IEEE
Signal Processing
MAGAZINE

[VOLUME 31 NUMBER 2 MARCH 2014]

DSP AND CODING FOR MULTI-Tb/s OPTICAL TRANSPORT
RECENT ADVANCES

MELODY EXTRACTION FROM POLYPHONIC MUSIC SIGNALS

ON THE RECURSIVE SOLUTION TO THE NORMAL EQUATIONS

HAPTICS AND SIGNAL PROCESSING

THE HOPPING DISCRETE FOURIER TRANSFORM

photo courtesy of the
U.S. Military & NASA

ULTRA-REL® *10 MHz to 6 GHz*

CERAMIC MMIC AMPLIFIERS

Low NF 0.5 dB High IP3 up to 38 dBm Low DC current 65 mA $4.95 ea. (qty 20)

When failure is not an option. Our new CMA MMIC amplifiers deliver outstanding performance in a rugged, nitrogen-filled, hermetic LTCC design, just 0.045" high. These models are so tough, they've qualified for use under MIL environmental conditions:

Robust performance across wide bandwidths makes them ideal for instrumentation, or anywhere long-term reliability adds bottom-line value. Go to minicircuits.com for all the details today, and get them in your hands as soon as tomorrow!

MIL Qualifications *(see website for complete list and details)*

Gross and Fine Leak	*HTOL (1700 hours+ @ +105°C)*
Mechanical Shock	*Thermal Shock*
Vibration	*Steam Aging*
Acceleration	*Solder Heat Resistance*
PIND	*Autoclave (and more)*

Electrical Specifications *(-55 to +105°C)*

3 x 3 x 1.14 mm

Model	Freq. (GHz)	Gain (dB)	P_{OUT} (dBm)	IP3 (dBm)	NF (dB)	DC (V)	Price $ ea. (qty 20)
CMA-62+	0.01-6	15	19	33	5	5	4.95
CMA-63+	0.01-6	20	18	32	4	5	4.95
CMA-545+	0.05-6	15	20	37	1	3	4.95
NEW CMA-5043+	0.05-4	18	20	33	0.8	5	4.95
NEW CMA-54SG1+	0.4-2.2	32	23	36	0.9	5	5.45
NEW CMA-162LN+	0.7-1.6	23	19	30	0.5	4	4.95
NEW CMA-252LN+	1.5-2.5	17	18	30	1	4	4.95

RoHS compliant

Mini-Circuits®

www.minicircuits.com P.O. Box 35166, Brooklyn, NY 11235-0003 (718) 934-4500 sales@minicircuits.com

503 Rev D

CONTENTS

[VOLUME 31 NUMBER 2]

SPECIAL SECTION—DSP AND CODING FOR MULTI-Tb/s OPTICAL TRANSPORT

15 **FROM THE GUEST EDITORS**
Ivan B. Djordjevic, William Shieh, Xiang Liu, and Moshe Nazarathy

16 **DIGITAL SIGNAL PROCESSING TECHNIQUES ENABLING MULTI-Tb/s SUPERCHANNEL TRANSMISSION**
Xiang Liu, S. Chandrasekhar, and Peter J. Winzer

25 **MIMO SIGNAL PROCESSING FOR MODE-DIVISION MULTIPLEXING**
Sercan Ö. Arık, Joseph M. Kahn, and Keang-Po Ho

35 **EFFICIENT CLOCK AND CARRIER RECOVERY ALGORITHMS FOR SINGLE-CARRIER COHERENT OPTICAL SYSTEMS**
Xiang Zhou

46 **DIGITAL FIBER NONLINEARITY COMPENSATION**
Liang B. Du, Danish Rafique, Antonio Napoli, Bernhard Spinnler, Andrew D. Ellis, Maxim Kuschnerov, and Arthur J. Lowery

57 **CYCLE SLIP COMPENSATION WITH POLARIZATION BLOCK CODING FOR COHERENT OPTICAL TRANSMISSION**
Tsuyoshi Yoshida, Takashi Sugihara, Kazuyuki Ishida, and Takashi Mizuochi

70 **SUBBANDED DSP ARCHITECTURES BASED ON UNDERDECIMATED FILTER BANKS FOR COHERENT OFDM RECEIVERS**
Moshe Nazarathy and Alex Tolmachev

82 **ADVANCED DSP TECHNIQUES ENABLING HIGH SPECTRAL EFFICIENCY AND FLEXIBLE TRANSMISSIONS**
Alan Pak Tao Lau, Yuliang Gao, Qi Sui, Dawei Wang, Qunbi Zhuge, Mohamed H. Morsy-Osman, Mathieu Chagnon, Xian Xu, Chao Lu, and David V. Plant

93 **CODED MODULATION FOR FIBER-OPTIC NETWORKS**
Lotfollah Beygi, Erik Agrell, Joseph M. Kahn, and Magnus Karlsson

104 **MULTIDIMENSIONAL SIGNALING AND CODING ENABLING MULTI-Tb/s OPTICAL TRANSPORT AND NETWORKING**
Ivan B. Djordjevic, Milorad Cvijetic, and Changyu Lin

FEATURE

118 **MELODY EXTRACTION FROM POLYPHONIC MUSIC SIGNALS**
Justin Salamon, Emilia Gómez, Daniel P.W. Ellis, and Gaël Richard

COLUMNS

4 **FROM THE EDITOR**
Peer Reviews: What Would Be the Best Model?
Abdelhak Zoubir

6 **PRESIDENT'S MESSAGE**
The Only Constant Is Change
Alex Acero

11 **SPECIAL REPORTS**
Touching Research: Haptics and Signal Processing
John Edwards

135 **SP TIPS & TRICKS**
The Hopping Discrete Fourier Transform
Chun-Su Park and Sung-Jea Ko

140 **LECTURE NOTES**
On the Recursive Solution to the Normal Equations: Some Other Results
Manuel D. Ortigueira

DEPARTMENTS

8 **SOCIETY NEWS**
SPS Fellows and Award Winners Recognized

143 **DATES AHEAD**

Digital Object Identifier 10.1109/MSP.2014.2299114

IEEE SIGNAL PROCESSING magazine

IEEE SIGNAL PROCESSING MAGAZINE

Abdelhak Zoubir—*Editor-in-Chief*
Technische Universität Darmstadt, Germany

AREA EDITORS

Feature Articles
Marc Moonen—KU Leuven, Belgium

Columns and Forum
Andrea Cavallaro—Queen Mary, University of London, UK
Andres Kwasinski—Rochester Institute of Technology

Special Issues
Fulvio Gini—University of Pisa, Italy

e-Newsletter
Z. Jane Wang, —University of British Columbia

EDITORIAL BOARD

Moeness G. Amin—Villanova University
Sergio Barbarossa—University of Rome "La Sapienza," Italy
Mauro Barni—Università di Siena, Italy
Helmut Bölcskei—ETH Zürich, Switzerland
A. Enis Cetin—Bilkent University, Turkey
Yen-Kuang Chen—Intel Corporation
Bernd Girod—Stanford University
Michael Jordan—University of California, Berkeley
Bastiaan Kleijn—Victoria University of Wellington, New Zealand
Visa Koivunen—Aalto University, Finland
Ying-Chang Liang—Institute for Infocomm Research, Singapore
Mark Liao —National Chiao-Tung University, Taiwan
Hongwei Liu—Xidian University, China
V. John Mathews—University of Utah
Satoshi Nakamura—Nara Institute of Science and Technology, Japan
Kuldip Paliwal—Griffith University, Australia
Béatrice Pesquet-Popescu—Télécom ParisTech, France
H. Vincent Poor—Princeton University
Eli Saber—Rochester Institute of Technology
Ali Sayed—University of California, Los Angeles
Victor Solo—University of New South Wales, Australia
Sergios Theodoridis—University of Athens, Greece
Michail K. Tsatsanis—Entropic Communications
Min Wu—University of Maryland

ASSOCIATE EDITORS—COLUMNS AND FORUM

Rodrigo Capobianco Guido — University of Sao Paulo, Brazil
Aleksandra Mojsilovic — IBM T.J. Watson Research Center
Douglas O'Shaughnessy — INRS, Canada
Gene Cheung — National Institute of Informatics
Alessandro Vinciarelli — IDIAP–EPFL
Michael Gormish — Ricoh Innovations, Inc.
Xiaodong He — Microsoft Research
Fatih Porikli — MERL
Stefan Winkler — UIUC/ADSC, Singapore
Saeid Sanei, — University of Surrey, UK
Azadeh Vosoughi — University of Central Florida
Danilo Mandic — Imperial College, UK
Roberto Togneri — The University of Western Australia
Gail Rosen — Drexel University

ASSOCIATE EDITORS—E-NEWSLETTER

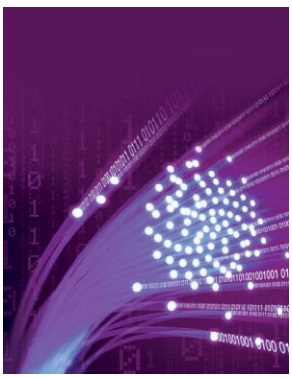
Gwenael Doerr—Technicolor, France
Vítor Nascimento—University of São Paulo, Brazil
Shantanu Rane—MERL
Yan Lindsay Sun—University of Rhode Island

