion coefficient of 17 ps/nm/km.

The performance of time transfer is evaluated by the difference, ΔΤ, between the one-way propagation delay obtained by proposed scheme and the one directly measured by TIC 3 in Fig.1. Fig.4 shows the TDEVs of bidirectional TDM based round-trip time transfer over optical fibers of 100 km and 200 km, respectively. To compensate the fiber loss, a special-designed single fiber bidirectional amplifier (SFBA), see the inset in Fig. 1, is added after about 100 km transmission in the time transfer over 200 km fiber. From the figure, we can see that the stabilities of the time transfer in both cases are better than 40ps/s, and 10ps/d, respectively. We can also see that the inserted SFBA in 200km fiber-optic link has no significant effect on the stability of time transfer.

Figure 5 shows the averages and standard deviations (Std) of ΔΤ in one hour over different lengths of fiber. In the experiment, the measured ΔΤ over 2 m fiber, 1.919 ns can be considered as the asymmetry of sending and receiving delay between two sites, (τ_P^+ - τ_P^- + τ_B^+ - τ_B^-). We can see that the average of the ΔΤ almost keeps constant with the increase of fiber length from 2 m to 200 km. The maximum fluctuation of the average of ΔΤ with fiber length is less than 27 ps. Since the resolution of SR620 is about 25ps, the measured fluctuation of ΔΤ is mainly come from SR620.

The above results indicate that the proposed scheme can achieve a time transfer over several hundred (even thousand) kilometers optical fiber with a precision of less than 100ps without any calibration of fiber links. It is worth to note that based on the point-to-point scheme shown in Fig.1, a
point-to-multipoint fiber-optic time transfer scheme based on bidirectional TDM transmission can also be realized to perform high-precision distributed time dissemination, see Fig.6.

![Fig.6. The scheme of distributed round-trip fiber optic time transfer based on bidirectional TDM transmission.](image)

IV. CONCLUSIONS

We propose a high-precision bidirectional TDM based round-trip fiber-optic time transfer scheme, which can suppress the impact of the Rayleigh backscattering and reaches the highest bidirectional symmetry of fiber propagation delays at the same time, to disseminate time signals from reference clock directly to users. The principle of the scheme is experimentally validated over 200 km optical fiber with a SFBA. The stability of less than 40ps/s is reached over 200km optical fiber link. The measured wavelength difference shows that the uncertainty induced by 200 km uncalibrated fiber can be less than 2 ps. The uncertainty induced by 200 km uncalibrated fiber measured in the experiment is less than 27 ps, which is mainly limited by the uncertainty of used TICs. A point-to-multipoint fiber-optic time transfer scheme based on bidirectional TDM transmission is also presented. Our next work includes to demonstrate the fiber-optic time transfer over more than 1000 kilometer and the distributed fiber-optic time dissemination by using the proposed scheme.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (61127016 and 61107041), SRFDP of MOE (Grant No. 20130073130005).

REFERENCES


Abstract—There have been several investigations [1], [2], [3], that demonstrated benefits of adding dopants such as (Sc) or combination of other materials, like Zr/Mg for example, to the aluminum nitride (AlN) films in order to increase coupling coefficient (kt^2) of the Bulk Acoustic Wave (BAW) devices. For concentrations below 10% atomic Sc, it is possible to use a single composite target with a standard magnetron design [4]. Most R&D systems that performed initial investigations on AlScN films with high concentration of Sc dopant, used two separate targets with two separate magnetrons: one with pure Al and one with pure Sc with different applying power to compensate for the large difference in sputtering rates of the two materials and get stoichiometric composition. Unfortunately, depositing from two different targets is only viable for low volume R&D experiments. The system described in this article uses standard dual conical magnetron with AC deposition source. Targets are cut into multiple segments as shown in Figure 1 [5].

Figure 1. Multiple material targets

Based on simple geometry of target’s surface, deposited film composition is proportional to the surface of specific pieces of target material. Unfortunately, Al is eroded at much higher rate than Sc at the same potential and same magnetic field. Over the target life, concentration of Sc increases in the deposited films. In order to maintain same Sc composition over the entire target life, it is necessary to vary magnetic field locally over the surface of the Al and Sc pieces to provide same erosion rate of Al vs. Sc at the same target potential. Adjusting magnetic field for each segment of both Al and Sc allows for constant deposited film composition over the entire target life solves this problem.

I. INTRODUCTION

Most PVD deposition processes that require multiple materials in a sputtering target use a composite target. Although it works well for most materials, some composite targets are extremely difficult to manufacture or even impossible. Also, they are very expensive. Binary Al/Sc targets are notoriously difficult to manufacture at Sc compositions above 10%. Theoretical maximum Sc concentration in aluminum with stable FCC crystal structure is about 22%, see Al-Sc phase diagram on figure 2 and table 1 and 2. However, the crystallographic texture should be remained relatively stable, retaining the preferred orientation ratios for good uniformity and stoichiometry during sputtering. Orientation ratios, compared to grain size, are the more critical parameter for sputtered film properties.

The AlSc phase diagram below shows the Al-Al3Sc eutectic reaction at very high temperature [6]. This process causes hot cracking of composite material, which were observed as black spots even below 10% of scandium composition in aluminum target. Higher Sc composition in Al target produces even worse problems.
Composite Targets with three different materials, such as ternary compounds Al-Er-Mg or Al-Zr-Mg, are even more difficult to make. For example, Al-Er-Mg alloy may show multiple intermediate phases [7], [8], which results unstable properties of the deposited film.

On other hand, sputtering from multiple targets from different magnetrons is not practical for high volume production due to low deposition rate and poor film thickness uniformity.

In this investigation we used multiple piece targets to produce highly piezoelectric films with Al, Sc, Mg and Er materials.

II. EQUIPMENT

In this investigation we used Advanced Modular Systems cluster tool with AlN deposition chambers and ion beam trimming module (shown in Figure 4).

AlN deposition uses a dual magnetron system with positive plasma column and with AC power applied between targets. Frequency of AC power is 40 kHz and power may vary from 3 to 10 kW. It is a reactive deposition process in deep poison mode using targets composed of Al and Sc pieces (tiles), see figure 5.
Figure 5. Multiple material targets

High purity research grade 99.9999% argon and nitrogen process gasses we used for all depositions.

Substrate rotation is using to compensate variation of scattering for different materials and composition non-uniformity across the substrate.

The trimming module uses DC focused ion source with argon process gas to improve thickness/uniformity of deposited films. Film thickness trimming/tuning is processing based on ion beam scanning across a wafer with power variation based on film thickness map. Use of the trimming process opens up a much wider process window for stress and composition control, because it allows avoiding of spending too much effort on controlling thickness uniformity during deposition.

III. FILM COMPOSITION CONTROL

The system uses standard dual conical magnetron with AC deposition source. Targets are cut into multiple pieces/tiles as shown in Figure 5. Films containing various concentrations Scandium (Sc) and Erbium (Er) in AlSc(x)Er(y)N films, Scandium and Magnesium in AlSc(x)Mg(y)N films or Scandium in Al(x)Sc(y)N films have been demonstrated using different number of Sc and/or Er, or Sc and Mg pieces compare to the number of the Al pieces. As sputtering area of specific material increases, composition of this material in deposited film also becomes higher.

But not only pieces/tiles sputtering area effects on film composition. Sputtering rate of specific material has a very significant effect on film composition. Typically, for example, concentration of Sc in the AlScN film is about 66% of what it is in the deposition target with either composite targets or targets made out of pieces. It is much more obvious on the targets made out pieces of Al and Sc, that erosion of Al piece is significantly higher than the erosion of the adjacent Sc piece, if the same magnetic field is used. Sputtering rate for different material vs. magnetron magnetic field is shown below on figure 6.

It is clear, that for Al-Sc deposition, for example, a constant magnetic field across of the magnetron will produce lower scandium composition film on the wafer than on the target if standard AlN magnetron is using. Changing the local magnetic field on the surface of tiles/segments with different sputtered materials results higher or lower concentration of desire dopant and a uniform target race track. We found that if we produce the same erosion of the Al and Sc targets, we get the same composition of Sc on the wafers as on the target. Based on Sputtering rate vs Magnetic field, we calculated compositions of different dopant’s concentration in AlN film depends on concentration of the same dopant in the target.

For Al-Sc deposition, increasing magnetic field on the surface of the aluminum pieces, increases Sc composition on the deposited film on the substrate, but it produces opposite result for Al-Mg deposition, as shown in figure 7.

Some mismatch between calculated and actual film/target composition ratio versus magnetic field could be considered as an error for different scattering shape for different type of dopant (heavier atoms, normally at given target voltage, have broader scattering and some material is losing due to deposition on shields. Lighter atoms, at given target voltage,
have, usually, more narrow scattering and more material reaches a substrate). Also, we didn’t consider smooth variation of magnetic field from tile/piece to tile/piece. Since magnetic field has continuous (smooth) function, and we can’t produce it as a step function.

Film’s compositions were measured by Rutherford Backscattering Spectrometry. Rutherford Backscattering Spectrometry spectra are acquired at a backscattering angle of 160° and an appropriate grazing angle (with the sample oriented perpendicular to the incident ion beam). The schematic diagram below shows the scattering geometry in a typical RBS experiments.

Figure 8. Scattering Geometry of an Rutherford Backscattering Spectrometry Experiments

Uniformity of sputtering rate changes over the target life. Fortunately Sc or Mg concentration across wafer and over the target life doesn’t change appreciably. This allows for a robust production process that can easily be fixed by ion beam trimming. Figures 9, 10 and 11 below, show AlScN deposited film properties over the target life, including thickness uniformity with parallel trimming process, with adjusted local magnetic field for Al and Sc tiles/pieces.

Figure 9. Film stress of AlScN deposition as a function of target life

Figure 10. Film uniformity of AlScN deposition as a function of target life

Figure 11. Film properties of AlScN deposition as a function of target life

IV. SUMMARY

Using custom and local adjusted magnetic field and thickness trimming, we were able to demonstrate production worthy highly doped AlN deposition process.
REFERENCES


781
Observation of Strong Temperature Hysteresis in Molybdenum Disulfide (MoS$_2$) Vibrating Nanomechanical Resonators

Zenghui Wang*, Rui Yang, Arnob Islam, Philip X.-L. Feng*
Department of Electrical Engineering and Computer Science, Case School of Engineering
Case Western Reserve University, Cleveland, OH 44106, USA
*Email: zenghui.wang@case.edu, philip.feng@case.edu

Abstract—We report experimental investigation of resonant responses of molybdenum disulfide (MoS$_2$) nanomechanical resonators at different temperatures. We observe strong temperature hysteresis in measurements. By examining devices with different geometries under different air pressures, we determine that surface adsorption plays an important role in the observed temperature hysteresis. This opens new possibilities for studying surface processes using 2D nanomechanical resonators.

Keywords—two-dimensional nanoelectromechanical systems (2D NEMS); resonator; temperature coefficient of frequency (TC$_f$); adsorption; hysteresis; molybdenum disulfide (MoS$_2$)

I. INTRODUCTION

Temperature-related hysteresis is an important or critical phenomenon that can affect the performance of resonators and oscillators based on microelectromechanical systems (MEMS). For timing applications the hysteresis is highly undesirable, and can be largely mitigated through techniques such as encapsulation [1]. Resonant nanoelectromechanical systems (NEMS) based on atomic-layer two-dimensional (2D) crystals have demonstrated interesting properties such as high tunability of frequency [2] and very broad dynamic range [3]. However, their temperature-dependent resonant responses, such as temperature coefficient of frequency (TC$_f$) and possible temperature hysteresis, remain to be characterized. In this work, we report observations of a strong temperature hysteresis in molybdenum disulfide (MoS$_2$) diaphragm nanomechanical resonators, with new features that differ from those known in mainstream MEMS devices.

II. EXPERIMENTAL TECHNIQUES

A. Device Fabrication

To fabricate MoS$_2$ nanomechanical resonators with the structures and structural variations uniquely suited for this study, we perform mechanical exfoliation of MoS$_2$ nanosheets onto prefabricated device structures [4]. First, we define circular microtrenches of different sizes (0.5–5μm diameter) on a silicon (Si) wafer coated with 290 nm of thermal oxide (SiO$_2$), using photolithography followed by buffered oxide etch (BOE). We carefully choose the etch time such that the flat Si surface is exposed at the bottom of the trench, which helps improve reflected light intensity and thus the interferometric motion transduction. We then exfoliate MoS$_2$ nanosheets onto this structured substrate using scotch tape method. We notice that this method results in various types of suspended MoS$_2$ devices, including those with the MoS$_2$ crystals fully covering the circular microtrenches and those partially covering the microtrenches. Some larger crystalline flakes can even cover multiple trenches, resulting in several devices with the same MoS$_2$ thickness but with different types of MoS$_2$ coverage (Fig. 1). Suspended MoS$_2$ devices are then identified under an optical microscope (Olympus MX50) with a 50× objective. We then use an AFM (Agilent N9610A) in tapping mode to measure the device thickness.

Fig. 1. Illustrations of the device fabrication process and typical devices.

B. Resonance Detection

We use a fully-optical actuation/detection scheme for the experiments in this study [5]. We use an intensity-modulated 405nm blue diode laser to optothermally excite the vibrational motion, and detect device resonance interferometrically using a 633nm He-Ne red laser. As shown in Fig. 2, we use the AC driving signal supplied by a network analyzer to modulate the 405nm blue laser. The DC bias of the laser modulator is set at ~25%, corresponding to an output power of 3.33mW into the microscope, and 100% modulation (amplitude of AC signal equals to DC level) is achieved by adjusting the modulation depth (the electrical driving signal amplitude is fixed at...
200mV, which is optimized for the laser modulation input. The resonance of the MoS2 device is detected interferometrically with the 633nm red laser focused on the MoS2 flake surface (∼1mW into the microscope, corresponding to ∼300μW on device). During the sweeping of the driving frequency, the output signal measured by a photodetector (New Focus 1801) is recorded using the same network analyzer, allowing identification of nanomechanical resonances in the devices. We finally fit the data to a damped simple harmonic oscillator, from which \( f_{\text{res}} \) and \( Q \) are carefully extracted (Fig. 2b). Prior to each measurement, the effect of laser heating (from both the driving 405nm laser and the probing 633nm laser) are carefully calibrated, and we have found that for the laser power levels we used the heating effect does not affect the conclusion from the TC\( f \) measurements [6].

![Image](image1.png)

**Fig. 2.** Resonance measurement. (a) Schematic of measurement system. (b) An example of measured resonance data with fitting (dashed red curve).

C. Measurement of Temp. Coefficient of Frequency (TC\( f \))

The samples are mounted on a custom-built heating stage composed of a Peltier heater and a large heat sink (Fig. 3). To switch from heating to cooling, we simply reverse the connection on the Peltier heater so the heat flow reverses the direction. The heat sink is important in maintaining the sample temperature, as it stabilizes the temperature on one side of the Peltier device, thus the temperature on the sample side is only determined by the temperature difference across the Peltier device, which is controlled by the output of the power supply.

![Image](image2.png)

**Fig. 3.** The sample temperature-regulation assembly: the actual picture and the circuit schematic. Electrical wirings are shown in the schematic.

To measure the temperature of the sample, we employ a diode temperature sensor (Lake Shore DT-670-SD). We pass a constant 10μA current through the device, and monitor its voltage drop which is then converted to temperature reading through a standard Si diode temperature calibration curve.