IEEE SIGNAL PROCESSING SOCIETY

Alex Acero—*President*
Rabab Ward—*President-Elect*
Konstantinos (Kostas) N. Plataniotis—*Vice President, Awards and Membership*
Wan-Chi Siu—*Vice President, Conferences*
Alex Kot—*Vice President, Finance*
Mari Ostendorf —*Vice President, Publications*
Charles Bouman —*Vice President, Technical Directions*

COVER

©ISTOCKPHOTO.COM/HENRIK5000



IEEE PERIODICALS MAGAZINES DEPARTMENT

Jessica Barragué
Managing Editor

Geraldine Krolin-Taylor
Senior Managing Editor

Susan Schneiderman
Business Development Manager
+1 732 562 3946 Fax: +1 732 981 1855

Felicia Spagnoli
Advertising Production Manager

Janet Dudar
Senior Art Director

Gail A. Schnitzer
Assistant Art Director

Theresa L. Smith
Production Coordinator

Dawn M. Melley
Editorial Director

Peter M. Tuohy
Production Director

Fran Zappulla
Staff Director, Publishing Operations

IEEE prohibits discrimination, harassment, and bullying. For more information, visit http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

IEEE SIGNAL PROCESSING MAGAZINE (ISSN 1053-5888) (ISPREG) is published bimonthly by the Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, 17th Floor, New York, NY 10016-5997 USA (+1 212 419 7900). Responsibility for the contents rests upon the authors and not the IEEE, the Society, or its members. Annual member subscriptions included in Society fee. Nonmember subscriptions available upon request. Individual copies: IEEE Members US$20.00 (first copy only), nonmembers US$201.00 per copy. Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of U.S. Copyright Law for private use of patrons: 1) those post-1977 articles that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 USA; 2) pre-1978 articles without fee. Instructors are permitted to photocopy isolated articles for noncommercial classroom use without fee. For all other copying, reprint, or republication permission, write to IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854 USA. Copyright ©2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals postage paid at New York, NY, and at additional mailing offices. Postmaster: Send address changes to IEEE Signal Processing Magazine, IEEE, 445 Hoes Lane, Piscataway, NJ 08854 USA. Canadian GST #125634188

Printed in the U.S.A.

Digital Object Identifier 10.1109/MSP.2014.2299115

SUSTAINABLE FORESTRY INITIATIVE Certified Chain of Custody At Least 25%Certified Forest Content www.sfiprogram.org SFI-01042

SCOPE: *IEEE Signal Processing Magazine* publishes tutorial-style articles on signal processing research and applications, as well as columns and forums on issues of interest. Its coverage ranges from fundamental principles to practical implementation, reflecting the multidimensional facets of interests and concerns of the community. Its mission is to bring up-to-date, emerging and active technical developments, issues, and events to the research, educational, and professional communities. It is also the main Society communication platform addressing important issues concerning all members.

Now...

2 Ways to Access the IEEE Member Digital Library

With two great options designed to meet the needs—and budget—of every member, the IEEE Member Digital Library provides full-text access to any IEEE journal article or conference paper in the IEEE *Xplore*® digital library.

Simply choose the subscription that's right for you:

IEEE Member Digital Library

Designed for the power researcher who needs a more robust plan. Access all the IEEE content you need to explore ideas and develop better technology.

- 25 article downloads every month

IEEE Member Digital Library Basic

Created for members who want to stay up-to-date with current research. Access IEEE content and rollover unused downloads for 12 months.

- 3 new article downloads every month

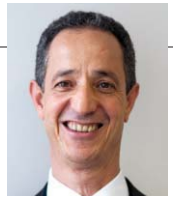
Get the latest technology research.

Try the IEEE Member Digital Library—FREE!
www.ieee.org/go/trymdl

IEEE
Advancing Technology for Humanity

IEEE Member Digital Library is an exclusive subscription available only to active IEEE members.

[from the EDITOR]



Abdelhak Zoubir
Editor-in-Chief
zoubir@spg.tu-darmstadt.de
http://signalprocessingsociety.org/publications/periodicals/spm

Peer Reviews: What Would Be the Best Model?

This editorial will appear in print approximately two months from now, i.e., about one month after the acceptance of papers for the 39th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014) has been disseminated. The last few months of 2013 were very busy and hectic for many of us as we finished paper submissions for ICASSP 2014, which will be held 4–9 May 2014 in Florence, Italy. We all felt relieved once the papers had been uploaded before the due date of 4 November 2013. But for many of us, it was the beginning of even more hard work, reviewing dozens of papers quickly so that the authors can get timely reviews and find out whether they can present their work at ICASSP, the most comprehensive technical conference on signal processing and its applications—which is also the flagship conference of the IEEE Signal Processing Society.

As of now, the review process is ongoing and should be concluded by 7 January. Experience shows, however, that the due date of completing the reviews is rarely respected by many of our esteemed colleagues, and those responsible for the coordination of reviews and the technical program of the conference must organize "speedy" reviews in a short and most critical time.

The current peer-review system has some drawbacks—this does not only apply to conference papers but also to journal submissions. The major challenge is to find experienced reviewers who would be able to complete reviews in the short time frame provided.

Digital Object Identifier 10.1109/MSP.2013.2295043
Date of publication: 12 February 2014

If one checked the "IEEE Signal Processing Society Guide for Reviewers" [1], one would read, "You are asked to complete the review within six weeks of receipt of the manuscript. Please note the peer-review cycle for *IEEE Signal Processing Letters* is three weeks." There is no doubt that the shorter the review time allocated to experts is, the better it is for the authors and the journals in general; every one of us would like to see his or her fresh, original ideas published as soon as possible, and every editor wishes a minimum turnaround time of paper reviews. The challenge for editors is therefore to secure comprehensive reviews so as to make an expert decision, acceptable to the authors, and in the shortest amount of time.

A central question has always been a topic of discussion at many editorial board meetings I have attended over the years: "How can one secure high-quality reviews by experienced researchers?" Most previous proposals to address this problem were based on an incentive principle. This makes sense, but one immediately comes to the next question: "What incentive would have an effect to change the current situation?" For example, what may appear to be an incentive to a young professional building his or her career may be unattractive to a senior and well-established colleague.

I recall that several years ago a discussion was held regarding the building of a network of selected reviewers who would perform reviews based on a small remuneration. The idea was quite attractive for numerous professionals at that time, but it never materialized. It is not uncommon that many professionals perform remuneration-based reviews for certain institutions on promotion or grant applications. The system works for them, and it is possible to secure experienced professionals for these tasks. The remuneration-based review is also a common practice for reviews of research centers and the like.

Many nonprofit professional organizations such as the IEEE work with volunteers, and a remuneration-based approach would not be an option: there are several reasons for its justification. The question of concern to all of us will remain: "Which model would work to ensure that 'busy' experts would immediately agree to perform a paper review, and that the task would be completed in the minimum given time?" I believe that there is no simple answer to this question, as we all work under constraints, either imposed by our institutions or professional organizations. We will continue to struggle to find three reviewers for each single submitted paper and, in a situation like ICASSP 2014, with a number of over 3,500 submissions, simple mathematics will show that this is an overwhelming task for all involved.

I take this opportunity to wholeheartedly thank all of the colleagues and friends who are always willing to help review articles for *IEEE Signal Processing Magazine*. It is always a great feeling to know that there are many busy professionals who will always find the time despite their busy schedules to assist with reviews to ensure a high-quality publication.

REFERENCE

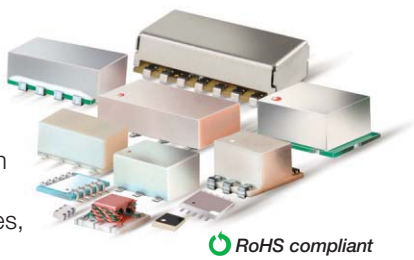
[1] IEEE Signal Processing Society Guide for Reviewers. [Online]. Available: http://www.signalprocessingsociety.org/publications/guides/guide-for-reviewers

[SP]

90° SPLITTERS

5 MHz to 8 GHz

from $1.99 ea. qty. 1000



RoHS compliant

With over 70 different models, our two-way 90° splitters make the perfect building blocks for many designs including balanced amplifiers, IQ modulator/demodulators, single sideband modulators, image rejection mixers, voltage variable attenuators, phase shifters, and more! Use them for signal processing designs requiring 90° phase offset or to insulate your circuit from reflective elements. The industry's widest range of frequencies, extra low amplitude and phase unbalance, and packages as small as 0.08 x 0.05" make these hybrids essential tools for your RF design toolbox.

Tiny, robust, low-cost LTCC models are now available in small quantity reels, with standard counts of 20, 50, 100, 200, 500, 1000, or 2000 at no extra cost! For full performance details and product availability, visit ***minicircuits.com***.

Order online today and have units in-hand as soon as tomorrow!

Mini-Circuits®

www.minicircuits.com P.O. Box 35166, Brooklyn, NY 11235-0003 (718) 934-4500 sales@minicircuits.com

463 rev J

[president's MESSAGE]

Alex Acero
2014–2015 SPS President
a.acero@ieee.org

The Only Constant Is Change

This is my first column as the president of the IEEE Signal Processing Society (SPS), and I am thrilled to serve all the members of this vigorous engineering Society. Part of the reason SPS is so vibrant is because of its volunteers, so I would like to offer my most sincere token of appreciation to all the volunteers whose term ended in 2013. First, I would like to thank Past President Ray Liu for his impressive dedication to the Society and the vast number of initiatives he has pushed through during his two-year tenure. I have big shoes to fill. I would also like to express my gratitude to our first Vice President, Awards and Membership, John Treichler, for getting the Membership Board off the ground. I'd also like to thank Petar Djuric, Clem Karl, and Sergios Theodoridis for their service as members-at-large of the Society's Board of Governors; Roxana Saint-Nom and Ta-Sung Lee for breaking ground as the Society's first regional directors; and IEEE Division IX Director José Moura for representing the SPS on the IEEE Board of Directors during the past two years.

My job as president is to lead a talented and energetic team of volunteers on our boards and committees and shepherd the Society through these times of change. I'd like to welcome Rabab Ward as our incoming president-elect and Kostas Plataniotis as our incoming vice president, membership. Rabab and Kostas will be part of the SPS Executive Committee—the group of volunteer leaders that guides the Society's operations—together with Vice President, Publications Mari Ostendorf; Vice President, Conferences Wan-Chi Siu; Vice President, Technical Directions Charlie Bouman; and Vice President, Finance Alex Kot. I'd also like to welcome Geert Leus, Helen Meng, and Fernando Pereira as incoming members-at-large of the Board of Governors and Doug O'Shaughnessy and Hong-Yuan Mark Liao as incoming regional directors. I also thank José Moura for becoming the Awards Board chair. I am very fortunate to count on the help of SPS Executive Director Rich Baseil and his professional staff.

For me, the journey through the IEEE started 30 years ago, when I was an undergraduate student at the Polytechnic University of Madrid in Spain looking for a thesis project topic. At an open house in our department, I saw a demo of a talking calculator for the blind and a speech analysis tool to help deaf children vocalize better. I immediately knew I wanted my project to focus on speech recognition. One of my professors told me about the IEEE so I joined in 1984 as an undergraduate Student Member. At that time, I also joined the IEEE Acoustics, Speech, and Signal Processing Society, which, in 1989, was renamed as the IEEE SPS. I continued as a graduate student member, first at Rice University and later at Carnegie Mellon. Once I obtained my Ph.D. degree, I became an IEEE Member, and later IEEE Senior Member and IEEE Fellow. I have been a volunteer for the IEEE SPS for many years, first handing out the heavy four-volume proceedings at the registration desk at the International Conference on Acoustics, Speech, and Signal Processing (ICASSP '87) in Dallas, Texas. Years later I became a reviewer of some of our Society's conferences and transactions and then as an associate editor of *IEEE Transactions on Speech and Audio Processing*. I was then invited to join the Speech Technical Committee and later became its chair. I have also served on the Society's Board of Governors as a member-at-large; director of industrial relations; vice president, technical directions; and president-elect.

During my 30 years as a member of the Society, I have seen many changes. I wrote my engineering degree's thesis using one of the first word processors running on a minicomputer. Yet, for my first conference submission (ICASSP '90) I had to print each column separately, glue them on a large paper, and mail it with a cardboard backer. As publications chair for ICASSP '98, I introduced the concept of electronic submission to ICASSP for the first time so that authors could submit their pdf file via a Web site. Not only has the paper submission process changed, but manuscripts also went from being mailed to reviewers to being available for download on a Web site. Soon after joining SPS in 1984, my mom started complaining about the rapidly shrinking space on my small shelf taken by *IEEE Transactions on Acoustics, Speech, and Signal Processing*, which, back then, was the only transactions in SPS. Over time, that single transactions became four separate transactions, so I stopped getting the hard copy years ago, as everything I wanted is available in IEEE *Xplore*. The information technology revolution has changed not only our lives but also the way SPS does business. As Heraclitus once said, "The only constant is change," so I welcome your suggestions on how to continue to change SPS in the 21st century.

[SP]

Digital Object Identifier 10.1109/MSP.2013.2285053
Date of publication: 12 February 2014

48th Annual Asilomar Conference on Signals, Systems, and Computers

November 2-5, 2014

www.asilomarssc.org

Submit papers by May 1, 2014 in the following areas:

Architecture and Implementation
Array Signal Processing
Biomedical Signal and Image Processing
Communications Systems
MIMO Communications and Signal Processing
Networks
Signal Processing and Adaptive Systems
Speech, Image and Video Processing

IEEE Signal Processing Society

General Chair: Roger Woods, *Queen's University of Belfast*
Technical Program Chair: Geert Leus, *Delft University of Technology*
Conference Coordinator: Monique P. Fargues, *Naval Postgraduate School*
Publication Chair: Michael Matthews, *ATK Space Systems*
Publicity Chair: Linda S. DeBrunner, *Florida State University*
Finance Chair: Ric Romero, *Naval Postgraduate School*
Electronic Media Chair: Marios S. Pattichis, *University of New Mexico*

The Conference is organized by the non-profit Signals, Systems and Computers Conference Corporation.

The conference will be held at the **Asilomar Conference Grounds**, *in Pacific Grove, CA. The grounds border the Pacific Ocean and are close to Monterey, Carmel, and the Seventeen Mile Drive in Pebble Beach.*

Photo Credit: L. S. DeBrunner

society NEWS

SPS Fellows and Award Winners Recognized

In this column of *IEEE Signal Processing Magazine*, 41 IEEE Signal Processing Society (SPS) members are recognized as Fellows, the IEEE SPS Italian Chapter receives the 2013 Chapter of the Year Award, and award recipients are announced.

41 SPS MEMBERS ELEVATED TO FELLOW

Each year, the IEEE Board of Directors confers the grade of Fellow on up to one-tenth of 1% of the Members. To qualify for consideration, an individual must have been a Member, normally for five years or more, and a Senior Member at the time of nomination to Fellow. The grade of Fellow recognizes unusual distinction in IEEE's designated fields.

The IEEE SPS congratulates the following 41 SPS members who were recognized with the grade of Fellow as of 1 January 2014.

Mohammad Alam, Mobile, Alabama, United States: For contributions to pattern recognition and high-resolution image reconstruction.

Alejandro Frangi Caregnato, Sheffield, United Kingdom: For contributions to medical image analysis and image-based computational physiology.

Tihao Chiang, Hsinchu, Taiwan: For contributions to the theory and applications of video coding algorithms.

Shuguang Cui, College Station, Texas, United States: For contributions to cognitive communications and energy-efficient system design.

Minh Do, Urbana, Illinois, United States: For contributions to image representation and computational imaging.

Aly Farag, Louisville, Kentucky, United States: For contributions to image modeling and biomedical applications.

James Glass, Cambridge, Massachusetts, United States: For contributions to probabilistic segment-based speech recognition and spoken dialogue interfaces.

Vivek Goyal, Cambridge, Massachusetts, United States: For contributions to information representations and their applications in acquisition, communication, and estimation.

Remi Gribonval, Rennes Cedex, France: For contributions to the theory and applications of sparse signal processing.

Martin Haenggi, Notre Dame, Indiana, United States: For contributions to the spatial modeling and analysis of wireless networks.

Dilek Hakkani-Tur, Los Altos, California, United States: For contributions to spoken language processing.

Yun He, Beijing, China: For contributions to video coding and communication technologies.

Syed Jafar, Irvine, California, United States: For contributions to analyzing the capacity of wireless communication networks.

Mohan Kankanhalli, Singapore: For contributions to multimedia content processing and security.

W. Clem Karl, Boston, Massachusetts, United States: For contributions to statistical signal processing and image reconstruction.

Sam Kwong, Hong Kong, China: For contributions to optimization techniques in cybernetics and video coding.

J. Laneman, Notre Dame, Indiana, United States: For contributions to multihop relaying and cooperative communication for wireless networks.

Haizhou Li, Singapore: For leadership in multilingual speaker and language recognition.

Yu Morton, Oxford, Ohio, United States: For contributions to the understanding of ionospherice effects on global navigation satellite signals.

Haldun Ozaktas, Bilkent, Turkey: For contributions to transforms For signal processing in optics.

Fatih Porikli, Cambridge, Massachusetts, United States: For contributions to computer vision and video surveillance.

Jinyi Qi, Davis, California, United States: For contributions to statistical image reconstruction for emission-computed tomography.

Stephen Renals, Edinburgh, United Kingdom: For contributions to speech recognition technology and its use in spoken language processing.

Amir Said, Cupertino, California, United States: For contributions to compression and processing of images and videos.

Mark Sandler, London, United Kingdom: For contributions to digital signal processing techniques in audio and music applications.