The TC\( f \) measurement is performed in the following procedure: (1). Prior to the measurement, we optimize the resonance detection conditions, and mark the position of the MoS2 flake under the microscope. We then start automatic, continuous measurement of resonance frequency. (2). At each new temperature setting (controlled by the power output to the Peltier heater), we wait sufficiently long time till the temperature reading stabilizes. (3). We then re-adjust the focus and sample location, to keep the relative position between the sample and the microscope objective constant throughout the experiment. This compensates for any possible heating related change of the focus and sample position under the microscope. (4). We then wait till the resonance frequency stabilizes (by monitoring the real-time frequency response curve on both the network analyzer and the computer screen); we note down the measurement run number (an automatically recorded and incremented number used to identify each individual frequency sweep), before changing the temperature to the next set point. The temperature reading is automatically recorded and attached to the data file of each run.

III. RESULTS AND DISCUSSION

The measurement of a fully-sealed 5μm-diameter, 56nm-thick MoS2 diaphragm resonator reveals that the device has very large TC\( f \) (>4000ppm/°C). More strikingly, it clearly exhibits strong temperature hysteresis (Fig. 4): As temperature \( T \) increases, initially resonance frequency \( f_{\text{res}} \) rapidly decreases (see the red curves in Fig. 4), exemplifying the very large TC\( f \). However, as \( T \) further increases, the \( f_{\text{res}} \) shift significantly slows down. During cooling, \( f_{\text{res}} \) follows a very different path (see the green curves in Fig. 4).

![Image](image3.png)

**Fig. 4.** TC\( f \) measurement of a fully-sealed device. (a) Optical image of the MoS2 crystalline flake. The device being measured is indicated by the dashed circle. (b) AFM measurement of the device thickness. Inset: AFM image. Black line in the image indicates where the height trace is taken. (c) A set of measured resonances as the device temperature is swept up and down.
Upon measuring the device for multiple times, we have found that such behavior is reliably reproducible and does not depend on the step size in temperature sweep (Fig. 5).

![Graph](image-url)

Fig. 5. $f_{res}$ as a function of $T$ for several cycles of temperature sweep for the device shown in Fig. 4. Arrows indicate the sweep direction and sequence.

This initially observed hysteresis exhibits several attributes that are in contrast to typical temperature hysteresis observed in conventional MEMS resonators [7]: (i) it is much stronger (~15% near 50°C vs. ~1–100 ppm levels in typical MEMS devices); (ii) it is strongly $T$-dependent; (iii) the $f_{res}$ shift precedes $T$-change (vs. that $f_{res}$ always lags behind $T$ change in hysteresis due to thermal lag). These phenomena suggest that the observed temperature hysteresis may be dominated by a different mechanism than those commonly found in MEMS and quartz oscillators [7,8], such as thermal lag, changes in strain or the crystal, and circuit hysteresis.

Compared with conventional devices, 2D resonators have very high surface-to-volume ratios (~$10^6$ m$^{-1}$), which can make their $f_{res}$ more sensitive to surface adsorption. Thus adsorption may play an important role in the observed strong hysteresis in $f_{res}$ with varying temperature. To further investigate the origin of the observed strong hysteresis, we perform additional experiments. We first vary the chamber pressure. The reasoning is that physisorption of a given atomic or molecular species on a given surface is largely dominated by both temperature and pressure [9]. Therefore, if the device temperature is varied under different background air pressures, one expects to observe different responses.

A. Pressure Dependence

We vary the chamber pressure by partially venting the sample chamber while monitoring the vacuum gauge. Fig. 6a and 6b show the measured temperature hysteresis of the same device shown in Fig. 5 under different pressures spanning ~2 orders of magnitude. The data show significant hysteresis over the entire pressure range, but the difference across different data sets (under different pressures) is small.

At first glance, the lack of pressure dependence might suggest that physisorption should not play a dominant role in the observed hysteresis. However, upon carefully examining the device structure as well as performance of similar devices in the literature, we discover that one could not exclude the possibility that the air volume trapped in the microtrench underneath the MoS$_2$ crystal during device fabrication might never escape the cavity, and maintains ~1 atm pressure on the bottom side of the MoS$_2$ diaphragm. In fact, previous study shows that it is possible for the MoS$_2$ crystal to completely seal the underneath microcavity with trapped air, and such device can exhibit robust nanomechanical resonance over a large pressure range [10]. This trapped volume of air remains ~1 atm pressure (largely insensitive to the chamber pressure), and thus could lead to significant adsorption and desorption processes during temperature sweeps, regardless of chamber pressure.

To exclude the adsorption effect from the trapped air volume, we explore the device geometry degrees of freedom by performing additional experiments using MoS$_2$ resonators with different structural features.

B. Device Geometry Dependence

We perform TC$T/f$ measurements on a partially-sealed device, in which the MoS$_2$ crystal only partially covers the microtrench, and thus exposing the volume underneath to the outside. Previous experiments show that such "nonideal" device geometries can exhibit the same robust nanomechanical resonances as geometrically 'regular' (e.g., fully-sealed diaphragm) devices, and in addition may give rise to interesting multimode resonance responses unavailable in the 'regular' structures [11]. However, to date no temperature dependent measurement has been performed across different device geometry types yet.

We choose a partially sealed device on the exact same MoS$_2$ flake as the fully-sealed device in Fig. 6a & 6b, in order to exclude any possible effect from variation in device thickness and crystal quality, and thus focus on the difference in device geometry. We perform the measurements in both medium vacuum (10 mTorr) and elevated pressure (90 Torr). The data clearly show that this device exhibit little hysteresis when under vacuum (Fig. 6c), while measurable hysteresis can be introduced as the chamber pressure increases (Fig. 6d). This is in striking contrast with the fully-sealed device (Fig. 6a & 6b), and suggest that physisorption plays an important role in the observed temperature hysteresis in this device.

C. Discussions

The observed temperature hysteresis in both devices types, though exhibiting apparently different pressure dependences, shows that the higher the vapor pressure actually experienced by device surface, the greater the temperature hysteresis.

Physisorption is known to give rise to hysteretic behavior. The most commonly known type is adsorption on porous solids, which is caused by the cavitations of adsorption phase in the pores [12]. In contrast, adsorption hysteresis on atomically-flat surfaces, such as exfoliated graphite, exhibit hysteresis due to different physical origins such as capillary condensation and phase transitions [13-15]. To identify the origin of the observed hysteresis in the MoS$_2$ resonators, new experiments (such as isotherm measurements, i.e., fixing temperature and sweeping pressure) is required in addition to the isobar (fixing pressure while sweeping temperature) measurements presented in this digest paper.
In summary, through examining temperature-dependent resonance responses in NEMS devices with different structures under different pressures, we observe strong temperature hysteresis in NEMS resonators based on MoS$_2$ crystals, which appears to be dominated by surface adsorption. The observed temperature hysteresis can be pronounced, given the fact that most NEMS structures based on 2D layered materials are made through either mechanical exfoliation [4] or transferring [16] onto pre-defined structures, and thus can trap and seal air volume underneath the 2D crystalline flake. Such strong temperature hysteresis, nevertheless, may be exploited for engineering temperature-programmed surface processes such as surface diffusion [17] and phase transitions [18] in adsorbed atoms or molecules, for which these NEMS resonators based on 2D layered materials may serve as an ideal platform.

IV. CONCLUSIONS

ACKNOWLEDGMENT

We thank J. Lee, K. He, J. Shan for helpful discussions and technical support. We thank support from Case School of Engineering, National Academy of Engineering (NAE) Grainger Foundation Frontier of Engineering (FOE) Award (FOE 2013-005), the CWRU Provost’s ACES+ Advance Opportunity Award, the National Science Foundation CAREER Award (ECCS-1454570), and the CSC Fellowship (No. 2011625071). Part of the device fabrication was performed at the Cornell NanoScale Science and Technology Facility (CNF), a member of the National Nanotechnology Infrastructure Network (NNIN), supported by the National Science Foundation (Grant ECCS-0335765).

REFERENCES

New Capacitive Micro-Acoustic Resonators Machined in Single-Crystal Silicon Stacked Structures

Nesrine Belkadi, Thomas Baron, Bernard Dulmet
Time and Frequency department
FEMTO-ST Institute
Besançon, France
nesrine.belkadi@femto-st.fr

Laurent Robert, Etienne Herth, Florent Bernard
Micro Nano Sciences & Systems department
FEMTO-ST Institute
Besançon, France

Abstract—We present a new type of acoustic resonator technology aimed to undoing the technological locks encountered during the realization of capacitive silicon MEMS resonators exploiting true Bulk Acoustic Wave resonances instead of structural ones. The single-crystal silicon resonators are driven through a combination of a static bias and dynamic voltage applied across a 700 nm-thick electrostatic gap parallel to the surface of the resonator substrate. The electrostatic actuation is realized by a local over-thickening of the gold layers used in the gold-gold bonding process of a resonant plate from resistive silicon with an external electrode-support structure, also made from resistive silicon. Thus, the electrostatic actuation is actually applied through existing interconnections between the existing conductive layers. This technology aims to the realization of very thin gaps on surfaces in the range of a few square millimeters while avoiding the shallow machining in the silicon that would require an additional technological step. We detail this technology through the process steps and designs implemented during the so-called ORSEPEE R&T project from CNES. Finally, the first characterization results are provided.

Keywords—Capacitive micro-acoustic resonator; electrostatic actuation; MEMS; longitudinal wave;

I. INTRODUCTION

It is now widely accepted that MEMS resonators play an increasing role in new architectures for RF communication systems [1]. Membranes designed were widely used together in various engineering applications such as the design of MEMS [2], [3], packaging [4], transducers [5] and other acoustical applications (e.g. BAW and SAW) [6], [7]. Development of high quality factor silicon-based capacitive micro-resonators can have a great impact on the future of wireless communication systems by providing unprecedented levels of power consumption and system integration. Indeed, this design can be manufactured with high parallelism, high yield requiring a small area, and further achieving compatibility with standard integrated circuit processes. Most of these new MEMS resonators are fabricated from silicon due to its crystallographic structure which enables the propagation of the acoustic wave with a sufficient level of quality to fulfill the requirements of the targeted applications [8], [9]. Some efforts have been done in order to obtain bulk resonators in the CMOS-MEMS approach [10]. In this study we present a new technological framework aimed to provide CMOS compatible MEMS resonators using thickness-extensional modes with higher yield and higher resonant frequencies. It is known that very small gaps are needed to drive silicon MEMS resonators with acceptable electromechanical coupling factors. Furthermore, obtaining sub-micronic gaps on the typical square-millimeters range of substrate surfaces needed to host elastic waves allowing an efficient energy-trapping mechanism is an issue for the development of true Capacitive-Micro Acoustic Resonators (C-MAR). In addition, controlling stiction with extended gap surfaces can be expected to represent a major problem in that case. In this paper, we report on the fabrication process method and characterization of C-MAR membrane-type resonators using electrostatic actuation.

II. MOTIVATION AND PRINCIPLE OF NEW RESONATOR DESIGN

Here-presented new technology of C-MAR structures is intended to improve a previous design of electrostatically-actuated bulk acoustic wave resonators presented in [11]. It was proven that thickness-extensional modes can be excited by single-side electrostatic actuation and that they can be localized in the center of plate by means of removing an aluminum layer in this region (cf. Fig.1). Using this technique equivalent thinning of the silicon plate was motivated by the negative value of the dispersion constant of thickness-extensional waves at small lateral wave-numbers.

Fig. 1. Schematic view of previous capacitive BAW resonators in Silicon-Glass stacked technology with electrical vias made by ultrasonic machining.
Then, typical Q-factors of 8,000 were obtained in air at 10.3 MHz on resonators from doped conductive silicon assembled with a glass wafer supporting the external electrode needed for the electrostatic actuation. We can turn the expression of the squared electromechanical coupling factor in the following form:

\[
k^2 = \frac{\pi^2}{2m\omega_0^2 g} \left[ \int_S u_n \, dS \right]^2
\]

where \( m \) is the mass of the vibrating part of the device, \( \omega_0 \) is the angular frequency of the considered mode, \( u_n \) is the out-of-plane component of the mechanical displacement, \( S \) is the acoustically active surface of the device, \( g \) is the effective averaged value of the static gap and \( E_e \) is the maximum allowed value of the electrical field, taken at the smaller value of the disruptive value and of the pull-in threshold. This expression indicates significantly smaller values of \( k^2 \) for thickness-extensional modes than for flexural modes. Exploiting such values in an oscillator loop was impeded by parasitic resistances attributed to the ultrasonic machining technology used to realize the electrical vias in the glass wafer supporting the external electrode. In addition, ultrasonic machining strongly restricted the possibility of further down-scaling of the devices. Consequently, we decided to start a new technology where both wafers of the stack are made from either conductive or resistive silicon. Simultaneously, the Glass-Silicon anodic bonding is substituted by a Gold-Gold bonding which eliminates the need to apply a high voltage during the bonding process, which put severe design constraints on the first generation of C-MAR resonators. At the same time, Gold-Gold bonding is expected to allow as thin or even thinner gaps as the RIE technique previously used to etch the electrostatic gap in the glass wafer.

III. DEVICE DESCRIPTION AND FABRICATION PROCESS

The new C-MAR devices were fabricated by combining two generic technologies: bulk micromachining, in which silicon (Si) was thinned, to define the vibrating area of the resonator; and surface micromachining, in which the conductive materials are added in order to structure the spacer used to define the electrostatic gap and the electrodes necessary to its excitation. The manufacturing process requires the use of seven masks in total. The starting material for technological realization is n-type silicon single-crystal oriented (100) with a resistivity of 1–10 \( \Omega \)cm. The wafers have an optical polish on both faces, since one face is a boundary of the resonator and the other one must start from a high-quality surface to minimize the final roughness of the etched surface of the membrane.

Fig. 2 shows the device architecture of the MEMS resonators which consists of two silicon plates called active part and base, which will support the upper and the lower electrode respectively. After completion of the process and before sealing the two wafers with the gold-gold bonding technique, two square vias placed over the end of the resonator are etched to allow the electrical access to the electrostatic gap by means of the two electrodes.

A. Opening masking oxide step by wet etching

Bulk micromachining was the first of the two technologies to be explored. We chose to realize the membranes by wet etching because of lower fabrication costs in comparison with dry etching. Nevertheless, membranes obtained by DRIE were realized for the purpose to compare the surface roughness since this parameter can significantly affect the electromechanical coupling of the device. Further on in this paper, Fig. 10-a and 10-b illustrate the steps required to prepare the substrate for the etching of the membranes and of the upper and lower contact areas. In order to electrically isolate the resonator from its substrate, thermal oxide was grown all over the substrate at 1100°C, the oxide measured thickness by ellipsometer equipment indicates 1.4 \( \mu \)m of SiO\(_2\). This oxide is locally removed in BHF bath after photolithography step. Square and rectangular openings are thus created, acting as a mask for the silicon etching step.

B. Membranes Etching SiO\(_2\)/Si

Once the etching of the masking oxide performed, a first top silicon plate 500 \( \mu \)m-thick was patterned to achieve a resonator body consisting of a 150 \( \mu \)m-thick membrane. Additionally, the two square vias are etched at the same time in order to obtain the growth initiation sites vias needed for the resumption of contact and thus reduce the dry etching time (as Fig. 3-c). Two silicon etching techniques widely used in microelectronics without barrier layer are implemented to ensure this step, using wet anisotropic silicon etchants and dry etchants. Although Silicon is known to be a generally inert material to chemical attack due to the native oxide protective layer present at its surface, sets of fluorinated (usually based on hydrofluoric acid coupled with oxidative species) or alkaline solutions (based on potassium hydroxide) are suitable to etch it [12], [13], [14].