Guillermo Sapiro, Durham, North Carolina, United States: For contributions to computational mathematics for computer vision.

Philip Schniter, Columbus, Ohio, United States: For contributions to signal processing in communications.

Richard Stern, Pittsburgh, Pennsylvania, United States: For contributions to robust speech recognition and auditory perception.

Anthony Tether, Falls Church, Virginia, United States: For leadership in the advancement of commercial and defense technologies.

Digital Object Identifier 10.1109/MSP.2013.2294947
Date of publication: 12 February 2014

Keiichi Tokuda, Nagoya, Aichi, Japan: For contributions to hidden Markov model-based speech synthesis.

Trac Tran, Baltimore, Maryland, United States: For contributions to multirate and sparse signal processing.

Wade Trappe, North Brunswick, New Jersey, United States: For contributions to information and communication security.

Marc Van Hulle, Leuven, Belgium: For contributions to biomedical signal processing and biological modeling.

Rene Vidal, Baltimore, Maryland, United States: For contributions to subspace clustering and motion segmentation in computer vision.

Zhou Jane Wang, Waterloo, Canada: For contributions to perceptual image processing and quality assessment.

Iram Weinstein, McLean, Virginia, United States: For leadership in signal processing and test methods for radars detecting advanced aircraft and cruise missiles in severe terrain clutter.

Sarah Wilson, Santa Clara, California, United States: For contributions to orthogonal frequency division multiplexing.

Boon-Lock Yeo, Sunnyvale, California, United States: For contributions to and leadership in image and video processing.

Wei Yu, Toronto, Canada: For contributions to optimization techniques for multiple-input, multiple-output communications.

Wei-Xing Zheng, Penrith, Australia: For contributions to signal processing and system identification.

Shengli Zhou, Storrs, Connecticut, United States: For contributions to wireless and underwater acoustic communications.

IEEE SPS ITALIAN CHAPTER RECEIVES THE 2013 CHAPTER OF THE YEAR AWARD

The IEEE SPS Italian Chapter has been selected as the third recipient of the 2013 Chapter of the Year Award.

The Chapter of the Year Award will be presented at the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) Awards Ceremony in Florence, Italy, 4–9 May. The award is presented annually to a Chapter that has provided their membership with the highest quality of programs, activities, and services. The SPS Italian Chapter will receive a certificate and a check in the amount of US$1,000 to support local Chapter activities. The Chapter will also publish an article in a future issue of *IEEE Inside Signal Processing eNewsletter*.

2013 IEEE SPS AWARDS TO BE PRESENTED IN FLORENCE, ITALY

The IEEE SPS congratulates the following SPS members who will receive the Society's prestigious awards during ICASSP 2014 in Florence, Italy.

The Society Award honors outstanding technical contributions in a field within the scope of the IEEE SPS and outstanding leadership within that field. The Society Award comprises a plaque, a certificate, and a monetary award of US$2,500. It is the highest-level award bestowed by the IEEE SPS. This year's recipients are Al Bovik "for fundamental contributions to digital image processing theory, technology, leadership, and education" and Alan S. Willsky "for fundamental contributions to probabilistic modeling and for pioneering work in the development and application of multiresolution statistical methods."

The IEEE Signal Processing Magazine Best Paper Award honors the author(s) of an article of exceptional merit and broad interest on a subject related to the Society's technical scope and appearing in the Society's magazine. The prize comprises US$500 per author (up to a maximum of US$1,500 per award) and a certificate. In the event that there are more than three authors, the maximum prize shall be divided equally among all authors, and each shall receive a certificate. This year, the IEEE Signal Processing Magazine Best Paper Award recipients are Zhou Wang and Alan C. Bovik for their article "Mean Squared Error: Love it or Leave it? A New Look at Signal Fidelity Measures," published in *IEEE Signal Processing Magazine*, vol. 26, no. 1, Jan. 2009.

Two Technical Achievement Awards will be presented this year. Yonina Eldar will receive the award "for fundamental contributions to sub-Nyquist and compressed sampling, convex optimization and statistical signal processing." Al Hero will be recognized "for information-theoretic advances in statistical signal processing and machine learning." The Technical Achievement Award honors a person who, over a period of years, has made outstanding technical contributions to theory and/or practice in technical areas within the scope of the Society, as demonstrated by publications, patents, or recognized impact on this field. The prize for the award is US$1,500, a plaque, and a certificate.

Two Meritorious Service Awards will be presented this year to Ali H. Sayed and Rabab Ward "for exemplary service to and leadership in the IEEE SPS." The award comprises a plaque and a certificate; judging is based on dedication, effort, and contributions to the Society.

The SPS Education Award honors educators who have made pioneering and significant contributions to signal processing education. Judging is based on a career of meritorious achievement in signal processing education as exemplified by the writing of scholarly books and texts, course materials, and papers on education; inspirational and innovative teaching; and creativity in the development of new curricula and methodology. The award comprises a plaque, a monetary award of US$1,500, and a certificate. The recipient of the SPS Education Award is Dimitris Manolakis, "for fundamental contributions to education in signal processing and algorithms for adaptive filtering and hyperspectral target."

Five Best Paper Awards will be awarded, honoring the author(s) of a paper of exceptional merit dealing with a subject related to the Society's technical scope and appearing in one of the Society's transactions, irrespective of the author's age. The prize is US$500 per author (up to a maximum of US$1,500 per award) and a certificate. Eligibility is based on a five-year window preceding the year of election, and judging is based on general quality, originality, subject matter, and timeliness. Up to six Best Paper Awards may be presented each year. This year, the awardees are:

- Amir Beck, Petre Stoica, and Jian Li, "Exact and Approximate Solutions of Source Localization Problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, May 2008

- Matthew A. Herman and Thomas Strohmer, "High-Resolution Radar via Compressed Sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, June 2009
- Chunming Li, Chiu-Yen Kao, John C. Gore, and Zhaohua Ding, "Minimization of Region-Scalable Fitting Energy for Image Segmentation," *IEEE Transactions on Image Processing*, vol. 17, no. 10, Oct. 2008
- Robert W. Heath, Jr., Tao Wu, Young Hoon Kwon, and Anthony C.K. Soong, "Multiuser MIMO in Distributed Antenna Systems with Out-of-Cell Interference," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, Oct. 2011
- George E. Dahl, Dong Yu, Li Deng, and Alex Acero, "Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 1, Jan. 2012.

The Young Author Best Paper Award honors the author(s) of an especially meritorious paper dealing with a subject related to the Society's technical scope and appearing in one of the Society's transactions and who, upon date of submission of the paper, is fewer than 30 years of age. The prize is US$500 per author (up to a maximum of US$1,500 per award) and a certificate. Eligibility is based on a three-year window preceding the year of election, and judging is based on general quality, originality, subject matter, and timeliness. Three Young Author Best Paper Awards are being presented this year:

- Yuejie Chi, for the paper coauthored with Louis L. Scharf, Ali Pezeshki, and A. Robert Calderbank, "Sensitivity to Basis Mismatch in Compressed Sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, May 2011
- Kalpana Seshadrinathan, for the paper coauthored with Alan Conrad Bovik, "Motion Tuned Spatio-Temporal Quality Assessment of Natural Videos," *IEEE Transactions on Image Processing*, vol. 19, No. 2, Feb. 2010
- Lin Li, for the paper coauthored with Anna Scaglione and Jonathan H. Manton, "Distributed Principal Subspace Estimation in Wireless Sensor Networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, Aug. 2011.

One IEEE Signal Processing Letters Best Paper Award will be awarded, honoring the author(s) of a letter article of exceptional merit and broad interest on a subject related to the Society's technical scope and appearing in *IEEE Signal Processing Letters*. The prize shall consist of US$500 per author (up to a maximum of US$1,500 per award) and a certificate. To be eligible for consideration, an article must have appeared in *IEEE Signal Processing Letters* in an issue predating the Spring Awards Board meeting by five years (typically held in conjunction with ICASSP). Judging shall be on the basis of the technical novelty, the research significance of the work, quality, and effectiveness in presenting subjects in an area of high impact to the Society's members. The recipient of the IEEE Signal Processing Letters Best Paper Award is:

- Gan Zheng, Kai-Kit Wong, Arogyaswami Paulraj, and Bjorn Ottersten, "Collaborative-Relay Beamforming with Perfect CSI: Optimum and Distributed Implementation," *IEEE Signal Processing Letters*, vol. 16, no. 4, Apr. 2009.

SPS MEMBERS RECEIVE 2014 IEEE AWARDS

The following SPS members will receive 2014 IEEE Awards.

Thomas P. Barnwell, III (Georgia Institute of Technology, Vinings) has been selected as the IEEE Jack S. Kilby Signal Processing Medal recipient "for leadership in and contributions to speech processing, filter banks and wavelets, DSP hardware and architectures, and technology-enhanced education." The medal will be presented to Dr. Barnwell at the IEEE Honors Ceremonies.