For the sake of simplicity, we chose to use the alkaline etching process. Then, a solution of 10 mol/l heated at 70°C to improve silicon etching kinetics in the alkaline solution was used to produce the 150 \( \mu \)m-thick membranes. The membranes are rectangular and include a 1.4 \( \mu \)m-thick external layer of thermal oxide. The anisotropy of the silicon etching by KOH solutions induces the formation of micro-pyramids inverted in the open oxide areas. In addition, the relaxation of differential

---

DRIE: Deep Reactive Ion Etching.
TCF: Temperature Coefficient of Frequency.
stress due to the TCE difference between silicon and thermal oxide induces a bow of the membrane at room temperature. To control these effects, the remaining oxide after the etching was removed and the wafers were re-oxidized as illustrated on Fig. 3-b.

Alternatively, we investigated the realization of membranes by DRIE Bosch Process. Fig. 4 presents a comparison of the membrane profile roughness obtained with the two techniques. It is clear that the alkaline etching has a certain advantage in terms of low roughness (60 nm on Fig. 4-a, against 323 nm obtained by DRIE on Fig. 4-b).

C. Definition of Electrostatic Gap Using Spacer Technique

The metal deposition steps occur after the micro-machining. Both wafers have to be metallized. Using Ti-Ni spacer technique between resonator wafer and base wafer, 200 nm and 700 nm gaps between both wafers could be defined (cf. Fig. 10-d). The spacer metals were deposited on the full wafer by RF sputtering. A 50 nm-thick layer of Ti was covered by a 250 nm-thick layer of Ni for the resonator #1, passivated with a 100 nm-thick Al₂O₃ layer. For the resonator #2, a 650 nm layer was used in order to provide a final 700 nm gap (cf. Fig. 5). We have preferred Ni to Au for this step because of its much lower cost, its good conductivity, its better stress resistance, and its positive contribution to the robustness of the 2-wafers stack. Since the process is direct, it requires a good control of the chemical etching.

D. Excitation Electrodes

The electrodes must be made from non-oxidisable metals. For this purpose, gold remains the most suitable one, especially considering its ductility in view of the stack bonding. Then, 250 nm-thick gold layers deposited both on the active part and the base substrates are patterned to obtain symmetrical electrodes on the two banks of the gap, as shown on Figs. 10-e and 10-g. These same gold layers are used to bond the two parts in the final stack.

E. Passivation

In order to reduce the breakdown rate of the devices at high polarization values during the characterizations, the interposition of a thin layer about 100 nm of alumina deposited by evaporation between electrostatic gap and the bottom electrode becomes necessary (cf. Fig. 6).

Fig. 3. Membrane etching: (a) Top view of the wafer after removal of the oxide. (b) Top view of the wafer after thermal re-oxidation. (c) Cross-sectional illustration of the membrane.

Fig. 4. Profiles roughness of (a) KOH etching and (b) DRIE etching.

Fig. 5. Microscopic view of the spacer.

Fig. 6. Microscopic view of the passivation layer of bottom electrode.
F. Contact pads

Before the bonding of the two wafers, the electrical contacts are finally achieved by DRIE holes granting a direct access to the top conductive layer at the inner core of stack. (cf. Fig. 10-f). We used the spray-coating technique to obtain a homogeneous distribution of the photosist over the entire surface of the substrate. The 13 µm-thick layer of photosist was found sufficient to etch the relatively thin 100 µm layer of silicon subsisting after the KOH wet etching. Fig. 7 below shows the photolithography step achieving the etching of the conductive holes and the active substrate after etching.

IV. GOLD THERMOCOMPRESSION WAVER BONDING

The final structure is assembled by means of the adhesive effect of two identical metallic layers as shown in Fig. 10-h. The assembly is performed under a 4000 N force exerted at 80 °C, thereby realizing a so-called “thermo-compression process”. Gold is actually employed to perform this bonding step. By combining the gold-gold bonding and the spacer techniques between active and base substrates, we were able to define 200 nm and 700 nm-thick gaps between both silicon layers. This step requires a good state of the surface. Note that a scouring step of the substrates is necessary before sealing. Fig. 8 shows the SEM image of the 700 nm gap after bonding.

Fig. 7. DRIE etching of the electrical vias. (a) Photolithography step using spray coating technique. (b) Image of active substrate with vias.

Fig. 8. SEM Close view of the 700 nm gap after the thermocompression process.

Fig. 9. Characterization of the square membrane by Laser Doppler vibrometry a) Resonant frequencies responses and b) 3D Deflection Shape of the flexural resonance.

LVDS: Laser Doppler Vibrometer System.

V. CHARACTERIZATIONS

To obtain the frequency response, two variants membranes were characterized. They were designed to be electrostatically actuated through the two conductive vias through a transduction gap of 700 nm.

Applying an alternative voltage across the capacitive gap between the top and bottom electrodes induces the mechanical vibration associated to the out-of-plane displacement of the movable electrode. These experiments were measured using a Polytec MSA 500 LDVS where the beam was focalized at the chosen scanning area. Thus, we could control the frequency of the oscillating strain.

Fig. 9-a below shows the response of the flexural mode of the resonator with square membrane submitted to a 9V AC signal. The resonance peak is clearly visible at 471 kHz with some parasitic signal floor. The LDV scan of Fig. 9-b shows the out-of-plane displacement pattern of the square membrane at resonance.
VI. CONCLUSION

We developed a new technology to fabricate Capacitive Micro-Acoustical Resonators with a transduction gap realized by a metallic spacer layer embedded in the bond process of two silicon wafers. A first sample of such C-MAR resonator achieved with this technology was characterized with a fundamental flexural frequency at 470 kHz with a 700 nm-thick gap, although this technology allows to achieve quite thinner gap. A simple method of silicon etching without barrier layer was implemented to obtain thin SiO₂/Si membranes. The proposed design reduces the assembly steps and eliminates wire interconnections between the different levels of metallization thanks to the conductive vias. Nevertheless, further work is still necessary to optimize and to stabilize the process parameters in order to reach the
objective of transducing higher-frequency thickness-extensional modes and to develop a technology of through-silicon conductive vias filled with metal, suitable for the C-MAR class of resonators [15].

ACKNOWLEDGMENT

This work has been supported by the ORSEPEE R&T project from CNES. The authors acknowledge the clean room and characterization laboratory staff at FEMTO-ST technological facility. This work was partly supported by the French RENATECH network and its FEMTO-ST technological facility.

REFERENCES


Evaluation of Elastic Properties of SiO\textsubscript{2} Thin Films by Ultrasonic Microscopy

Kensuke Sakamoto\textsuperscript{1}, Tatsuya Omori\textsuperscript{1}, Jun-ichi Kushibiki\textsuperscript{1,2}, Satoru Matsuda\textsuperscript{3}, and Ken-ya Hashimoto\textsuperscript{1},
\textsuperscript{1}Graduate school of Engineering Chiba University, Chiba Japan \textsuperscript{2}Tohoku University, Sendai Japan \textsuperscript{3}Microdevices R&D Department, Taiyo Yuden Ltd., Akashi Japan

Abstract—This paper describes evaluation of stiffnesses of SiO\textsubscript{2} thin films when the anisotropy in elasticity is taken into account. The authors measured the propagation direction \( \theta \) dependence of the water-loaded surface acoustic waves (SAW) velocity, and tried to estimate stiffnesses of SiO\textsubscript{2} films from the measured \( \theta \) dependence. The result indicates that SiO\textsubscript{2} films possess the strong 6mm anisotropy. Namely, stiffnesses normal to the surface are significantly larger than those along the surface. This anisotropy may be induced during the deposition or caused by variation of film properties in the thickness direction.

I. INTRODUCTION

Amorphous silicon dioxide (SiO\textsubscript{2}) thin films are often used in these devices for their temperature compensation [1]–[3]. It is known that elastic properties of SiO\textsubscript{2} change dramatically with used deposition apparatus and conditions, and give serious impacts on device performances.

Matsuda et al. proposed use of the Fourier transform infrared (FT-IR) spectroscopy for assessing elastic properties of SiO\textsubscript{2}. They showed that the FT-IR spectra of SiO\textsubscript{2} films are closely related to TCF and the propagation loss [4], [5] of SAW propagating on the SiO\textsubscript{2}/LiNbO\textsubscript{3} structure. Although this technique is quite effective for this purpose, it does not give the elastic constants.

The authors measured the propagation direction \( \theta \) dependence of the water-loaded SAW velocity on SiO\textsubscript{2} thin films deposited on a Si(001) substrate by Line-Focus-Beam (LFB) ultrasonic microscopy [6], and tried to estimate stiffnesses of SiO\textsubscript{2} films from the measured \( \theta \) dependence [7]. However, agreement was poor between the measurement with the fitted result when isotropy was assumed to SiO\textsubscript{2} films.

This paper describes evaluation of stiffnesses of SiO\textsubscript{2} films when the anisotropy in elasticity is taken into account. The result indicates that SiO\textsubscript{2} films possess the strong 6mm anisotropy. Namely, stiffnesses normal to the surface are significantly larger than those along the surface. This anisotropy may be induced during deposition.

We have applied the same fitting procedure to four SiO\textsubscript{2} samples prepared by sputtering or chemical vapor deposition (CVD) with two different deposition conditions each, and the agreement was excellent for all these samples. In addition, estimated result of elastic constants shows the large variation of elasticity among four samples.

II. LINE-FOCUS-BEAM ULTRASONIC MICROSCOPY

Here brief introduction of the LFB ultrasonic microscopy used in this work is given. Detailed discussions on its use for material characterization can be found in Refs. [6], [8], [9].

Basic configuration of the LFB microscope is illustrated in Fig. 1. The acoustic lens focuses ultrasonic RF pulses excited by the transducer on vicinity of the boundary between water and the sample to be evaluated.

Then leaky-SAWs are generated and propagated to the \( x \) direction, and a portion of leaked ultrasonic components will be detected by the transducer. It should be noted that the transducer also receives ultrasonic echoes caused by the simple geometric reflection.

As a result of the interference between these two components, the output signal \( V(z) \) of the transducer varies periodically with the lens height along the \( z \) axis as shown in Fig. 2.

![Fig. 1. Basic configuration of LFB microscope [6]](image)

One can estimate the SAW velocity \( v \) from this periodicity
\[ \Delta z = \frac{V_w}{\sqrt{1 - \left(1 - \frac{V_w}{2f \Delta z}\right)^2}}, \]

where \( V_w = \sqrt{c_w/\rho_w} \) is the acoustic velocity in water, \( c_w \) and \( \rho_w \) are the mass density and elastic constant of water, and \( f \) is the carrier frequency of the RF burst. In the following experiments, \( f \) was fixed at \( f = 225 \text{ MHz} \).

Repeatability and accuracy of this measurement are mainly governed by those of the mechanical translation and temperature control. Careful system setup allows us to achieve extremely high repeatability and accuracy; measurement repeatability better than 0.01% was reported [8].

III. CHARACTERIZATION OF ELASTIC PROPERTIES OF ANISOTROPIC SUBSTRATE USING LFB MICROSCOPY

Since the LFB microscope used in this work equips a sample rotation mechanism, one can easily obtain \( v \) as a function of \( \theta \). If the sample is anisotropic, its elastic constants may be estimated by fitting a numerically evaluated SAW velocity \( v_c(\theta) \) to the measured one \( v_m(\theta) \) [10].

At first, a Si (001) bare wafer is used as a sample to check applicability of this technique. The SAW velocity on a water-loaded Si surface is governed by three independent elastic constants (\( c_{11}, c_{12}, c_{44} \)) and the mass density \( \rho \) when \( \theta \) is specified. Although the SAW velocity also depends on \( c_w \) and \( \rho_w \), they are well controlled in the present system.

Elastic constants of Si can be estimated by minimizing the sum of square error defined by

\[ \delta^2(c_{11}, c_{12}, c_{44}) = \sum_{n=1}^{N} \left( v_m(\theta_n) - v_c(\theta_n - \theta_c; c_{11}, c_{12}, c_{44}) \right)^2, \]

where \( N \) is the number of measured points, and \( \theta_c \) is a constant to compensate misalignment caused at the sample setting on the rotation stage. The global minimization of \( \delta^2(c_{11}, c_{12}, c_{44}) \) is performed by the simulated annealing (SA) [11], [12].

Note that \( c \) and \( \rho \) appear in a form of \( (c/\rho) \) in SAW velocity calculation, and they can not be determined independently in this procedure. Thus \( \rho \) given in Ref. [13] is used for the estimation.

Figure 3 shows \( v_m \) and fitted \( v_c(\theta) \) as a function of \( \theta \). Here, \( \theta = 0 \) corresponds to the \( \langle 110 \rangle \) direction. Good agreement can be seen between \( v_c(\theta) \) and \( v_m(\theta) \). Four-fold symmetry is seen. It is due to the crystallographic symmetry of the Si (001) surface.

Both \( v_c(\theta) \) and \( v_m(\theta) \) change abruptly at \( \theta \approx \pm 15^\circ \) and \( \theta \approx 75^\circ \). This is due to change of the detected wave from the Rayleigh SAW (R-SAW) to the pseudo SAW (P-SAW) or vice versa [14].

Table I shows elastic constants estimated by the fitting. They agree well with those given in Ref. [13] in reasonable accuracy.

<table>
<thead>
<tr>
<th>TABLE I. ESTIMATED ELASTIC CONSTANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stiffness [GPa]</td>
</tr>
<tr>
<td>( c_{11} )</td>
</tr>
<tr>
<td>Estimated</td>
</tr>
<tr>
<td>Reference [13]</td>
</tr>
</tbody>
</table>

IV. EVALUATION OF ELASTIC CHARACTERISTICS SiO\textsubscript{2} THIN FILM DEPOSITED ON Si (001) SUBSTRATE

The method described in the previous section is also applied for determining elastic constants of SiO\textsubscript{2} thin films deposited on a Si (001) substrate.

Amorphous SiO\textsubscript{2} thin films were deposited by CVD for samples (a) and (b) while sputtering was used for samples (c) and (d). Note different deposition conditions were applied for (a) and (b), and (c) and (d) (Table II). Thickness of SiO\textsubscript{2} films was adjusted to be 1 \( \mu \)m.

TABLE II. LIST OF PREPARED SAMPLES TO BE EVALUATED

<table>
<thead>
<tr>
<th>Sample name</th>
<th>Condition</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>#1</td>
<td>CVD</td>
</tr>
<tr>
<td>(b)</td>
<td>#2</td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>#1</td>
<td>Sputtering</td>
</tr>
<tr>
<td>(d)</td>
<td>#3</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4 shows the measured SAW velocity propagating on the SiO\textsubscript{2} thin film / Si (001) substrate structure as a function of \( \theta \). Although the film thickness is much smaller than the SAW wavelength (~ 20 \( \mu \)m), it is seen that

- SAWs propagating on the SiO\textsubscript{2}-thin film/Si-substrate structure are about 5% slower than that propagating on the bare Si substrate.
- This reduction of SAW speeds is larger for the samples deposited by sputtering compared with that deposited by CVD.
- The SAW velocity varies significantly with the depositing condition even when the same apparatus is adopted.
- The \( \theta \) dependence is significantly changed by the SiO\textsubscript{2} deposition.
- The change is noticeable especially for P-SAW.

The result indicates elastic properties of SiO\textsubscript{2} films vary significantly with the deposition technique and/or condition.

Then the elastic constants of SiO\textsubscript{2} films \( (c_{11}, c_{44}) \) were estimated. For the numerical calculation of \( v_c(\theta) \), it was assumed that: 1) the deposited amorphous SiO\textsubscript{2} films are isotropic, 2) variation of the mass density of SiO\textsubscript{2} films are small and it is approximated to the one given in Ref. [13].
Fig. 4. Phase velocities of SAWs propagating on Si wafer or SiO₂-thin film/Si-substrate structures as a function of propagation direction.

Fig. 5 show the measured and fitted results for Sample (a). In this case, agreement is poor under the comparison with Fig. 3.

Elastic constants estimated for four samples are listed in Table III. In the table, $\tilde{c}_{12}$ was calculated by $\tilde{c}_{11} - 2\tilde{c}_{44}$. We believe that estimated $\tilde{c}_{11}$ and $\tilde{c}_{44}$ include considerable errors because estimated $\tilde{c}_{12}$ was negative.

Figure 6 (a) shows the fitting result when the fitting range was restricted to $-70^\circ \sim -25^\circ$ for Sample (a). In the range, the R-SAW is responsible. Although $v_\nu(\theta)$ fits well for the limited area, the agreement is poor for the $\theta$ region where P-SAW is responsible.

Figure 6 (b) shows the result when that the range was restricted in $-16^\circ \sim +12^\circ$ where the P-SAW is responsible. In the case, $v_\nu(\theta)$ fits well only for the limited area.

Elastic constants estimated in these cases are given in Table IV. For all cases, estimated $\tilde{c}_{12}$ was negative.

![Table III](attachment:table_iii.png)

![Table IV](attachment:table_iv.png)

This fact implies that SiO₂ films possess anisotropy. Namely, elastic properties normal to the surface seem to be significantly different from those along the surface. Therefore existence of the 6mm anisotropy is assumed to SiO₂ films, and elastic constants were estimated by extending the fitting procedure described in Section III.

Fig. 7 show the fitted results when the 6mm anisotropy is assumed to the SiO₂ film for the Sample (a). In this case, the fitted result agreed well with the experiment.

Estimated elastic constants are listed in Table V. In the table, $\tilde{c}_{12}$ was calculated by $\tilde{c}_{11} - 2\tilde{c}_{66}$. It is seen that $\tilde{c}_{11}$ includes considerable errors.
is considerably higher than $c_{33}$ for all specimens. This fact indicates SiO$_2$ films possess the 6mm anisotropy, which may be induced during the deposition or caused by variation of film properties in the thickness direction.

![Graph](image)

**Fig. 7.** Fitting results when the 6mm anisotropy is assumed. (Sample (a); SiO$_2$/Si, deposition condition = CVD#1)

V. CONCLUSION

Elastic properties of SiO$_2$ films were evaluated by the LFB ultrasonic microscopy.

First, a simple technique was proposed to estimate elastic constants. Namely, the SAW velocity is measured by the LFB microscopy as a function of $\theta$, and elastic constants are estimated by fitting the measured $\theta$ dependence with the simulation. Effectiveness of this technique was demonstrated by its application to the Si (001) surface.

Then the technique was applied to the SiO$_2$ thin film / Si substrate structure. Large variation of elastic properties are observed among four specimens prepared with different apparatus and conditions. Elastic properties normal to the surface seemed to be significantly different from those along the surface. Namely, SiO$_2$ films possess the strong 6mm anisotropy which may be induced during the deposition or caused by variation of film properties in the thickness direction.

Acknowledgment

This work was partially supported by a Grant-in-Aid for Scientific Research (KAKENHI) from the Japan Society for the Promotion of Science (JSPS).

REFERENCES


Study on double-modulation coherent population trapping resonance

Peter Yun, Sinda Mejri, François Tricot, David Holleville, Emeric de Clercq, Stéphane Guérandel
LNE-SYRTE, Observatoire de Paris, CNRS, UPMC
75014 Paris, France
enxue.yun@obspm.fr

Abstract—Spectroscopy studies of coherent population trapping (CPT) with constructive polarization modulation are presented. In these studies, the effects of modulation frequency and laser intensity on the CPT contrast of the clock transition are investigated. We experimentally show that no CPT transition with \( \Delta m_F=2 \) is excited. Moreover, the constructive polarization modulation CPT signals are observed in both cases, with excited levels \( F'=3 \) and \( F'=4 \) in the Dl line of Cesium. Our studies show that this scheme, with potential to implement a compact and high performance clock, could also be applied to high pressure buffer gas cell, e.g., chip scale atomic clock.

Keywords—CPT; polarization modulation; spectroscopy;

I. INTRODUCTION

A passive coherent population trapping (CPT) [1] clock with high performance [2] is undergoing at least two approaches: one is Michelson interferometer based push-pull configuration [3] at FEMTO-ST [4], the other is the constructive polarization modulation scheme [5] at SYRTE, which employs a polarization switch to replace the interferometer and also noted as double-modulation scheme. The latter could be seen as a time version of push-pull optical pumping [3], thus it is interesting to study the behavior of atomic system in a vapor cell interrogated by bichromatic laser with constructive polarization modulation, like the study of its counterpart with only the polarization modulation [6]. These studies also pave the road to a compact and high performance CPT atomic clock.

II. EXPERIMENTAL SETUP

Our setup shown in Fig. 1 is similar to the one used in [5]. For simplicity of the experimental setup and reduction of the intensity noise induced by a fiber MZM as depicted in [7], we replaced it with an electro-optic phase modulator (EOPM). This EOPM is modulated at 4.596 GHz with 26.5dBm power, where it transfers maximum of carrier power to the ±1st sidebands light, which is about 70%.

The multi-chromatic laser generated by EOPM, in which a Raman phase modulation is also applied, is chopped by an AOM for one pulse CPT spectrum experiments. With a polarization modulation switch, which is synchronized with Raman phase modulation, we get the double-modulated laser. The laser beam diameter is expanded to 15mm before the vapor cell. The cylindrical Cs vapor cell, 20 mm diameter and 50 mm long, is filled with 21 Torr of mixed buffer gas (argon and nitrogen). The cell temperature is stabilized at 32.5 °C. In our experiment unless stated otherwise, a uniform magnetic field of 6.88 \( \mu \)T along the direction of cell axis is applied to remove the Zeeman degeneracy.

Fig. 1. Setup of double-modulation CPT.

Fig. 2. Time sequence.

This work is supported in part by ANR and DGA (ISIMAC project ANR-11-ASTR-0004). This work has been funded by the EMRP program (IND55 Mclocks). The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.
A. Contrast vs $f_m$

Fig. 3. The contrast of the clock transition as function of modulation rate. Pulse sequence parameters: $t_w=9\text{ms}$, $t_r=2.28\text{ms}$, $t_w=10\mu\text{s}$ (except for $f_w=83\text{kHz}$ and $100\text{kHz}$, with $t_w=5\mu\text{s}$). Light power: 2.63 mW.

As we can see in Fig. 3, the optimal modulation frequency for maximum contrast of clock transition is about 2 kHz. For a lower frequency, the atomic population in the clock CPT states is pumped out to the maximum $m_F$ ground Zeeman states, which reduces the contrast of the clock transition. At high frequency, the finite polarization switch time of our EOM (about 2.5 $\mu\text{s}$) is one of the reasons for contrast drop. During this time, both polarizations coexist, while the Raman phase is the same for both bichromatic lasers with counter-rotating circularly polarizations. This laser configuration will induce destructive interference for the two CPT states built by the two counter-rotating polarizations. The ratio of switch time to modulation period $t_s=1/f_m$ increases as $f_m$, thus the destructive effect of CPT states becomes more noticeable.

B. Contrast vs laser intensity

Fig. 4. The contrast of the clock transition as function of the laser intensity in one pulse CPT. Pulse sequence parameters: $t_w=1\text{ms}$, $t_r=3.19\text{ms}$, $t_w=10\mu\text{s}$, $f_w=2\text{kHz}$.

The saturation of the clock transition contrast versus the laser intensity is shown in Fig. 4, which is a typical behavior of the CPT resonance.

C. CPT spectrum without $\Delta m_F=2$ transition

Fig. 5. CPT spectrum of clock transition with C-field at 6.88 $\mu\text{T}$ and 100 $\mu\text{T}$. Pulse sequence parameters: $t_w=1\text{ms}$, $t_r=3.19\text{ms}$, $t_w=10\mu\text{s}$, $f_w=2\text{kHz}$. Light power is 0.325 mW.

In push-pull and lin lin CPT experiment, the CPT transitions with $\Delta m_F=2$ are observed with large enough static magnetic field [4, 8]. However, there is no such CPT peak observed in our scheme as shown in Fig. 5, because the two counter-rotating circularly polarized bichromatic laser beams do not coexist at the same time in the ideally double-modulation scheme. The shift of two CPT resonances at C-field of 6.88 $\mu\text{T}$ and 100 $\mu\text{T}$ are in agreement with the theoretical prediction: $\nu_z(B_1)-\nu_z(B_2)=425$ Hz, using the formula of the quadratic Zeeman shift of the clock transition: $\nu_z=0.0427453B^2$, here $\nu_z$ is expressed in Hz, and B in $\mu\text{T}$.

With this light power of 0.325 mW in Fig. 5, the contrast and full-width at half maximum (FWHM) of the clock transition are 8.5% and 615 Hz, respectively. These values are close to the values reported in [7], which demonstrated a push-pull CPT clock with frequency stability at the level of $4.2\times10^{-13}$ at one second in the continuous mode.

D. CPT spectrum with excited level $F'=3$ and $F'=4$

Fig. 6. CPT spectrum with $F'=3$ and $F'=4$ in one-pulse CPT. Pulse sequence parameters: $t_w=1\text{ms}$, $t_r=3.19\text{ms}$, $t_w=10\mu\text{s}$, $f_w=2\text{kHz}$. Light power is 2.23 mW.
The CPT spectrum with excited level $F'=3$ and $F'=4$ in the D1 line of Cesium is shown in Fig. 6, the corresponding contrasts of clock transition are 9.6% and 14.3%, respectively. The high contrast signals show the validity of our scheme when the excited levels are spectrally unresolved in the case a cell filled with high pressure buffer gas, e.g., chip scale atomic vapor cell.

IV. CONCLUSION

We present the spectroscopic studies of constructive polarization modulation for CPT. In these studies, we found that the optimal modulation frequency for maximum contrast of clock transition is about 2 kHz. The laser intensity dependence of clock transition contrast shows a saturated behavior. We experimentally demonstrate there is no CPT transition with $\Delta m_F=2$ and the validity of our scheme in both case of excited level $F'=3$ and $F'=4$ in D1 line of Cesium. These spectroscopic studies, besides providing key parameters for implementing a high performance clock, also indicate the potential application for chip scale atomic cell.

Acknowledgment

We would like to thank Thomas Zanon-Willette, Jean-Marie Danet, Natascia Castagna, Olga Kozlova, Rodolphe Boudot and Moustafa Abdel Hafiz for helpful discussions. We are grateful to José Pinto Fernandes for technical assistance and the realization of various electronic devices. We are also pleased to acknowledge Pierre Bonnay and Annie Gérard for manufacturing Cs cells.

References

Compact and high-performance Rb clock based on pulsed optical pumping for industrial application

Songbai Kang*, Mohammadreza Gharavipour, Florian Gruet, Christoph Affolderbach, Gaetano Mileti
Laboratoire Temps-Fréquence, Institute of Physics, University of Neuchâtel, Neuchâtel, Switzerland.
Email: gaetano.mileti@unine.ch

Abstract—We report on the development of a compact laser-pumped Rb clock based on the pulsed optical pumping (POP) technique, in view of future industrial applications. The clock Physics Package (PP) is based on a compact magnetron-type microwave cavity of 45 cm$^3$ volume, and our current clock PP has a volume of only 0.8 liters, including temperature control and magnetic shields. This clock PP is completed by a newly-developed frequency-stabilized laser head of 2.5 liters overall volume, with an acoustic optical modulator (AOM) integrated within the laser head for switching the laser output power. Due to the highly uniform magnetic field inside the microwave cavity, Ramsey signals with high contrast of up to 35% and with a linewidth of 160 Hz have been demonstrated. A typical short-term clock stability of $2.4\times10^{-13}\tau^{1/2}$ is measured. Thanks to the pulsed operation, the light-shift effect has been considerably suppressed as compared to previously demonstrated continuous-wave (CW) clock operation using the same clock PP, which is expected to enable improved long-term clock stabilities down to the $10^{-14}$ level or better.

Keywords—Rb clock; POP; frequency stability; magnetron-type cavity.

I. INTRODUCTION

Commercial lamp-pumped Rb atomic clocks have been providing the stable and reliable frequency and timing signals in a wide range of industrial applications, such as telecommunication, navigation, space application and others [1]. However, with the ongoing quest for improved stability performances (especially on long-term timescales of one day or more), laser-pumped vapor cell clocks [2-5] have been considered as the most promising candidates to meet the next generation’s industrial demand and technical applications, because they feature the combined advantages of high frequency stability, compactness and reliability.

Recently, a laser-pumped Rb clock prototype based on the continuous-wave (CW) double-resonance (DR) principle has achieved a state-of-the-art short-term stability of $1.4\times10^{-13}\tau^{1/2}$ [6]. With a compact magnetron-type cavity [7] and the simplicity of its optical setup (no need for AOM), the prototype has a greatly reduced volume (by a factor of 10) with respect to a passive H-maser. However, for the long-term stability (considered here at $10^5$ s) its light shift (LS) induced frequency instability is still at the level of $4\times10^{-14}$. Thanks to the pulsed optically pumped (POP) technique [8, 9], the LS effect can be considerably suppressed due to the separation of light-atom and microwave-atom interactions. A POP Rb clock principle prototype developed in INRIM has already demonstrated an unprecedented long-term stability of $\sim 1\times10^{-15}$ at $10^5$ s [10].

We are developing a compact laser-pumped Rb clock based on POP technique. For the physics package (PP), we use the same compact magnetron-type cavity [7] as in the previous CW laser-pumped Rb clock, which has been proven to be very suitable for application to POP Rb clocks [11, 12]. For the clock’s optical setup, we recently have developed a fully integrated frequency-stabilized laser system integrating an AOM for switching the laser output. In this communication, we report on the clock prototype’s Ramsey fringes, characterizations of light-shifts, microwave power shifts and short-term stability performance.

II. PROTOTYPE SETUP

Our POP clock prototype is composed of three main components, 1) the PP containing the cavity-cell assembly; 2) the frequency-stabilized laser system; 3) the local oscillator (LO).

The PP has been detailed previously in [11, 12]. It has a volume of 0.8 liters, including temperature control and magnetic shields. The magnetron-type cavity, core of the PP, has a low quality factor of about 200 (when loaded with the vapor cell). A highly uniform magnetic field inside has been proven which realizes well the required $\pi/2$ microwave pulses for a large fraction of the Rb atoms across the cell.

The laser head is similar to the laser systems developed previously, where the AOM served only for shifting the laser frequency [13]. In the newly-developed laser head used here, the AOM is implemented to realize both the laser output switching and frequency shifting. The schematics of the optical setup and the photograph for the newly-developed laser head.
are shown in Fig. 1 (a) and (b). A DFB laser diode is used as light source, emitting at 780 nm and with a measured linewidth of 1.7 MHz. The laser frequency is locked to the $F_g=1$ to $F_e=1-2$ cross-over sub-Doppler saturated-absorption line obtained from an evacuated reference Rb cell. The AOM is operated in the double-pass technique, and the final output laser is frequency-shifted by -160 MHz and shows an extinction ratio of 40 dB (laser on vs. laser off). Other relevant performance characteristics are a relative intensity noise (RIN) $5\times10^{-12}$ Hz$^{-1}$ at 100 Hz, and a fractional frequency stability of $6\times10^{-12}$ and fractional intensity stability of $7\times10^{-4}$, both at a time scale of $10^4$ s. The whole volume of the assembly laser head is 2.5 liters. The optical setup only occupies less than half of the laser head volume (1 liter), leaving room for further reduction in size in view of the realization of a highly compact atomic clock.

The LO consists of the microwave synthesizer, the servo loop electronics, and a quartz oscillator (OCXO) [14]. Under the POP scheme, the estimated clock stability limit due to Dick effect is at the level of $7.4\times10^{-14} \tau^{-1/2}$.