The IEEE James H. Mulligan, Jr. Education Medal, which recognizes a career of outstanding contributions to education in the fields of interest of the IEEE, will be presented to John G. Proakis (Northeastern University, Winchester, Massachusetts) "for contributions to electrical engineering education through influential textbooks and inspiring leadership in integrating research and education."

The IEEE Dennis J. Picard Medal for Radar Technologies and Applications recognizes outstanding accomplishments in advancing the fields of radar technologies and their applications. The 2014 recipient is Yuri Abramovich (WR Systems Ltd., Fairfax, Virginia) "for seminal contributions to adaptive radar signal processing algorithms and over-the-horizon radar."

The IEEE W.R.G. Baker Paper Award recognizing the most outstanding paper reporting original work published in any IEEE archival publications (such as transactions, journals, and letters), magazines, or proceedings is awarded to Robert Nowak and Mário A.T. Figueiredo for their paper, "Sparse Reconstruction by Separable Approximation," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, July 2009.

The IEEE James L. Flanagan Speech and Audio Processing Technical Field Award will be presented to Biing-Hwang Juang (Georgia Institute of Technology, Atlanta) "for pioneering contributions to automatic speech recognition and speech coding." This award was founded and is sponsored by the IEEE SPS.

The IEEE Kiyo Tomiyasu Award recognizes outstanding early to midcareer contributions to technologies holding the promise of innovative applications. The award is being presented to George Chrisikos (Qualcomm Inc., San Diego, California) "for contributions to heterogeneous network architectures with ubiquitous wireless access."

The IEEE Leon K. Kirchmayer Graduate Teaching Award recognizing inspirational teaching of graduate students in the IEEE fields of interest will be presented to John M. Cioffi (Stanford University, Stanford, California) "for educating a stellar array of graduate students in digital communications and for inspiring them to make a difference."

[SP]

[special REPORTS]

John Edwards

Touching Research: Haptics and Signal Processing

The science of haptics gets its name from the Greek word *haptikos*, which means the ability to grasp or touch.

Haptics researchers investigate ways of using tactile sensation and control to help people interact more intuitively with computers, robots, and various other types of electronic and electromechanical systems. Through the use of specially designed input/output devices, such as haptic joysticks, gloves, and surfaces, users receive feedback in the form of sensations that are felt in a finger, hand, or other part of the body. Haptic devices typically use resistance motors to supply force, vibration, or motion feedback.

When used in conjunction with a video display, haptics can help users perform complex tasks requiring precise hand–eye coordination, such as surgery, physical rehabilitation, flying a drone, or manipulating very large or very small objects, ranging from space vehicles to microscopic organisms. Haptics technology is also finding a home in consumer applications, such as computer games, enabling users to feel as well as see and hear play interactions.

"Haptics technology can provide the user with a representation of the sense of touch and corresponding force feedback in interacting with a virtual or a remote environment," says Robert Diraddo, medical devices' group leader for the National Research Council of Canada. "These (sensations) are most important in virtual reality and remote manipulation applications, where the interactive sense of touch is important."

Gabriel Robles-De-La-Torre, founder of the International Society for Haptics, notes that signal processing is crucial in haptic system design and operation. "[Applications] range from filtering signals coming from position and torque sensors in haptic interfaces to processing the signals coming from touch sensors intended for robot control," he says. "One can also use fast Fourier transform (FFT) and other techniques to analyze the results of measurements in human performance experiments."

WHEN USED IN CONJUNCTION WITH A VIDEO DISPLAY, HAPTICS CAN HELP USERS PERFORM COMPLEX TASKS REQUIRING PRECISE HAND–EYE COORDINATION.

The most difficult task facing haptic system developers is creating accurate representations of real-world tactile sensations without consuming vast amounts of processor power or driving up costs excessively. "Ensuring suspension of multisensorial user disbelief while maintaining hardware affordability is the biggest challenge for haptics-driven virtual reality," Diraddo says. "If not addressed, user-adoption of the technology is limited."

VIRTUAL HANDS-ON TREATMENT

University of Texas (UT) at Dallas researchers are developing a haptics-based technology that promises to make it possible for physical therapists and other health-care professionals to work with patients located almost anywhere. The system utilizes multiple three-dimensional (3-D) cameras to create avatars of both the care provider and patient and then places them inside a virtual space where they can interact with each other in a realistic fashion.

"Our technology allows real-time capturing, transmission, and rendering of 3-D avatars of multiple human actors as well as force feedback," says Balakrishnan "Prabha" Prabhakaran, a professor of computer science at UT Dallas and the project's principal investigator (Figure 1).

With conventional telemedicine, a caregiver and patient can both appear on the same display, talk to each other, and even exchange medical data, yet still have no practical way of physically interacting. "Our project's goal is to facilitate high-quality multimodal 3-D teleimmersion incorporating a force-feedback experience in a 3-D virtual space using haptic devices," Prabhakaran says.

The researchers' current platform allows applied force to be sent from one user to the other. This way, a care provider could, for example, feel an accurate representation of the strength of a patient's muscle as it is being flexed. The platform uses force-feedback haptic devices manufactured by Geomagic of Morrisville, North Carolina. The units rely on built-in motors to create forces that push back on the user's hand to simulate touch and interaction with virtual objects. "We are focusing on upper-arm movement, because it's easier in terms of the off-the-shelf devices that already exist," Prabhakaran says.

Haptic devices come in different forms, and the one the researchers are currently using is not wearable. The device, which sits on a table, features a small, grabbable handle. "You can feel the force feedback when you are moving the handle," Prabhakaran says. The device can generate, or receive, force-feedback in three translational degrees of freedom as

Digital Object Identifier 10.1109/MSP.2013.2292433
Date of publication: 12 February 2014

[FIG1] Balakrishnan "Prabha" Prabhakaran. (Photo courtesy of the University of Texas at Dallas.)

well as torque feedback in three rotational degrees of freedom in yaw, pitch, and roll directions.

The current device isn't suitable for rehabilitation use, yet it is helping the researchers to coordinate haptic movements with the 3-D virtual space. More sophisticated haptic technologies are currently under development, Prabhakaran says. "What we are trying to do is to build our own haptic device using resistance motors and microcontrollers," he notes. Prabhakaran says the researchers are striving to achieve maximum realism and accuracy via force-feedback technology. "For instance, if I am moving my arm up, and the doctor is moving his or her arm down, then it would be like offering resistance in the opposite direction. That gives you a feel for how healthy the patient's muscles are."

In the years ahead, Prabhakaran sees care providers and patients using several different types of haptics devices, with each designed to mimic and communicate data about a specific body part. "We use our arms and legs in different ways," Prabhakaran says. "It's like when you go to the gym; you have different equipment for different muscle groups."

Transmitting large amounts of audio, video, and haptics data increases the likelihood of significant lag time and other types of transmission delays. The researchers are addressing this challenge by developing sophisticated algorithms and various types of software designed to speed data through the Internet in real time. "All the data that we are dealing with in realizing our goal of multimodal 3-D teleimmersion are digitized signals, representing time-varying and spatially varying physical quantities," Prabhakaran says. "These data need to be synchronously captured in real time, analyzed, compressed if needed, transmitted, and processed on the receiver's side before rendering them to the user. Hence, signal processing is an inherent component of our research."

[FIG2] The Bristol Robotics Laboratory haptic robot palpitates artificial tissue samples. (Photo courtesy of Bristol Robotics Laboratory.)

Prabhakaran notes that signal processing also plays key roles in other areas, including filtering, image processing, and pattern recognition, as well as in the development of different haptics devices. "Multidimensional and biomedical signal processing is also critical to the operation of the inertial body sensors," he adds.

Innovative algorithms will also help minimize the amount of data that needs to be exchanged. Xiaohu Guo, an associate professor of computer science at UT Dallas and a project coprincipal investigator, is an expert in computer graphics, animation, and modeling. Guo is refining techniques to allow data moving between haptic devices to be transmitted over a network more efficiently, as well as creating 3-D visual images of user movements in real time. Gao has also reported success in downsizing massive data sets with spectral transformation techniques. The technique relies on manifold harmonics to transform 3-D images into points that represent the surface of an object. The data is then compressed into a more compact form that can be transmitted more rapidly over networks.

Prabhakaran believes that the system will eventually provide benefits to users of both telerehabilitation and in-home rehabilitation services. "Current in-home rehabilitation relies on patients' word of mouth—or diary—for having done the prescribed exercises with the prescribed number of repetitions," he says. Limited oversight capabilities requires caregivers to trust that the patient has actually performed the exercises, and in the correct manner. "With 3-D teleimmersion incorporating haptics, the experts can have an exercise record ... that contains the trajectories of the body joint movements as the exercises were done, the number of repetitions and so on," Prabhakaran says.

Prabhakaran predicts a bright future for haptic medical and rehabilitation technologies. "Sense of touch is really important in areas such as physical medicine and rehabilitation, where the experts would like to have a 'feel' for the strength

of muscles or the trajectory of body joint movement," he says. "It will be appreciated by both caregivers and patients."