### III. EXPERIMENTAL RESULTS

For the POP clock operation in Ramsey scheme, the durations of the three interrogation phases are set as follows: 1) optical pumping time $T_p=0.4$ ms; 2) microwave pulse time $T_1=0.4$ ms; 3) Ramsey time $T_{\text{Ramsey}}=3$ ms; 4) optical detection time $T_d=0.7$ ms. The total cycle period $T_c$ is 4.94 ms (including some pauses).

#### A. Ramsey Pattern

Figure 2 shows typically detected Ramsey fringes. The pumping and detection light power are 15 mW and 120 $\mu$W, respectively. The microwave power sent to the cavity is set to -20 dBm to realize the $\pi/2$ pulse. The central fringe contrast is of 35% and its full-width at half-maximum (FWHM) $\Delta \nu$ is 160 Hz, which is consistent with the theoretical prediction of $\Delta \nu=1/T_{\text{Ramsey}}$. The POP prototype’s estimated shot-noise limit is below $2\times10^{-14} \tau^{-1/2}$ [11]. We note that, based on the same PP, the DR signal for the previous CW laser-pumped Rb clock has a 26% contrast and 334 Hz linewidth and its shot-noise limit is at $4.9\times10^{-14} \tau^{-1/2}$ [6].

#### B. Light-shifts

Fig. 3 (a) and (b) show the measured POP Rb prototype’s fractional frequency shifts depending on the normalized laser output power and the laser frequency detuning. The LS intensity and frequency coefficients are $-2.1\times10^{-14}/%$ and $4.6\times10^{-13}/$MHz respectively, both of which are more than one order of magnitude lower than in the CW scheme [6, 15, 16]. Although LSs effect can be fully suppressed for the POP scheme, the residual LSs still exist in our case probably due to the position shift effect (non-uniformity of microwave field and spatial distribution of laser absorption rate) [17], residual coherence at the beginning of the first Ramsey pulse [18] and residual non-zero light power during the Ramsey phase.

Table I shows the estimated long-term stability limit due to LS effects, for both POP and CW schemes. Comparing with the $-4\times10^{-14}$ instability limit for the CW scheme, the LS induced instability limit in the POP scheme reduces to the level of $-1\times10^{-15}$ which is sufficient to satisfy requirements for the envisaged industrial applications. Similarly, the estimated short-term instability contribution induced by the LS effects for the POP clock prototype is $-1\times10^{-14} \tau^{-1/2}$.
Microwave power shift

The prototype’s fractional output frequency depending on the microwave power is shown in Fig. 4. The fitted power shift coefficient is $1.8 \times 10^{-12} / \mu W$. Two potential sources can contribute to this microwave power shift. One is the position shift effect due to the spatial inhomogeneity of light and microwave field, and also the residual gradients in the static magnetic field. The other is the cavity pulling effect [19]. Further investigation will follow to confirm the origin of these shifts in our prototype. The POP prototype’s microwave power shift coefficient is slightly lower than that in CW scheme which is $2.2 \times 10^{-12} / \mu W$ [6]. If considering a typical LO performance of 1 nW power fluctuation at long term scale ($10^5$ s), we obtain a power-shift instability limit of $-2 \times 10^{-15}$.

Short-term stability

Figure 5 depicts a typical short-term stability performance for our compact POP Rb clock prototype, when the laser frequency is locked to the CO1-12 reference transition. The measured short-term clock stability of $2.4 \times 10^{-13} \tau^{-1/2}$ can guarantee an instability level of $2.4 \times 10^{-15}$ at $10^5$ s. Currently, the dominant source for this stability is the optical detection noise. Furthermore, it mainly originates from the laser’s FM-to-AM conversion when the detection light goes through the Rb vapor cell in the clock PP [20]. In addition, we found that currently the RF signal driving the AOM also introduces some noise that can be totally avoided in the CW scheme [11], or potentially by active stabilization.

IV. CONCLUSION

We reported the latest results for our compact Rb clock prototype based on the POP technique. A clock PP with 0.8 liters volume based on the magnetron-type cavity was used for these studies. A newly-developed frequency-stabilized laser head with an AOM integrated and an overall volume of 2.5 liters was used as the light source for optical pumping and detection, replacing the previous tabletop laser system. This presents a further step towards the miniaturization of a POP Rb clock for industrial application.

| Estimations of Short- and Long-Term Clock Stability Limits Due to the Light Shift Effects for POP and CW Schemes. |
|-------------------------------------------------|----------|----------|
| POP | CW |
| Long-term instability (at $10^5$s) | Frequency LS | $-1 \times 10^{-15}$ | $3.7 \times 10^{-14}$ |
| | Intensity LS | $-1 \times 10^{-15}$ | $9.9 \times 10^{-15}$ |
| | Total | $-1 \times 10^{-15}$ | $3.8 \times 10^{-14}$ |

C. Microwave power shift

The prototype’s fractional output frequency depending on the microwave power is shown in Fig. 4. The fitted power shift coefficient is $1.8 \times 10^{-12} / \mu W$. Two potential sources can contribute to this microwave power shift. One is the position shift effect due to the spatial inhomogeneity of light and microwave field, and also the residual gradients in the static magnetic field. The other is the cavity pulling effect [19]. Further investigation will follow to confirm the origin of these shifts in our prototype. The POP prototype’s microwave power shift coefficient is slightly lower than that in CW scheme which is $2.2 \times 10^{-12} / \mu W$ [6]. If considering a typical LO performance of 1 nW power fluctuation at long term scale ($10^5$ s), we obtain a power-shift instability limit of $-2 \times 10^{-15}$.
The compact POP Rb clock prototype shows a Ramsey signal of 35% with a linewidth of 160 Hz from which the shot-noise limit is estimated as below $2 \times 10^{-14} \tau^{-1/2}$. The most relevant achievement for the prototype is on the light shifts effect suppression (frequency LS: $4.6 \times 10^{-13}$ MHz; intensity LS: $-2.1 \times 10^{-15}$%) which restricts the clock’s light-shift related instability limit to the level of $-1 \times 10^{-15}$ at $10^8$ s. This LS-induced instability limit is more than one order of magnitude lower than the CW case. The measured microwave power shift coefficient is $1.8 \times 10^{-12}/\mu$W and the resulting estimated long-term frequency fluctuation is $-2 \times 10^{-15}$. Finally, a typical short-term stability of $2.4 \times 10^{-13} \tau^{-1/2}$ has been achieved, which is currently limited by the optical detection noise. This work paves the way for the development of a highly compact and high-performance POP Rb clock that could find its applications in industrial applications like local timing sources or metrology references, in telecommunication, space navigation, or could also be used as high-performance local oscillator (LO) reference.

ACKNOWLEDGMENT

We thank A. K. Skrivervik, C. Stefanucci (both EPFL-LEMA) and M. Pellaton and T. Bandi (both UniNe-LTF) for their contributions to realizing the cell and cavity, and C. Calosso (INRIM, Italy) for providing the LO.

REFERENCES


Identification and Calibration of Ground System Biases in Ground to Space Laser Time Transfer

Ivan Prochazka*, Josef Blazej*, Jan Kodet†

Email: {ivan.prochazka,blazej,jan.kodet}@fjfi.cvut.cz

* Dept. of Physical electronics, Czech Technical University in Prague, Prague 1, Czech Republic
† Technical University Munich, Fundamental Station Wettzell, Germany

Abstract— Laser time transfer is an attractive technique to transfer time ground to space with picosecond precision and systematic errors on the level of tens of picoseconds. Recently the European Laser Timing experiment is under construction in the frame of the European Space Agency mission Atomic Clock Ensemble in Space. The objective of this laser time transfer is synchronization of the ground based clocks and the clock on board the International Space Station with picosecond precision and the accuracy better than 50 picoseconds. We are reporting on a progress in identification and calibration of the biases associated to both ground and space segment. To characterize the delays of a ground segment the Calibration Device has been developed and tested. It enables to calibrate the ground based laser systems for their systematic timing biases. As a result the ground to ground time transfer in an ELT experiment should be accomplished with systematic errors not higher than 25 picoseconds. To determine the delays of a space segment the new calibration procedure has been designed and tested as well. The calibration test results will be presented.

I. INTRODUCTION

The precise and accurate time transfer is a prerequisite of a number of experiments in fundamental physics, Earth science, global navigation and many other disciplines. One of the most challenging disciplines is the time transfer ground to space. The standard techniques used recently are based on radio frequencies [1]. Significant improvement in time transfer accuracy ground to space might be achieved using optical frequencies [2]. In optical frequency range the systematic error contributors to the overall error budget may be limited to the level of 10 picoseconds.

A. Operating Principle

The laser time transfer (LTT) ground to satellite is an extension of the standard measurement technique of satellite laser ranging (SLR) [2], its principle is drawn in Fig. 1.

The satellite equipped with optical retro reflectors is ranged using short laser pulses. A short and powerful laser pulse is transmitted toward a satellite, part of the energy is reflected by the retro reflector onboard the satellite back to the ground. The reflected optical pulse is detected at the ground station and the pulse propagation time is evaluated. The range \( D \) is determined on the basis of the measured laser pulse propagation time toward the target satellite and back again. The epoch of transmission of laser pulse \( T \) is monitored with respect to the local clock for each laser pulse emission. For the LTT experiment the existing satellite laser ranging ground stations is used. The satellite is equipped with retro reflectors to enable the laser ranging and, additionally, with an optical detector which detects and time tags the arrival of laser pulse at the satellite. The satellite range \( D \) is measured by laser ranging to the on-board retro reflectors and the arrival time of the laser pulse to the satellite \( E \) is recorded by on board clock and the recorded time tags are transmitted to ground via satellite telemetry channel. Combining the laser pulse emission times, propagation and instrumental delays and satellite arrival times, the space clock and the station clock may be compared. The SLR technique has been well developed in last years, ranging and epoch timing precision of the order of \( 1 \times 10^{-11} \) s may be achieved. Fortunately, the influence of the atmosphere in propagation time ground to space is completely compensated in a laser time transfer when combined with satellite laser ranging at the same time.

Fig. 1. The principle of ground to space clock synchronization by means of laser pulses. The satellite range \( D \) is measured by laser ranging technique, the laser emission time \( T \) is recorded by ground clock, the arrival time of the laser pulse to the satellite \( E \) is recorded by on-board clock.

The key problem of any time transfer is its systematic error budget, namely the list and characterization of all its contributors [3]. The block scheme of LTT is plotted in Fig. 2 where the individual components of system delay are displayed.

The ultimate experiment goal of laser time transfer is to determine the relation between the ground and space time scales. These time scales are characterized by “1pps” pulses. Their epochs are denoted \( TG \) and \( TS \) for ground and space
time scales respectively. From the block scheme one can identify these contributors:

- **DR** signal propagation time from ground to space is determined on the basis of satellite laser ranging result. Its accuracy is limited by determination of reference points and SLR system calibration. These parameters may be estimated with ~1 mm over all accuracy [4], what corresponds to the accuracy of the DR of 3 ps.

- **D3, D5** detection delays of the photo detectors may be determined by a new procedure proposed and tested in our laboratory with the accuracy of each contributor of 15 ps [5].

- **D1, D2** laser beam propagation inside the laser station.

- **D4, D6** signal propagation in electrical cables on ground and in space respectively.

![Fig. 2. Laser time transfer system delays](image)

### II. GROUND SEGMENT DELAYS CALIBRATION

To characterize the ground segment – namely to measure the systematic delays in a laser time transfer a new approach to calibration of system delays was developed. The method is based on a presumption, that all the ground stations participating in the European Laser Time transfer experiment will be calibrated versus a dedicated set consisting of the photon counting detector identical to the satellite one, epoch timing system and signal cable. These components will form an ELT Calibration Device, which has been developed in our lab [6]. The calibration principle is plotted in Fig. 3.

The Detector in Fig. 3 is a twin of a photon counting detector used in space segment. Its photon to electrical signal delay $D5'$ may be determined with accuracy better than 15 ps [3]. Both the Event Timers are referred to the common time scale and clock frequency. One common signal cable for “1pps” has to be used to synchronize the Event Timers $ETG$ and $ETS$ consecutively. Considering the experiment setup the calibration constant $B$ for ELT related to the particular ground station can be evaluated as (1).

\[
B = \frac{L}{c} - (ES - EG) \tag{1}
\]

where $B$ is the calibration constant, $L$ is a separation of reference points, $c$ is a group speed of light, $ES$ and $EG$ are the epoch readings of Even Timers of the ground and space segment respectively. Obviously the accuracy of this calibration constant $B$ is limited by the accuracy of determining the distance $L$ and epochs $ES$ and $EG$ only. As the distance $L$ is of the orders of meters, it can be determined with accuracy better than one millimeter. The systematic error contribution of the atmospheric propagation of optical pulse is negligible on this distance. This calibration scheme was tested in a series of indoor experiments and in a real field operation [7]. The delays associated with the SLR Station in Wettzell, Germany, have been measured. The resulting accuracy and stability of determining the calibration constant $B$ was better than 25 ps over a period of three months of operation [6]. This accuracy is better than 25 ps per site. It characterizes the ground–ground laser time transfer accuracy. Considering equal accuracy of delays related to both ground terminals one can conclude that the laser time transfer accuracy ground–ground should be 36 ps.

![Fig. 3. Block scheme of calibration of the laser time transfer ground segment by means of ELT Calibration Device.](image)

\[
C = B - (D5' + Dets + Dc) \tag{2}
\]

The absolute value of an optical to electrical delay $D5'$ of a detector was determined using a procedure described in [5]. The optical test pulses 80 ps long at a wavelength of 531 nm were generated by a laser PicoQuant LDH-P-FA-530B. The electrical signals were monitored using a 2.5 GHz bandwidth...
and 40 Gs/s oscilloscope. The results are summarized in Fig. 4 where the detection delay of an Engineering Model of the ELT detector is plotted.

Fig. 4. Detection delay of an Engineering Model of the ELT detector measured repeatedly 2 series per day.

In total six measurement series were completed within 3 days. Two different reference linear photo diodes were used to record a multiphoton signal [5]. The series 1, 2 and 5, 6 were completed using photodiode #1 and series 3 and 4 were completed using photodiode #2. The internal consistency of the series using identical photodiode was well within ±4 ps peak–to–peak. The overall data spread is ±14 ps peak–to–peak with the mean value of a detection delay is 2,168 ps.

The delay of a signal cable interconnecting the detector and event timing system of a Calibration Device was measured using conventional techniques applying the epoch timing device NPET [8]. The delay value of (7,114±2) ps peak–to–peak was measured.

The delay attributed to a response of the epoch timing device used in a calibration process may be estimated on a basis of experiments and results described in [9]. From these results one can conclude, that the uncertainty of epoch timing by means of this device is less than ±15 ps peak–to–peak.

Combining all the accuracies of contributors to components in (2) one can conclude that ground system calibration constant C may be determined with the accuracy of 33 ps. This value will characterize the accuracy of delay constant related to particular ground station in a laser time transfer ground–space. In this value the ground system stability over a period of several months of operation is included

III. SPACE SEGMENT DELAYS CALIBRATION

Just recently the requirement for the space segment delay values has appeared. Block scheme of the delay contributions is in Fig. 2. Space segment is composed of detector delay $D_5$, interconnecting cable delay $D_6$ and of the delay related to an event timing system response. The optical to electrical delay $D_5$ of the detector will be determined in a process of final assembly operational tests of the flying unit of the detector. The procedure [5] is identical to that one used for characterization of a detection delay of the Calibration Device in a previous chapter. The accuracy of this delay will be identical, typically ±15 ps peak–to–peak.

The delays of the signal cable and of the timing system have to be determined by a different technique, because the measurements have to be completed during assembly phase of a flying unit in an “open box” configuration. The following procedure has been designed and tested see Fig. 5 for its block scheme.

The sum of signal cable and timing system delays will be determined. The test signals will be generated using ELT detector simulator, which consists of a modified ELT detector breadboard device. The modification is such that the device is not responding to optical signals, but its output signals are generated with a fixed delay of several nanoseconds after the “Gate ON” signal is provided. For test purposes this gate signal will be generated synchronously to a local time scale by the microwave link (MWL) timing system. The time measurements will be accomplished using a high bandwidth ($\geq 1$ GHz) digitizing oscilloscope providing at least 10 Gsamples per second. The detector output signal and the local time scale “1pps” signal will be monitored consecutively using a single high bandwidth probe. Their times of arrival $ED$ and $Epps$ to the oscilloscope will be recorded in relative oscilloscope time scale. The epochs $ES$ of arrival of detector output pulses into the MWL timing system will be recorded as timing system output data. Mean values of a series of measurements for all the parameters will be used. The total delay $D$ of a space segment for laser time transfer may be expressed in a form (3):

$$D = D_5 + (ES - ED + Epps)$$

The procedure described above has been tested in a laboratory experiment. The MWL Timing unit has been simulated using a standard NPET timing system [8, 9] analogically to the calibration of Calibration Device delays. The events $ED$ and $Epps$ were recorded by means of a fast digitizing oscilloscope with a bandwidth of 2.5 GHz using a sampling rate of 20 Gs/s what corresponds to a timing resolution of 50 ps/dot. The oscilloscope was externally triggered by a “1pps” signal. For each measurement series of 10 wavefronts have been recorded. The time interval definition used the trigger levels and slopes identical to that ones
used in a real space device. The relative time delays of wavefronts were determined with a precision in a range of 12 to 25 picoseconds r.m.s. depending on a slew rate of a pulse under test.

The experiment was repeated five times for three different lengths of cables. The signal propagation delay of these cables was determined by an independent test with accuracy better than 5 ps. The delay related to the NPET timing system equals to zero with the accuracy of 15 ps [8, 9]. Considering these facts the accuracy of the delay determination method described above was estimated to be well below ±50 ps peak to peak. These accuracies are well below requirements put on the laser time transfer channel of the mission.

IV. CONCLUSION AND SUMMARY

We have presented the procedure and test results of determining the accuracy of delays associated with a ground to ground and ground to space laser time transfer in the European Laser Timing project. The real field measurement data indicate the worst case estimate of accuracy of the ground to ground laser time transfer of 36 picoseconds. The laboratory test simulating the measurements in an “open box configuration” at final flying unit assembly indicated the accuracy of determining the accuracy of the space segment delay well below 50 picoseconds.

ACKNOWLEDGMENT

This work has been carried out at the Czech Technical University in Prague. Numerous grants were provided by the Czech Grant Agency, Czech Ministry of Education and by international agencies. Recently, the research and development of time transfer technologies is supported by MSMT CR grant LH12005. The support of the Fundamental Station Wettzell, Germany, operated jointly by Federal Bureau of Cartography and Geodesy and the Technical University Munich is highly appreciated.

REFERENCES

Characterization of an ultra stable quartz oscillator thanks to Time Transfer by Laser Link (T2L2, Jason-2)

A. Belli1,2, P. Exertier1, E. Samain1 and C. Courde1
1: Géoazur 250 Rue A. Einstein
F - 06560 Valbonne
Email: belli@geoazur.unice.fr
F. Vernotte2
2: UTINAM/Observatory THETA
Université de Franche-Comté/CNRS
F - 25000 Besançon
Email: francois.vernotte@obs-besancon.fr
A. Auriol3 and C. Jayles3
3: CNES French Space Agency
18 Avenue E. Belin
F - 31455 Toulouse
Email: Albert.Auriol@cnes.fr

Abstract—The T2L2 experiment (Time Transfer by Laser Link), on-board Jason-2, with an orbit at 1335 km, since June 2008 allows the clock synchronization between ground clock (generally H-maser) and space clock (quartz Ultra Stable Oscillator (USO) DORIS) with a stability of a few picoseconds over 100 seconds. In common view, when two laser stations see T2L2, the time transfer stability is less than 10 picoseconds over few seconds. In order to perform non-common view time transfer for synchronizing distant ground clocks, it is important to precisely characterize the on-board oscillator at least on 10,000 seconds (maximal flight time between two distant stations). The key is to study the space environment on the Jason-2 orbit, to separate deterministic and stochastic behaviors of the USO (shift and drift). We show that T2L2 is able to provide accurate frequencies, which are deduced from the ground to space time transfer over each laser station (few \(10^{-13}\)). Since 2008, these time transfers helped us to create an on-board frequency data base.

The major contributors to these frequency variations on 10,000 seconds are temperature and space radiation especially due to the South Atlantic Anomaly (SAA) in which Jason-2 pass through. Aging can be considered as a linear drift during 10,000 seconds and the effect of radiation like a very small shift over each SAA overflight. The effect of the temperature is driven by the on-board temperature measurement. A model is realized to represent these effects on USO with a RMS of few \(10^{-13}\) over 10,000 seconds. Space phenomena are also playing an important role in long term. Actually, if we consider both accumulation dose received by radiation and aging, we can explain 99.9 % of the global frequency variation of the USO since the beginning of the T2L2 mission.

I. INTRODUCTION

The Time Transfer by Laser Link (T2L2) experiment was launched in 2008 on-board the oceanographic space mission Jason-2 at an altitude of 1335 km. T2L2 aims to synchronise ultra-stable clocks using the Satellite Laser Ranging (SLR) technique. T2L2 consists of an optical system for detection and an electronic device for timing [1]. The timing device is referenced to an Ultra Stable Oscillator (USO) which is provided by the DORIS (Doppler Orbitography and Radiopositioning Integrated on Satellite) system; DORIS is used to provide the Jason-2 orbit with precise and worldwide tracking Doppler measurements [2].
T2L2 is able to accurately date, in the time scale of the satellite, the optical events (with picosecond (ps) resolution) which come from the laser stations (a worldwide network which is operated by the International Laser Ranging Service, ILRS see Fig. 1). Several SLR stations (some of these operate several geodetic techniques on the same site, e.g. SLR, GNSS, and DORIS) are equipped with H-maser which provides the local time reference system to date the laser events. Each satellite pass, over a given station, is of 10 to 15 minutes duration. As a maximum, a given station is able to track the Jason-2 satellite 5 to 6 times per day.
T2L2 is a technology demonstrator with performances at the metrological level. The ground-to-space time transfer (from a SLR station to the satellite) reach a stability of 5-6 ps over around 75 s in best cases. It is thus an opportunity to compute a frequency bias of a few \(10^{-13}\) over this interval. When the satellite is in Common View (CV) of two SLR stations, it is thus possible to establish a ground-to-ground link between both stations at ground level by combining both ground-to-space data sets; the overall performance of this link is of 12-
15 ps rms over a common pass with a repeatability of around 50 ps over several CV passes (at 1 day) [3], [4]. In case of non CV time transfer (Jason-2 has an orbital period of around 6400 s), the USO unstability is the main source of limitation; it has been established at $3.5 \times 10^{-13}$ over 1000 s [2]. Thus, over around 10,000 s, it is necessary to study all possible sources of nonlinear effects which affect the USO. The goals of the present study are: i) to identify the main physical effects, e.g. aging, temperature and radiations, ii) to modelize each effects, by taking into account the sensitivities of the USO which were measured before the Jason-2 launch in laboratory conditions (ref), ii) to extract the relative frequency bias of the USO from ground-to-space time transfers along the time, and then to compare the observed frequency variations to the proposed model (over around 10 days).

II. HOW TO CHARACTERIZE THE USO

The stability of the USO is due to stochastic and deterministic behaviors. In space or laboratory conditions, several physical effects that affect the frequency of an oscillator have been studied by numerous authors (see e.g. [5], [6]). The new USO model (DGXX generation) onboard Jason-2 has been developed by the French Space Agency (CNES) for current and future space missions such as Cryosat, HY-2A, and the next Jason-3. The performance and, above all, the expected sensitivity of the oscillator frequency to several types of physical effects were carefully studied, and in particular the radiations [2]. The goal was to decrease the consequences of radiations in space (the altitude of Jason-like satellites is of 1335 km) as much as possible, knowing what happened on DORIS /Jason-1: the frequency was strongly affected above the South Atlantic Anomaly (SAA) area, making the geodetic products of DORIS unusable (the geocentric positioning of the ground beacons in South America notably) [7].

The new USO was pre-irradiated to be less sensitive to the radiation environment than the one on Jason-1 (gain factor >10). Because we are looking for establishing a time transfer in non-common view of 3–4 ns over 10,000 s and because we have at our disposal a high quality data set provided by the T2L2 (relative frequency biases at a few $10^{-13}$), we have only kept three major physical effects: aging, temperature and radiations. In addition, because the onboard epochs are in the proper time scale of the spacecraft we included also the main relativistic effects.

A. Relativity

At the level of $10^{-13}$, it is necessary to take into account the relative frequency shift due to General Relativity (in the context of the Schwarzschild solution) [8]:

$$\frac{\Delta \nu}{\nu} = \frac{1}{c^2} \left[ (U_{\text{sat}} - U_{\text{station}}) + \frac{1}{2} (v_{\text{sat}}^2 - V_{\text{station}}^2) \right]$$

(1)

On the one hand we simplified the model by removing the last term (in $V^2$, which is 0.003 times the term in $v^2$), on the other we considered the term $U_{\text{station}}$ as a constant term which only depend on the altitude of each SLR station above the geoid.

B. Aging

Aging is a proper effect of the oscillator. With time, the oscillator is becoming "older" what changes slowly its behavior. For example, except for the first 90 days in space (it can be represented by a logarithm function [9]), the linear drift of the USO frequency at the begining of the space mission could decreased (or increased) by a few pourcents a year. The aging has been largely studied generally.

In order to modelize this effect for a period of a few days, we choose a simple polynomial function of time $t$ (see Table I):

$$\frac{\Delta \nu}{\nu} = \sum_{j=0}^{3} \beta_j t^j$$

(2)

where: $\beta_j$ are empirical coefficients which are estimated from the data (see the Section dedicated to the data analysis).

C. Temperature

The DORIS USO is thermally controlled. This control is monitored and it allows us to have the temperature value every 30 seconds. These thermal variations are evolving between 9 to 11 deg. Celsius ($^\circ$C). First of all, there is a periodic behavior of the temperature of the onboard instruments (variations of $<0.5$ $^\circ$C) which is due to the passage of the satellite in the Earth’s shadow. In addition, due to the attitude law of the satellite [10] (with a period of 59 days) which is correlated to the orientation of the orbital plane with regard to the Earth-Sun direction, there is a slow increase (or decrease) of the temperature of around 2 $^\circ$C on few days.

By taking into account the expected sensitivity of the DORIS USO to the temperature $T$ (see Table I), and what has been studied before (e.g. [11]), we considered that the thermal changes of the frequency are quite comparable to a polynomial function of the temperature $T$, as:

$$\frac{\Delta \nu}{\nu} = \sum_{i=1}^{3} \alpha_i (T - T_0)^i$$

(3)

where: $T_0$ if the average value of $T$ during the studied period (a few days), and $\alpha_i$ are the coefficients of the model to be estimated from our data (see Table III).

D. Radiations

Radiations are induced by high-energy particles (protons) which are trapped by the Earth magnetic field. Thanks to some previous studies (notably with Jason-1 DORIS data, [7]), we know that the USO is affected by an energy higher than 87 MeV. In order to properly determine the received dose of radiations, it is necessary to estimate the flux (in $p^+cm^{-2}sr^{-1}s^{-1}$). The analysis which was conducted by [7] gave the geographical extension of the area (SAA) which dominates the radiations. In addition, the map provided the intensity of the particle flux which was later correlated with the CARMEN /Jason-2 data [12].
TABLE I
DORIS PERFORMANCES FOR THE DGXX GENERATION USO (CNES)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aging</td>
<td>&lt; 1 × 10^{-11}/day</td>
</tr>
<tr>
<td>Temperature</td>
<td>6.5 × 10^{-13}/degree</td>
</tr>
<tr>
<td>Radiations</td>
<td>6.7 × 10^{-12}/rad</td>
</tr>
</tbody>
</table>

The adopted model to represent the frequency changes of the USO due to the radiations received during the passages of the satellite in this area is based on i) an integration during the passage, and ii) a small period of relaxation. The idea behind is to take into account the fact that the Jason-2 USO is much less sensitive to the radiations than the other previous models (see Table I).

\[
\frac{\Delta \nu_{\text{SAA}}}{\nu} = \int_{t_0}^{t_{\text{SAA}}} \gamma D(t) dt \quad \text{during the SAA passage}
\]

\[
\Delta \nu(t) = \Delta \nu_{\text{SAA}} \exp\left(-\frac{t}{\tau}\right) \quad \text{after (4)}
\]

where:
- \( \gamma \) is a factor which is used to convert the received dose (in Gray or rad) in term of relative frequency change, \( \tau \) is the relaxation delay. Thanks to Jason-1, we know that this delay should be in the range of 15 to 7 min [7].
- In addition to this "local" effect, we should say that radiations occurs at a larger time scale. If the USO presents a local frequency drift during its exposure to the SAA radiations, there is also, after the relaxation period, a remaining effect (like a memory). The small but cumulative memory effect has to be integrated over time; it is thus an additional drift to the proper aging process.

III. HOW TO READ THE USO FROM TIME TRANSFER BY LASER

A. Ground to space time transfer

T2L2 is able to provide accurate relative frequency bias of the onboard oscillator, which can be deduced from the ground-to-space time transfer (passage) over each laser station. The ground-to-space time transfer is given by the following equation:

\[
\Delta \nu_{\text{SAA}}(t) = t_B + C_i - \left[ t_S + \frac{1}{2}(tof - C_S) \right] + \Delta \nu_{\text{SAA}}(t_0) + \int \frac{\Delta \nu_B}{\nu_B} dt + \text{Noise}
\]

where :
- \( \Delta \nu_{\text{SAA}}(t) \): is the ground-to-space (board minus station) time synchronization,
- \( t_B \): is the board epoch,
- \( t_S \): is the emission time of a given laser pulse from a station \( S_i \),
- \( tof \): is the time of flight of the pulse (two-way); T2L2 benefits from a two-way ranging thanks to the laser reflector array on Jason-2,
- \( C_i \): is the sum of the instrument corrections (around 0.3 ns),
- \( \Delta \nu_{\text{SAA}}(t_0) \): is the initial error,
- \( \Delta \nu_B/\nu_B \): is the frequency drift during its exposure to the SAA radiations, there occurs at a larger time scale. If the USO presents a local aging process, \( \Delta \nu_B/\nu_B \) is to take into account the fact that the Jason-2 USO is much less sensitive to the radiations than the other previous models

B. Building a complete model

As described by several authors, it is possible to consider that all effects, which are identified as the most sources of oscillator variations, can be added [13]:

\[
\frac{\Delta \nu}{\nu} = \sum_{i=1}^{3} \alpha_i (T - T_0)^i + \sum_{j=0}^{2} \beta_j t^j + \gamma_1 \int_{t}^{t_{\text{SAA}}} D(t) dt \quad (6)
\]

Now, we are able to compare the observations to the model. We choose a 10-day period over which a least square fit is performed. The goal is to estimate the set of coefficients \( \alpha_i, \beta_j \) and \( \gamma_i \) corresponding to this period. The initial values are given by CNES who estimated the sensitivity of the USO before the launch of Jason-2 (Table I).

In addition, the last coefficient \( \gamma_1 \) has been introduced with the aim of considering the proportion of dose absorbed by the USO.