ROBOTIC SURGERY WITH FEELING

Developing a haptics technology that can help a surgeon operate with great precision, even when using a robot as the primary surgical tool, is the goal of researchers at the Bristol Robotics Laboratory (BRL), a joint venture of the University of the West of England and the University of Bristol.

"At the moment, the da Vinci robot, which is the main robot used in surgery, doesn't actually provide any haptics feedback," explains Adam Spiers, a BRL research associate. "So surgeons need to make their diagnosis of tissue inside the body based purely on vision, which works but is also quite unnatural." Even with the assistance of the best video equipment, accidents still happen from time to time. "There have been examples of mistakes in operations because surgeons couldn't distinguish one type of tissue from another, or detect what is underlying a structure," Spiers notes.

The BRL researchers aim to give surgeons who operate with robots the ability to feel as well as see inside patients. The extra sensing ability, delivered from the robot to a surgeon's finger, promises to make surgical procedures more accurate and reduce the likelihood of potentially devastating errors.

"The area of surgery we've been looking at is keyhole—laparoscopic—surgery," Spiers says. "This type of surgery is very difficult to perform and has a lot of restrictions in terms of vision and touch feedback, which makes it quite easy to misidentify tissue structures."

Spiers notes that different types of organ tissues present various levels of stiffness, a characteristic that can be used to help robot-equipped surgeons operate more accurately and efficiently. "We work primarily with urologists, and the main thing they use their robot for is the removal of cancerous prostates," he says.

To find and remove a diseased prostate, a surgeon must slice through a great deal of tissue. "There's all these different types of tissue structures—fat, nerve tissue, and the prostate itself," Spiers says. "If you could find and detect those things by touch, you would potentially save errors where others have misidentified tissue and dissected it by accident."

Spirers says his team recently completed a study based on a prototype surgical system that demonstrated haptics ability in action (Figure 2). Project participants were asked to find hard tumors placed within opaque synthetic tissue samples located in a distant part of the lab (Figure 3). The system relayed the participants' finger movements to a robot that then precisely duplicated the motions and transmitted the corresponding sense information as it touched the tissue samples back to each participant. Using an Internet connection, there was a performance loss of under 1% compared to using one's own finger to detect the lumps, Spiers says. "If we can move this prototype system into a surgical scenario, the technology will have significant implications," he notes. "The main thing we're hoping for with the inclusion of haptics in surgical systems is a reduction of error—less error in surgery."

DEVELOPING A HAPTICS TECHNOLOGY THAT CAN HELP A SURGEON OPERATE WITH GREAT PRECISION, EVEN WHEN USING A ROBOT AS THE PRIMARY SURGICAL TOOL, IS THE GOAL OF RESEARCHERS AT THE BRISTOL ROBOTICS LABORATORY.

From a signal processing perspective, one of the biggest challenges facing the researchers is ensuring that the data moves from surgeon to robot and back again with as little delay as possible, Spiers says. "The whole idea is that we're trying to create the illusion to people that their finger is somewhere else, that they are feeling things on the other side of the room." Maintaining almost instantaneous feedback is essential to the technology's success. "As soon as things start slowing down, or getting jittery, or you feel something after you've finished touching it, then the illusion falls apart," Spiers explains.

[FIG3] A user operates the Bristol Robotics Laboratory haptic robot. (Photo courtesy of Bristol Robotics Laboratory.)

System integration work consumes a great deal of the researchers' time. "We have quite a few different elements in our system—the part on their user's fingertip, the robot, and the robot's fingertip—and we have a lot of signals going between all of those components," Spiers says. "We spend quite a bit of time trying to make sure that all of our signal processing and networking algorithms work as fast as possible."

HAPTIC TWEEZERS

While scientists can turn to a microscope to view and study microscopic structures, touching and feeling extremely minute objects has always been a far more challenging task. But perhaps not for much longer. Researchers at France's Université Pierre et Marie Curie say that the new "haptic optical tweezers" they have developed (Figure 4) will enable microscope users to manipulate samples within worlds up to a million times smaller than humans.

The tool promises to help scientists explore microscopic worlds by sensing and exerting piconewton-scale forces, allowing improved micromanipulation and microassembly dexterity. "The initial results obtained are promising and demonstrate that optical tweezers have a significant potential for haptic exploration of the microworld," says Cecile Pacoret,

[FIG4] **The haptic optical tweezers. (Photo courtesy of Université Pierre et Marie Curie.)**

a postdoctoral research fellow at Université Pierre et Marie Curie, who also works in collaboration with the Nanorobotics Laboratory at Carnegie Mellon University. "Haptic optical tweezers will become an invaluable tool for force-feedback micromanipulation of biological samples and nano- and microassembly parts," she predicts.

Conventional optical tweezers, which manipulate tiny objects with highly focused laser light that generates an attractive or repulsive force, are already widely used for the study of single molecules and cells. Yet biological researchers exploring 3-D cells or molecular phenomena like DNA stretching would benefit greatly from having access to touch feedback as they go about their work. "Our system already provides great improvement in maneuverability and training requirements," Pacoret states.

The technology is also relatively easy to learn and use. Pacoret notes that guests visiting the lab are generally able to perform their first micromanipulations within just a few minutes. "Our system allows an operator to touch and explore microscaled shapes like microelectromechanical systems or single cells," she says. "Moreover, it is able to interact, apply stress, and give feedback."

One of the biggest challenges facing the researchers as they developed the technology was magnifying piconewton-scale forces high enough to allow human operators to perceive interactions never before physically sensed, such as adhesion phenomena, extremely low inertia, and the high-frequency dynamics of extremely small objects. Design requirements included very high sensitivity and dynamic stability. "Our system merges different techniques and is highly multidisciplinary, combining vision, signal processing, real-time programming, optics, robotics, automatics, [and] haptics," Pacoret says. "The classic optical tweezers have been modified to meet the human perception requirement: safety, stability, reliability, and responsiveness."

Signal processing is used to ensure the performance and reliability of the tool's haptic coupling and teleoperation functions. "Information is exchanged bilaterally between two robots: the slave robot—the micromanipulator in our case—and the master robot, the haptic interface," Pacoret says. The arrangement creates an automatic closed loop that must remain stable under all circumstances for the safety of both the samples and the robots. "The originality of our system is to have designed all the components to have the smallest need of filtering, so our coupling loop can be reduced to the most transparent type of coupling," Pacret says. "Our thesis is that a well-designed haptic system allows high performance."

> **MOST HAPTICS RESEARCHERS FEEL THAT THE TECHNOLOGY'S POTENTIAL IS ONLY BEGINNING TO BE FULLY UNDERSTOOD AND APPRECIATED.**

Pacoret predicts that the haptic optical tweezers will one day become an important tool for the exploration, diagnosis, and assembly of sensors, microsystems, and biomedical elements, including cells, bacteria, viruses, and proteins. Such objects tend to be extremely fragile and difficult to manipulate. "This tool will provide a new degree of freedom and accessibility to researchers, including, for example, new versatility in the study and micromanipulation of cells," she says.

THE FUTURE OF HAPTICS

Most haptics researchers feel that the technology's potential is only beginning to be fully understood and appreciated. "Applications in medicine already include the interactive design of surgical devices, surgical robotics, patient-specific (surgical) rehearsal, and rehabilitation," Diraddo says. "Outside of medicine, there are many potential applications, such as the handling of dangerous materials, maintenance training, interactive automotive design, aerospace telemanipulation, and serious gaming."

Robles-De-La-Torre believes that technology developers in a variety of fields are now acquiring a growing respect for haptics, coming to the realization that the sense of touch is as essential in many situations as vision and hearing. If this trend continues to build, as now appears likely, it should lead to haptics finding a place in a wider number of systems and applications.

"The importance of the scientific part of haptics is fundamental, but largely overlooked in general," Robles-De-La-Torre says. He points out that removing someone's haptic abilities would leave the individual unable to walk, chew food, or articulate speech normally. "Skilled tool use would be out of the question," he says. (Try, for example, using a pen or a pair of scissors after your arm has fallen asleep.) "Basically, our haptic capabilities enable us to interact with and change our environment," Robles-De-La-Torre says.

AUTHOR

John Edwards (jedwards@johnedwardsmedia.com) is a technology writer based in the Phoenix area.