Fig. 2 shows the complete model and the data over a period of 10 days (at the top), and below after a linear fit (the overall frequency drift is of 0.810^{-11} per day). In addition, we can see a pseudo-periodic behavior (at the orbital period) the amplitude of which is of a few 10^{-12}.

Fig. 3 shows the model (only radiations) and the frequency residuals after applying only aging plus temperature. Now, the frequency variations are under the level of 10^{-12}; the adjusted value of the \( \gamma_1 \) coefficient is of 0.05 in average (we analysed a complete set of 10-day periods over several years). The adjusted value of the relaxation coefficient \( \tau \) is of 8 min. We can see that the USO frequency biases which were extracted from SLR stations equipped with a H-maser (as Grasse SLR) are in better adequation with the model.

IV. RESULTS AND DISCUSSION

We tested the model on several 10-day periods from 2010 to 2014. The rms of the fits are in the range from 4.5 to 6.5 10^{-13}. There is, in average, six involved SLR stations
which data weight ranges from 0.95 to 0.15, depending on the performance of the used time system (H-maser, Cesium, Rubidium or oscillator).

The adjusted coefficient $\alpha_1$ (Table II) is systematically greater than the expected value given by the previous CNES study (factor of 2.5 in average). Whereas we do not detect any variation in the temperature coefficients ($\alpha_i$) over more than 4 years, we note however a slow decrease of the drift (see Table III) from 1.7 to 0.8 $\times 10^{-12}$ per °C, 5-6 $\times 10^{-12}$ per rad (with $\approx$ 5% of the total dose transmitted to the quartz), and a drift of 1.7 to 0.8 $\times 10^{-11}$ per day, from 2010 to 2014. The rms of several 10-day fits of the model is of $5 \times 10^{-13}$ in average.

This short-term modeling of the DORIS USO is an opportunity to develop the time transfer, by T2L2, between SLR stations over inter-continental distances (non Common View). The numerical integration of the model between two epochs would allow us to synchronize ground remote cloks (laser observatories) at the level of a few nanoseconds over 10,000 seconds.

The study of the DORIS USO behavior during 7 years is also an opportunity to describe the long-term variations of the frequency of such an oscillator, in the real space environment at 1335 km altitude; we expect to explain 99.9% of the effects.

**ACKNOWLEDGMENT**

The authors thank the CNES space agency for the mission support, and the astrogéo team of OCA for observing Jason-2 regularly. We thank also the laser stations (in Herstmonceux, Matera, Wettzell, Zimmerwald, MacDonald, Tokyo, Changchung, Yaragadee, notably), and the colleagues from Paris Observatory.
REFERENCES


<table>
<thead>
<tr>
<th>Author</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbé, Philippe</td>
<td>158, 728</td>
</tr>
<tr>
<td>Abdallah, Zeina</td>
<td>602</td>
</tr>
<tr>
<td>Abdel Hafiz, Moustafa</td>
<td>33, 171</td>
</tr>
<tr>
<td>Abedi, Maryam</td>
<td>139, 257, 522</td>
</tr>
<tr>
<td>Abgrall, Michel</td>
<td>239, 257</td>
</tr>
<tr>
<td>Ablewski, Piotr</td>
<td>304</td>
</tr>
<tr>
<td>Accadia, Timothée</td>
<td>343</td>
</tr>
<tr>
<td>Adamowicz, Waldemar</td>
<td>583</td>
</tr>
<tr>
<td>Adel, Amr</td>
<td>151</td>
</tr>
<tr>
<td>Affolderbach, Christoph</td>
<td>21, 800</td>
</tr>
<tr>
<td>Agatsuma, Shu</td>
<td>528</td>
</tr>
<tr>
<td>Ahmed, Ayman</td>
<td>151</td>
</tr>
<tr>
<td>Akgul, Mehmet</td>
<td>5</td>
</tr>
<tr>
<td>Alaric Schäffer, Stefan</td>
<td>357, 625</td>
</tr>
<tr>
<td>AlDawood, Khalid</td>
<td>254</td>
</tr>
<tr>
<td>Alford, Terry</td>
<td>111</td>
</tr>
<tr>
<td>AIIani, Maroua</td>
<td>100</td>
</tr>
<tr>
<td>Almleaky, Alaa</td>
<td>384</td>
</tr>
<tr>
<td>Almleaky, Hamzah</td>
<td>384</td>
</tr>
<tr>
<td>Almleaky, Yaseen M.</td>
<td>384</td>
</tr>
<tr>
<td>Alpay, Pamir</td>
<td>392</td>
</tr>
<tr>
<td>Amber, Michel</td>
<td>509</td>
</tr>
<tr>
<td>Ambrosini, Roberto</td>
<td>769</td>
</tr>
<tr>
<td>Atashbar, Massood Z.</td>
<td>536</td>
</tr>
<tr>
<td>Atef, Mohamed</td>
<td>151</td>
</tr>
<tr>
<td>Aubry, Jean-Pierre</td>
<td>690</td>
</tr>
<tr>
<td>Auriol, Albert</td>
<td>808</td>
</tr>
<tr>
<td>Auroux, Vincent</td>
<td>602</td>
</tr>
<tr>
<td>Badawy, Ahmed</td>
<td>151</td>
</tr>
<tr>
<td>Bai, Lina</td>
<td>133, 445</td>
</tr>
<tr>
<td>Bai, Long</td>
<td>742</td>
</tr>
<tr>
<td>Bai, Yu</td>
<td>747</td>
</tr>
<tr>
<td>Bakir, Ahmed</td>
<td>158, 343</td>
</tr>
<tr>
<td>Bale, Simon</td>
<td>423</td>
</tr>
<tr>
<td>Ballandras, Sylvain</td>
<td>214</td>
</tr>
<tr>
<td>Ban, Kazuhiro</td>
<td>174</td>
</tr>
<tr>
<td>Bandi, Thejesh</td>
<td>21</td>
</tr>
<tr>
<td>Baron, Thomas</td>
<td>100, 125, 214, 787</td>
</tr>
<tr>
<td>Bartoszek-Bobe, Dobroslaw</td>
<td>304</td>
</tr>
<tr>
<td>Basetas, Charis</td>
<td>448</td>
</tr>
<tr>
<td>Basetas, Charis</td>
<td>452</td>
</tr>
<tr>
<td>Bauch, Andreas</td>
<td>245, 320, 379, 643, 649</td>
</tr>
<tr>
<td>Bel, Olivier</td>
<td>100</td>
</tr>
<tr>
<td>Belfi, Jacopo</td>
<td>51</td>
</tr>
<tr>
<td>Belkadi, Nesrine</td>
<td>787</td>
</tr>
<tr>
<td>Belli, Alexandre</td>
<td>808</td>
</tr>
<tr>
<td>Bernard, Florent</td>
<td>787</td>
</tr>
<tr>
<td>Beverini, Nicolò</td>
<td>51</td>
</tr>
<tr>
<td>Bhave, Sunil</td>
<td>68</td>
</tr>
<tr>
<td>Bielska, Katarzyna</td>
<td>304</td>
</tr>
<tr>
<td>Biczewski, Artur</td>
<td>280, 583</td>
</tr>
<tr>
<td>Name</td>
<td>Pages</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Bize, Sébastien</td>
<td>257</td>
</tr>
<tr>
<td>Blazej, Josef</td>
<td>804</td>
</tr>
<tr>
<td>Bober, Marcin</td>
<td>304</td>
</tr>
<tr>
<td>Bogacki, Wojbor</td>
<td>583</td>
</tr>
<tr>
<td>Boillat, Sébastien</td>
<td>637</td>
</tr>
<tr>
<td>Bortolotti, Claudio</td>
<td>769</td>
</tr>
<tr>
<td>Boschen, Dan</td>
<td>752</td>
</tr>
<tr>
<td>Botteron, Cyril</td>
<td>690</td>
</tr>
<tr>
<td>Boudot, Rodolphe</td>
<td>33, 81, 171, 456</td>
</tr>
<tr>
<td>Bourgeois, Pierre-Yves</td>
<td>338, 672</td>
</tr>
<tr>
<td>Bourquin, Roger</td>
<td>158, 728</td>
</tr>
<tr>
<td>Bowen, Ross</td>
<td>76</td>
</tr>
<tr>
<td>Boy, Jean-Jacques</td>
<td>100</td>
</tr>
<tr>
<td>Braxmaier, Claus</td>
<td>47</td>
</tr>
<tr>
<td>Bregolin, Filippo</td>
<td>300</td>
</tr>
<tr>
<td>Buchman, Sasha</td>
<td>47</td>
</tr>
<tr>
<td>Buczek, Lukasz</td>
<td>583</td>
</tr>
<tr>
<td>Burt, Eric</td>
<td>188</td>
</tr>
<tr>
<td>Cabane, Hugues</td>
<td>100</td>
</tr>
<tr>
<td>Cahill, James</td>
<td>765</td>
</tr>
<tr>
<td>Caligaris, Massimo</td>
<td>681</td>
</tr>
<tr>
<td>Calonico, Davide</td>
<td>260, 300, 579, 769</td>
</tr>
<tr>
<td>Cambon, O.</td>
<td>100</td>
</tr>
<tr>
<td>Camparo, James</td>
<td>25, 37, 180, 474</td>
</tr>
<tr>
<td>Cantonni, Elena</td>
<td>260</td>
</tr>
<tr>
<td>Cao, Xiaotian</td>
<td>133</td>
</tr>
<tr>
<td>Cao, Yuanhong</td>
<td>265</td>
</tr>
<tr>
<td>Cárdenas-Olaya, Andrea Carolina</td>
<td>676</td>
</tr>
<tr>
<td>Carelli, Giorgio</td>
<td>51</td>
</tr>
<tr>
<td>Carter, Gary</td>
<td>765</td>
</tr>
<tr>
<td>Cassella, Cristian</td>
<td>709</td>
</tr>
<tr>
<td>Cerretto, Giancarlo</td>
<td>260, 320, 643, 649</td>
</tr>
<tr>
<td>Chandhoke, Sundeep</td>
<td>684</td>
</tr>
<tr>
<td>Chang, David</td>
<td>11, 76</td>
</tr>
<tr>
<td>Chang, Jung-Hao</td>
<td>202</td>
</tr>
<tr>
<td>Charmet, Jerome</td>
<td>209</td>
</tr>
<tr>
<td>Chen, Chao-Yu</td>
<td>155</td>
</tr>
<tr>
<td>Chen, Cheng-Chi</td>
<td>202</td>
</tr>
<tr>
<td>Chen, Faxi</td>
<td>133</td>
</tr>
<tr>
<td>Chen, Hui</td>
<td>402</td>
</tr>
<tr>
<td>Chen, Jianfeng</td>
<td>11</td>
</tr>
<tr>
<td>Chen, Jianping</td>
<td>773</td>
</tr>
<tr>
<td>Chen, Jingbiao</td>
<td>363, 611, 614, 618, 622</td>
</tr>
<tr>
<td>Chen, Mo</td>
<td>363</td>
</tr>
<tr>
<td>Chen, Weiliang</td>
<td>492, 742</td>
</tr>
<tr>
<td>Chiu, Wan-Cheng</td>
<td>155</td>
</tr>
<tr>
<td>Choi, Gobong</td>
<td>84</td>
</tr>
<tr>
<td>Chupin, Baptiste</td>
<td>257, 643, 649</td>
</tr>
<tr>
<td>Chutani, Ravinder</td>
<td>33</td>
</tr>
<tr>
<td>Cibiel, Gilles</td>
<td>121, 125, 158</td>
</tr>
<tr>
<td>Clairet, Alexandre</td>
<td>100, 214</td>
</tr>
<tr>
<td>Clement, Marta</td>
<td>117</td>
</tr>
<tr>
<td>Clivati, Cecilia</td>
<td>579, 769</td>
</tr>
<tr>
<td>Name</td>
<td>Page</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Feng Liu, Nian</td>
<td>492</td>
</tr>
<tr>
<td>Fernandez, Arnaud</td>
<td>602</td>
</tr>
<tr>
<td>Fluhr, Christophe</td>
<td>343</td>
</tr>
<tr>
<td>François, Bruno</td>
<td>81, 456</td>
</tr>
<tr>
<td>Friedt, Jean Michel</td>
<td>676</td>
</tr>
<tr>
<td>Frittelli, Matteo</td>
<td>579, 769</td>
</tr>
<tr>
<td>Fu, Guitao</td>
<td>742</td>
</tr>
<tr>
<td>Fu, Wei</td>
<td>129</td>
</tr>
<tr>
<td>Gaied, David</td>
<td>151</td>
</tr>
<tr>
<td>Galindo, Francisco Javier</td>
<td>320, 643, 649</td>
</tr>
<tr>
<td>Gallagher, Daniel</td>
<td>530</td>
</tr>
<tr>
<td>Gallagher, Mark</td>
<td>530</td>
</tr>
<tr>
<td>Galliou, Serge</td>
<td>125, 728</td>
</tr>
<tr>
<td>Gamal, Mostafa</td>
<td>151</td>
</tr>
<tr>
<td>Gao, Anming</td>
<td>1</td>
</tr>
<tr>
<td>Gao, Chao</td>
<td>562, 747</td>
</tr>
<tr>
<td>Gao, Yuanyuan</td>
<td>265</td>
</tr>
<tr>
<td>Gao, Zhe</td>
<td>236</td>
</tr>
<tr>
<td>Ge, Jun</td>
<td>541, 553</td>
</tr>
<tr>
<td>Gharavipour, Mohammadreza</td>
<td>800</td>
</tr>
<tr>
<td>Ghenna, Sofiane</td>
<td>509</td>
</tr>
<tr>
<td>Ghosh, Santunu</td>
<td>158</td>
</tr>
<tr>
<td>Ghosh, Siddhartha</td>
<td>72</td>
</tr>
<tr>
<td>Gibson, Brian</td>
<td>709</td>
</tr>
<tr>
<td>Giordano, Vincent</td>
<td>343</td>
</tr>
<tr>
<td>Giraud, Frédric</td>
<td>509</td>
</tr>
<tr>
<td>Giraud-Audine, Christophe</td>
<td>509</td>
</tr>
<tr>
<td>Govaec-Merou, Gwenhaël</td>
<td>672</td>
</tr>
<tr>
<td>Godone, Aldo</td>
<td>456</td>
</tr>
<tr>
<td>Goka, Shigeyoshi</td>
<td>162, 467</td>
</tr>
<tr>
<td>Gomah, Gihan</td>
<td>379</td>
</tr>
<tr>
<td>Gong, Hang</td>
<td>284</td>
</tr>
<tr>
<td>Gong, Songbin</td>
<td>1</td>
</tr>
<tr>
<td>Gong, Wei</td>
<td>622</td>
</tr>
<tr>
<td>Gorecki, Christophe</td>
<td>33</td>
</tr>
<tr>
<td>Goryachev, Maxim</td>
<td>728</td>
</tr>
<tr>
<td>Gosavi, Tanay</td>
<td>68</td>
</tr>
<tr>
<td>Götzte, Jens</td>
<td>106</td>
</tr>
<tr>
<td>Gourlat, Guillaume</td>
<td>222</td>
</tr>
<tr>
<td>Grabow Westergaard, Philip</td>
<td>357</td>
</tr>
<tr>
<td>Gray, Robert</td>
<td>316</td>
</tr>
<tr>
<td>Greve, Graham</td>
<td>351</td>
</tr>
<tr>
<td>Grop, Serge</td>
<td>343</td>
</tr>
<tr>
<td>Gruet, Florian</td>
<td>800</td>
</tr>
<tr>
<td>Guang, Wei</td>
<td>236</td>
</tr>
<tr>
<td>Guéna, Jocelyne</td>
<td>257</td>
</tr>
<tr>
<td>Guérandel, Stéphane</td>
<td>33, 456, 797</td>
</tr>
<tr>
<td>Guerlin, Christine</td>
<td>557</td>
</tr>
<tr>
<td>Guillemot, Philippe</td>
<td>121</td>
</tr>
<tr>
<td>Guo, Hong</td>
<td>611, 622</td>
</tr>
<tr>
<td>Guo, Wenge</td>
<td>599</td>
</tr>
<tr>
<td>Gürlebeck, Norman</td>
<td>47</td>
</tr>
<tr>
<td>Hall, John</td>
<td>713</td>
</tr>
<tr>
<td>Name</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Josefsson, Börje</td>
<td>276</td>
</tr>
<tr>
<td>Jourdan, Guillaume</td>
<td>222</td>
</tr>
<tr>
<td>Joyce, Richard</td>
<td>76</td>
</tr>
<tr>
<td>Kadota, Michio</td>
<td>412</td>
</tr>
<tr>
<td>Kajita, Masatoshi</td>
<td>162</td>
</tr>
<tr>
<td>Kalinin, Victor</td>
<td>498</td>
</tr>
<tr>
<td>Kalkur, Thottam</td>
<td>392</td>
</tr>
<tr>
<td>Kang, Songbai</td>
<td>800</td>
</tr>
<tr>
<td>Kanteres, Anthimos</td>
<td>452</td>
</tr>
<tr>
<td>Karlsson, Magnus</td>
<td>276</td>
</tr>
<tr>
<td>Kashiwada, Shinji</td>
<td>17</td>
</tr>
<tr>
<td>Kellogg, James</td>
<td>752</td>
</tr>
<tr>
<td>Kersalé, Yann</td>
<td>343</td>
</tr>
<tr>
<td>Khadem-Al-Charieh, Samy</td>
<td>384</td>
</tr>
<tr>
<td>Khairy, Mohamed</td>
<td>151</td>
</tr>
<tr>
<td>Khomenko, Igor</td>
<td>439</td>
</tr>
<tr>
<td>Kirby, Deborah</td>
<td>11, 76</td>
</tr>
<tr>
<td>Klimcak, Charles</td>
<td>180</td>
</tr>
<tr>
<td>Kobayashi, Tetsuo</td>
<td>174</td>
</tr>
<tr>
<td>Kochhar, Abhay</td>
<td>633</td>
</tr>
<tr>
<td>Kodet, Jan</td>
<td>804</td>
</tr>
<tr>
<td>Kolodziej, Jacek</td>
<td>583</td>
</tr>
<tr>
<td>Kosvin, Igor</td>
<td>752</td>
</tr>
<tr>
<td>Kosykh, Anatoly</td>
<td>136, 439</td>
</tr>
<tr>
<td>Krehlik, Przemyslaw</td>
<td>280, 583, 761</td>
</tr>
<tr>
<td>Kubena, Randall</td>
<td>11, 76</td>
</tr>
<tr>
<td>Kuna, Alexander</td>
<td>245</td>
</tr>
<tr>
<td>Kundermann, Stefan</td>
<td>594</td>
</tr>
<tr>
<td>Kushibiki, Jun-Ichi</td>
<td>793</td>
</tr>
<tr>
<td>Kwashin, Gennady</td>
<td>94, 396</td>
</tr>
<tr>
<td>Kwiatkowski, Pawel</td>
<td>575</td>
</tr>
<tr>
<td>Lämmertz, Claus</td>
<td>47</td>
</tr>
<tr>
<td>Langlois, Mehdi</td>
<td>456</td>
</tr>
<tr>
<td>Laroche, Thierry</td>
<td>214</td>
</tr>
<tr>
<td>Larson III, John</td>
<td>205</td>
</tr>
<tr>
<td>Laurent, Philippe</td>
<td>257, 557</td>
</tr>
<tr>
<td>Le Poncin-Laffite, Christophe</td>
<td>557</td>
</tr>
<tr>
<td>Lecomte, Steve</td>
<td>594</td>
</tr>
<tr>
<td>Leeson, David</td>
<td>332</td>
</tr>
<tr>
<td>Lemaire-Semail, Betty</td>
<td>509</td>
</tr>
<tr>
<td>Lemanski, Dariusz</td>
<td>583</td>
</tr>
<tr>
<td>Lenczner, Michel</td>
<td>338</td>
</tr>
<tr>
<td>Lepetaev, Alexander</td>
<td>136, 439</td>
</tr>
<tr>
<td>Lesage, Jean-Marc</td>
<td>100, 214</td>
</tr>
<tr>
<td>Leute, Julia</td>
<td>245</td>
</tr>
<tr>
<td>Levi, Filippo</td>
<td>260, 300, 456, 579, 769</td>
</tr>
<tr>
<td>Levine, Judah</td>
<td>655</td>
</tr>
<tr>
<td>Li, Cheng-Syun</td>
<td>202</td>
</tr>
<tr>
<td>Li, Dawei</td>
<td>465</td>
</tr>
<tr>
<td>Li, Ming-Huang</td>
<td>155</td>
</tr>
<tr>
<td>Li, Scott</td>
<td>5</td>
</tr>
<tr>
<td>Li, Sheng-Shian</td>
<td>155, 202</td>
</tr>
<tr>
<td>Li, Tianchu</td>
<td>492, 562</td>
</tr>
</tbody>
</table>
Li, Wei ......................................................................................................................... 236
Li, Wenhao .................................................................................................................... 622
Li, Xiaohui .................................................................................................................... 297
Li, Zhiqi ........................................................................................................................ 133, 541, 553
Liang, Kun ................................................................................................................. 284, 545, 742
Liao, Chia-Shu ............................................................................................................ 696
Liashuk, Alexei ........................................................................................................... 439
Lin, Chih-Ming ............................................................................................................. 432
Lin, Quanming ............................................................................................................. 518
Lin, Shinn-Yan ........................................................................................................... 226, 696
Lin Naing, Thura ........................................................................................................ 700
Lingley, Andrew ......................................................................................................... 205, 218
Lipa, John .................................................................................................................... 47
Lipinski, Marcin ......................................................................................................... 280, 583
Lisak, Daniel ............................................................................................................... 304
Liu, Chang ................................................................................................................... 471, 483
Liu, Kun ...................................................................................................................... 492
Liu, Li .......................................................................................................................... 307, 465
Liu, Ruonan ................................................................................................................ 5
Liu, Tao ......................................................................................................................... 293
Liu, Ya ........................................................................................................................ 297
Llopis, Olivier ............................................................................................................. 602
Lombardi, Michael A. ............................................................................................... 570
Lou, Janet .................................................................................................................... 90
Lu, Haoyuan ............................................................................................................... 462
Lu, Jun ........................................................................................................................ 297
Lu, Ruochen ............................................................................................................... 1
Lu, Xuanhui ................................................................................................................. 606
Luo, Bin ....................................................................................................................... 611
Ma, Tingfeng .............................................................................................................. 402, 406
Maccioni, Enrico ....................................................................................................... 51
Mailloux, David ......................................................................................................... 752
Malocha, Donald ........................................................................................................ 530
Mansour, Almonir ...................................................................................................... 392
Martin, Gilles ............................................................................................................ 214
Maslowski, Piotr ....................................................................................................... 304
Matsakis, Demetrios ................................................................................................. 717
Matsuda, Matsuda ...................................................................................................... 793
Maurice, Vincent ....................................................................................................... 33
McCants, Andrew ..................................................................................................... 752
Mejri, Sinda ................................................................................................................. 797
Melvin, Hugh ............................................................................................................... 684
Menyuk, Curtis .......................................................................................................... 736, 765
Meynadier, Frédéric ................................................................................................. 557
Miao, Kai ..................................................................................................................... 758
Micalizio, Salvatore ................................................................................................. 456, 676
Mikhail, Aleynikov ................................................................................................... 480
Mikhov, Mikhail ........................................................................................................ 117
Milani, Gianmaria ...................................................................................................... 300
Mileti, Gaetano ......................................................................................................... 21, 800
Mirea, Teona ............................................................................................................. 117
Mishin, Sergey .......................................................................................................... 777
Mizutani, Natsuhiko ................................................................................................. 174
<table>
<thead>
<tr>
<th>Name</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shi, Fan</td>
<td>290, 541, 553, 587</td>
</tr>
<tr>
<td>Shibata, Takayuki</td>
<td>17</td>
</tr>
<tr>
<td>Sicard, Gilles</td>
<td>222</td>
</tr>
<tr>
<td>Siccardi, Marco</td>
<td>239</td>
</tr>
<tr>
<td>Signorile, Giovanna</td>
<td>260</td>
</tr>
<tr>
<td>Silaghi, Marius Alexandru</td>
<td>62, 326</td>
</tr>
<tr>
<td>Simonelli, Andreino</td>
<td>51</td>
</tr>
<tr>
<td>Sinoussi, Nabil</td>
<td>151</td>
</tr>
<tr>
<td>Siu, Sammy</td>
<td>696</td>
</tr>
<tr>
<td>Skakun, Ivan</td>
<td>655</td>
</tr>
<tr>
<td>Skovbo Adersen, Sigrd</td>
<td>625</td>
</tr>
<tr>
<td>Sleewaegen, Jean-Marie</td>
<td>662</td>
</tr>
<tr>
<td>Sliwczynski, Lukasz</td>
<td>280, 583, 761</td>
</tr>
<tr>
<td>Smirnova, Elena</td>
<td>106</td>
</tr>
<tr>
<td>Sorokin, Boris</td>
<td>94, 396</td>
</tr>
<tr>
<td>Sotiriadis, Paul</td>
<td>448, 452, 667</td>
</tr>
<tr>
<td>Sotnikov, Andrey</td>
<td>106</td>
</tr>
<tr>
<td>Sridaran, Suresh</td>
<td>205, 218</td>
</tr>
<tr>
<td>Staliuniene, Egle</td>
<td>320, 379</td>
</tr>
<tr>
<td>Sthal, Fabrice</td>
<td>158</td>
</tr>
<tr>
<td>Stochino, Alberto</td>
<td>47</td>
</tr>
<tr>
<td>Stroinski, Maciej</td>
<td>583</td>
</tr>
<tr>
<td>Sun, Fuyu</td>
<td>129, 487</td>
</tr>
<tr>
<td>Sun, Guangfu</td>
<td>284</td>
</tr>
<tr>
<td>Sun, Tianchi</td>
<td>749</td>
</tr>
<tr>
<td>Sun, Xiaolin</td>
<td>758</td>
</tr>
<tr>
<td>Suo, Rui</td>
<td>492</td>
</tr>
<tr>
<td>Szplet, Ryszard</td>
<td>575</td>
</tr>
<tr>
<td>Tabata, Osamu</td>
<td>174</td>
</tr>
<tr>
<td>Takashi Røjle Christensen, Bjarke</td>
<td>357, 625</td>
</tr>
<tr>
<td>Takizawa, Yoshinori</td>
<td>17</td>
</tr>
<tr>
<td>Tan, Si</td>
<td>47</td>
</tr>
<tr>
<td>Tanaka, Shuji</td>
<td>17, 412, 633</td>
</tr>
<tr>
<td>Tang, Gongbin</td>
<td>416</td>
</tr>
<tr>
<td>Tang, Hao</td>
<td>193</td>
</tr>
<tr>
<td>Tarot, Jean-Marie</td>
<td>171</td>
</tr>
<tr>
<td>Tavella, Patrizia</td>
<td>260</td>
</tr>
<tr>
<td>Telichko, Arseniy</td>
<td>94, 396</td>
</tr>
<tr>
<td>Terao, Akira</td>
<td>174</td>
</tr>
<tr>
<td>Teshigahara, Akihiko</td>
<td>416, 633</td>
</tr>
<tr>
<td>Thompson, James</td>
<td>351</td>
</tr>
<tr>
<td>Thoumany, Pierre</td>
<td>300</td>
</tr>
<tr>
<td>Tjoelker, Robert</td>
<td>188</td>
</tr>
<tr>
<td>Tobar, Michael</td>
<td>728</td>
</tr>
<tr>
<td>Tompa, Gary</td>
<td>392</td>
</tr>
<tr>
<td>Treutlein, Philipp</td>
<td>21</td>
</tr>
<tr>
<td>Tricot, Francois</td>
<td>797</td>
</tr>
<tr>
<td>Tseng, Wen-Hung</td>
<td>226, 696</td>
</tr>
<tr>
<td>Tsujimoto, Kazuya</td>
<td>174</td>
</tr>
<tr>
<td>Turner, Kimberly</td>
<td>709</td>
</tr>
<tr>
<td>Turza, Krzysztof</td>
<td>280, 583</td>
</tr>
<tr>
<td>Uhrich, Pierre</td>
<td>239, 257, 643, 649</td>
</tr>
<tr>
<td>Underhill, Michael</td>
<td>343, 549</td>
</tr>
</tbody>
</table>
Vacheret, Xavier ................................................................. 100
Varadan, Siddharth .......................................................... 111
Vernier, David .................................................................. 158
Vernotte, Francois ............................................................. 338, 808
Villard, Patrick .................................................................. 222
Vorobyev, Nikolay .............................................................. 125
Vuillemin, Cedric ............................................................... 158
Wabbeke, Jared .................................................................. 536
Wang, Bo.......................................................................... 562, 747
Wang, Haidong................................................................. 622, 250, 290, 541, 553, 587
Wang, He .......................................................................... 369
Wang, Hongbo .................................................................. 250, 290, 541, 553, 587
Wang, James H-C. ............................................................. 310
Wang, Ji ............................................................................ 402, 406
Wang, Jijun ....................................................................... 562, 747, 758
Wang, Peng ....................................................................... 316
Wang, Qinghua .................................................................. 374
Wang, Qing-Ming ............................................................... 310, 518
Wang, Weibo ..................................................................... 545, 742
Wang, Xueyun ................................................................... 250
Wang, Yanhui .................................................................... 471, 483
Wang, Yigen ...................................................................... 307
Wang, Yuzhu ..................................................................... 495
Wang, Zenghui .................................................................. 316, 783
Wang, Zhong ................................................................... 307, 462, 465
Wcislo, Piotr ...................................................................... 304
Weaver, Gregory ................................................................. 145
Weeks, Arthur .................................................................. 530
Wei, Pei ............................................................................ 591
Wei, Rong ......................................................................... 495
Wei, Yajing ........................................................................ 236
Weihnacht, Manfred ........................................................... 106
Weiner, Joshua ................................................................ 351
Weiss, Marc ..................................................................... 684
Wells, Nathan ................................................................... 25
Westenkaer Thomsen, Jan .................................................. 357, 625
Whibberley, Peter ............................................................... 649
Winrow, Edward ................................................................. 752
Wojtewicz, Szymon .............................................................. 304
Wolf, Peter ....................................................................... 557
Wu, Guiling ....................................................................... 773
Wu, Huiyan ....................................................................... 310, 518
Wu, Meifang ..................................................................... 591
Wu, Teng .......................................................................... 622
Wu, Wenjun ..................................................................... 717
Xu, Longfei ....................................................................... 133
Xu, Yongliang .................................................................... 236
Xu, Zhichao ....................................................................... 363
Xu, Zhouxiang ................................................................... 606
Xuan, Meina ....................................................................... 445
Xue, Xiaobo ...................................................................... 363, 611, 614
Yamada, Ken ..................................................................... 528
Yamamoto, Yasuo ................................................................. 17, 633
Zou, Fan .............................................................................................................................................................................. 495
Zou, Jie ................................................................................................................................................................................ 432
Zu, Hongfei............................................................................................................................................................................. 310, 518
Zucco, Massimo ..................................................................................................................................................................... 769