[SP]

from the GUEST EDITORS

Ivan B. Djordjevic, William Shieh,
Xiang Liu, and Moshe Nazarathy

Advanced Digital Signal Processing and Coding for Multi-Tb/s Optical Transport

The exponential Internet traffic growth places an enormous transmission capacity demand on the underlying information infrastructure at all levels, ranging from core to access networks. In response to the capacity demand, the IEEE ratified a 40/100 Gb/s Ethernet (GbE) standard (IEEE 802.3ba) in June 2010. The deployment of 100GbE has been underway at an accelerated pace. To meet the ever-increasing capacity demands, 1Tb/s Ethernet (TbE) rates and beyond (e.g., 4 TbE and 10 TbE) are expected to be standardized in the near future. There are several emerging technologies, such as Nyquist wavelength-division multiplexing (WDM), orthogonal frequency-division multiplexing (OFDM), polarization-division multiplexing (PDM), and high-level modulation formats, potentially usable in concert to deliver multi-TbE services. However, many challenges still exist with respect to practical and cost-effective implementations of these new technologies. The reality is that TbE or multi-TbE connectivity requires massive parallel processing in both the optical and electronic domains. The envisioned capacity growth in Internet traffic is about to place enormous demand not only on transmission speed at every level but also on the energy consumption required for information creation, distribution, and reception. Recent studies indicate that the power consumed by the information and communication technology (ICT), currently about 2–4% of total carbon emissions, is going to be doubled by the end of this decade, provided the current trend continues. Therefore, the Internet is becoming constrained not only by achievable speed and capacity but also by tolerable energy consumption. It is a well-known fact that large data centers are built closer to power plants to provide ample and cost-effective energy supply. It can be concluded that it is imperative for the research and development community to address, sooner rather than later, both bandwidth and energy constraints; any technological advances on this subject, even seemingly incremental, would be meaningful and beneficial.

This issue of *IEEE Signal Processing Magazine* aims to highlight diverse recent advances in digital signal processing (DSP) and coding, enabling TbE and multi-TbE optical transport while addressing bandwidth and energy constraints. The addressed topics range from sophisticated modulation and coding schemes to advanced detection schemes. The multi-Tb/s optical transports over either single-mode fibers or few-mode/few-core fibers are also topics covered in this special issue.

The main technologies covered by this special issue comprise multiple-input, multiple-output (MIMO) signal processing, advanced multilevel and multidimensional modulation schemes, advanced multiplexing schemes, signal processing for superchannel transmission, advanced DSP for signal detection and equalization, and advanced coding, all aiming at achieving multi-Tb/s optical transports.

Liu et al. introduce the concept of superchannel and review recent advances in the generation, transmission, and detection of optical superchannels at channel data rates of the order of Tb/s. The authors describe the multi-Tb/s enabling DSP techniques including Nyquist-WDM, multiband OFDM, software-defined modulation and detection, advanced coding, and joint DSP among the superchannel constituents.

Arık et al. present the fundamentals of MIMO signal processing for mode-division multiplexing (MDM) in multimode fiber (MMF). They further compare the performance and complexity of MIMO signal processing architectural candidates, establishing that programmable frequency-domain equalization (FDE) of chromatic dispersion and adaptive FDE of modal dispersion (MD) provide an attractive combination. The authors also review two major algorithms for adaptive FDE of MD, mainly least mean squares and recursive least squares, analyzing their complexity, throughput and adaptation speed.

Zhou provides a systematic review of the challenges and recent progress in timing and carrier synchronization techniques for high-speed optical transmission systems using single-carrier-based coherent optical modulation formats.

Du et al. review different approaches to the mitigation of fiber nonlinearity impairments, including methods directly compensating fiber nonlinearity, such as digital back-propagation, as well as the technique of optimizing the modulation format to reduce the transmission impairments resulting from fiber nonlinearities.

Yoshida et al. describe a critical technique for cycle slip (CS) compensation in nondifferential coded coherent optical transmission systems. They describe how CS can be estimated from a shorter block of symbols by monitoring sparse and asymmetric polarization block coded symbols.

Nazarathy and Tolmachev describe the use of underdecimated filter banks to digitally slice the optical channel bandwidth into multiple spectrally disjoint subbands to be processed in parallel. This is an alternative approach to the parallelization of

(continued on page 142)

Digital Object Identifier 10.1109/MSP.2013.2288652
Date of publication: 12 February 2014

Xiang Liu, S. Chandrasekhar, and Peter J. Winzer

Digital Signal Processing Techniques Enabling Multi-Tb/s Superchannel Transmission

An overview of recent advances in DSP-enabled superchannels

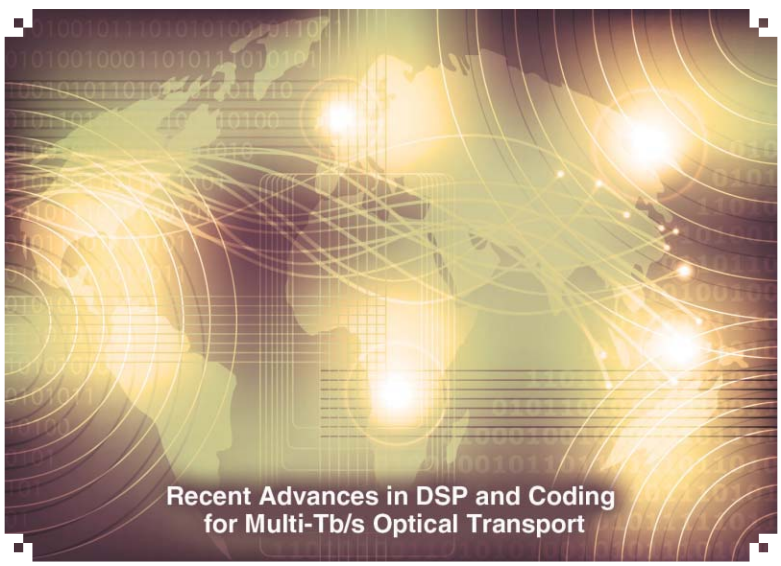


IMAGE LICENSED BY INGRAM PUBLISHING /JANMIKS

Digital signal processing (DSP) combined with coherent detection has played a central role in the recent capacity expansion of optical networks. Optical superchannels aim to increase per-channel interface rates as well as per-fiber capacities of wavelength-division multiplexed (WDM) systems in a cost-effective manner. Superchannels circumvent the electronic bottleneck via optical parallelism and provide high-per-channel data rates and better spectral utilization, especially in transparent optical mesh networks. This article reviews recent advances in the generation, detection, and transmission of optical superchannels with channel data rates on the order of terabits per second (Tb/s). Enabling DSP techniques such as Nyquist–WDM, orthogonal frequency-division multiplexing (OFDM), software-defined modulation and detection, advanced coding, and joint DSP among the superchannel constituents are presented. Future prospects of DSP techniques for high-capacity superchannel transmission are also discussed.

Digital Object Identifier 10.1109/MSP.2013.2285934
Date of publication: 12 February 2014

SUPERCHANNEL TRANSMISSION

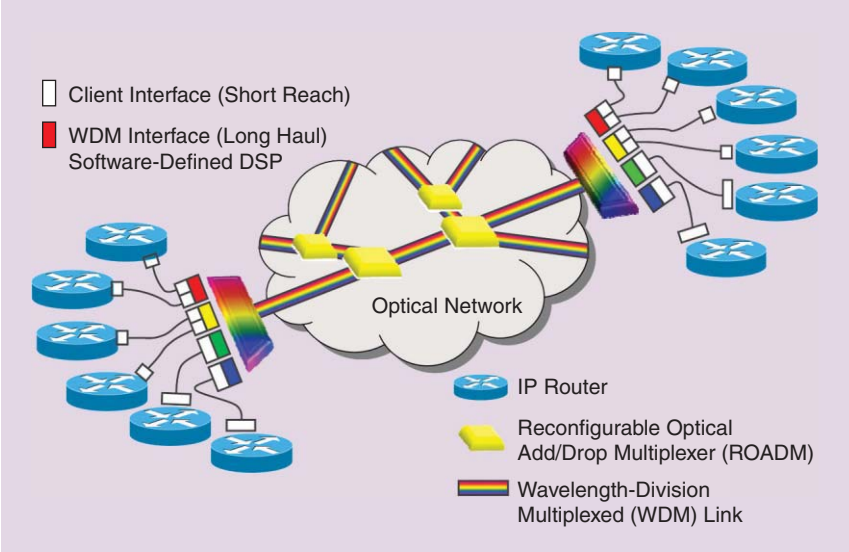
To satisfy the ever-increasing capacity demands of fiber-optic networks, the data rate carried by each wavelength channel in WDM

1053-5888/14/$31.00©2014IEEE

systems has been increasing exponentially for over two decades [1]. Figure 1 shows the schematic of a typical optical network, connecting Internet protocol (IP) routers via short-reach client interfaces to long-haul WDM transponders, which are wavelength multiplexed onto fibers making up an optically routed mesh network. The key enabling elements for the recent interface speed and fiber capacity expansion in optical networks are high-speed optical WDM transceivers based on coherent detection and advanced DSP [1]. High-speed digital-to-analog converters (DACs) and analog-to-digital converters (ADCs) built in 40-nm complementary metal–oxide–semiconductor (CMOS) technology are currently available up to 65 Gsamples/s with an effective number of bits (ENOB) of about six (e.g., see http://www.fujitsu.com/downloads/MICRO/fme/documentation/c63.pdf). Powerful digital signal processors, containing on the order of 100 million gates [2] and capable of operating at close to 100 tops/s, and form the basic building blocks of modern optical transceivers. Figure 2 shows the evolution of such application-specific integrated circuits (ASICs) for commercial optical communications applications, showing a gate count increase of ~70% per year. Single-carrier 100-Gb/s transponders have been commercially available since mid-2010 (e.g., see http://www3.alcatel-lucent.com/100g-coherent/). To scale to Tb/s interfaces and beyond (expected to be needed by 2015 [1]) in an economically attractive manner, however, optical superchannels that circumvent the electronic bottleneck via optical parallelism are the method of choice. Consequently, dual-carrier 400-Gb/s digital coherent transponders have been commercially available since early 2013 (e.g., see http://www3.alcatel-lucent.com/400g-pse/), and long-haul transmission with Tb/s superchannel data rates has been experimentally demonstrated [3]–[7]. In the context of optical transport networks, the term *superchannel* was first used in [3] to refer to multiple single-carrier-modulated signals that are seamlessly multiplexed under the coherent optical OFDM (CO-OFDM) conditions [3]–[5]. The superchannel concept was later generalized to any collection of optical signals that are

1) modulated and multiplexed together with high spectral efficiency (SE) at a common originating site
2) transmitted and routed together over a common optical link
3) received at a common destination site.

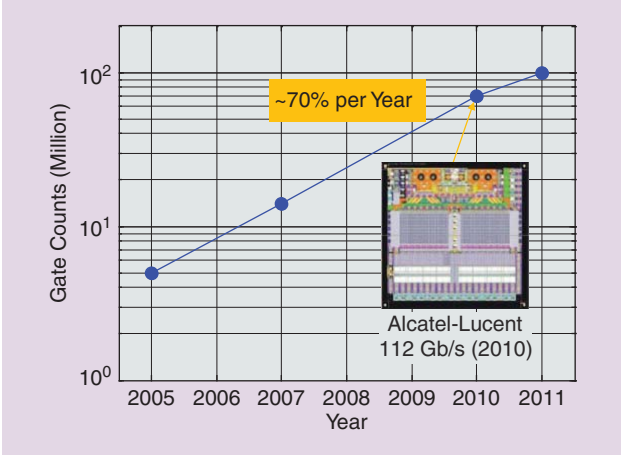
To achieve high-SE multiplexing, "Nyquist–WDM" and "quasi-Nyquist–WDM" with spectrally shaped single-carrier modulated signals have also been introduced [8], [9], offering an alternative to OFDM with tradeoffs in SE, DSP complexity, optoelectronic (O/E) hardware complexity, and subcarrier access possibilities [1]. From a networking point of view, the introduction of Tb/s-class superchannels has led to a rethinking of the spectral bandwidth allocation in optical fibers. While most current systems operate on a rigid 50-GHz WDM channel grid, Tb superchannels benefit from so-called "flexible grid" WDM systems that allow for more efficient utilization of the optical spectrum (see, e.g., ITU-T G.694.1, "Spectral Grids for WDM Applications: DWDM Frequency Grid").



[FIG1] **A schematic of a typical optical network.**

The key benefits of superchannels in WDM systems are as follows:

1) the ability to meet the demand of high-speed serial interface rates, which increases faster than the speed provided by O/E converters, electro-optic (E/O) converters, DACs, and ADCs
2) higher SE in WDM transmission, by reducing the percentage of wasted optical spectrum between individual channels
3) increased efficiency in DSP
4) better leverage of photonic integrated circuits and ASICs



[FIG2] **The evolution of commercial optical communications ASICs.**

5) native support of software-defined optical transmission, enabled by DSP at both the transmitter and receiver, to improve system throughput and flexibility.

OVERVIEW OF SUPERCHANNEL ARCHITECTURES

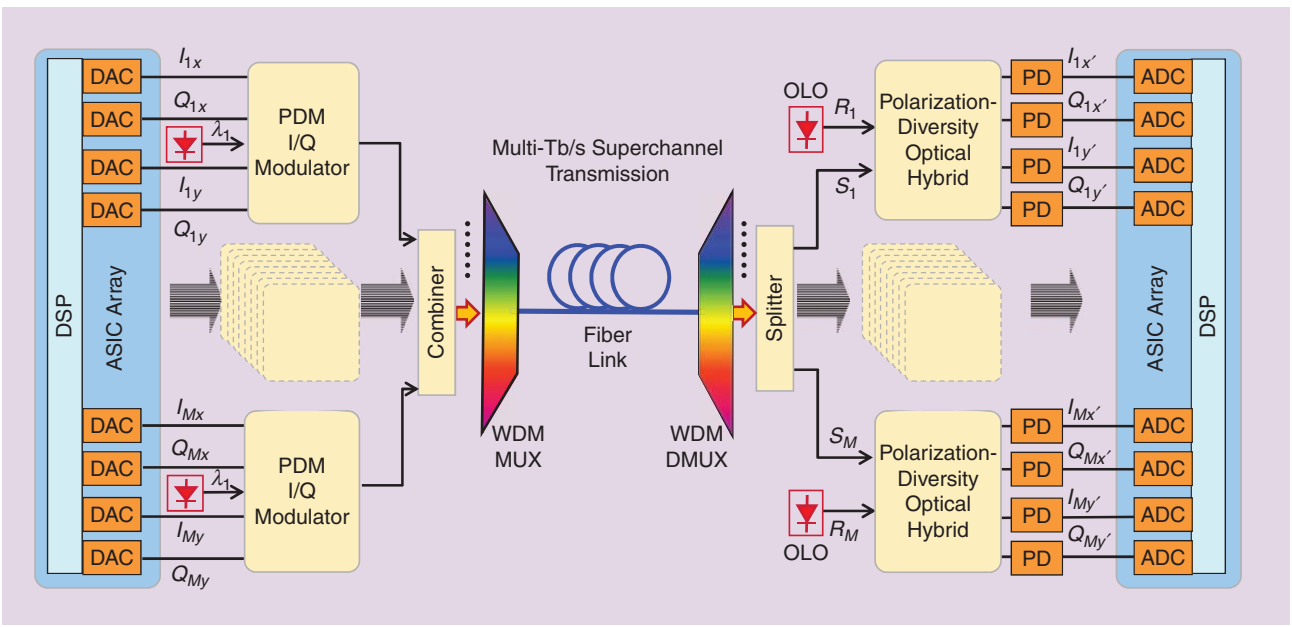
Figure 3 shows the schematic of a superchannel transponder embedded in a WDM optical network. The full optical field information of an optical signal can be decomposed into four orthogonal real-valued components, particularly the inphase (I) and quadrature (Q) components of the two orthogonal polarization states (x and y) supported by a single-mode optical fiber. At the transmitter, a polarization-division multiplexed (PDM) I/Q modulator is commonly used to imprint these four high-speed electrical waveforms onto an optical carrier, emerging from a semiconductor laser oscillating at 193 THz and stabilized to within about ± 1 GHz. At the receiver, a polarization-diversity 90° hybrid and an optical local oscillator laser (OLO) are used to decompose the received signal into four orthogonal components, which can then be detected by four photo-detectors (PDs) and processed in a digital signal processor to recover the original signal. To avoid the bandwidth bottleneck of electronic and O/E components in the transmitter and receiver, optical parallelism is utilized to generate and detect a superchannel, whose aggregate bandwidth by definition exceeds that of the individual transponder components.

THE KEY ENABLING ELEMENTS FOR THE RECENT INTERFACE SPEED AND FIBER CAPACITY EXPANSION IN OPTICAL NETWORKS ARE HIGH-SPEED OPTICAL WDM TRANSCEIVERS BASED ON COHERENT DETECTION AND ADVANCED DSP.

It is evident from Figure 3 that the generation and detection of superchannels can greatly benefit from large-scale integration of both optical and electronic components, leading to potential savings in cost, size, and power. With all the constituents of a superchannel being available at the transmitter and at the receiver, joint DSP may be leveraged to improve the transmission performance and/or to reduce DSP complexity, as will be discussed in more depth later.

SUPERCHANNEL CLASSIFICATION

Table 1 summarizes various superchannel demonstrations reported recently. Key performance indicators include intrachannel SE (ISE) and the ISE-distance product (ISEDP). Using OFDM-based seamless multiplexing, Chandrasekhar et al. demonstrated the transmission of a 1.2-Tb/s superchannel, consisting of 24 polarization division-multiplexed (PDM) quadrature phase-shift keying (QPSK) signals, over 7,200 km of ultra-large area fiber (ULAF) [3]. With the use of 16-QAM, Huang et al. demonstrated the transmission of a 1.5-Tb/s superchannel over 1,200 km of standard single-mode fiber (SSMF) [5]. With the use of reduced-guard-interval (RGI) OFDM, which essentially performs full chromatic dispersion compensation prior to OFDM processing and hence allows reduction of the guard-interval to accommodate mostly transmission effects with much shorter channel memory, Liu et al. demonstrated the transmission of a 485-Gb/s superchannel over an ISEDP as high as 30,000 km·b/s/Hz [10]. To avoid the use of accurately frequency-locked optical carriers, a small frequency guard-band between adjacent carriers within a superchannel can be used to allow one to trade a small fraction (e.g., <10%) of the link capacity for



[FIG3] An illustration of a superchannel transponder embedded in a WDM system.

[TABLE 1] RECENT Tb/s SUPERCHANNEL DEMONSTRATIONS.

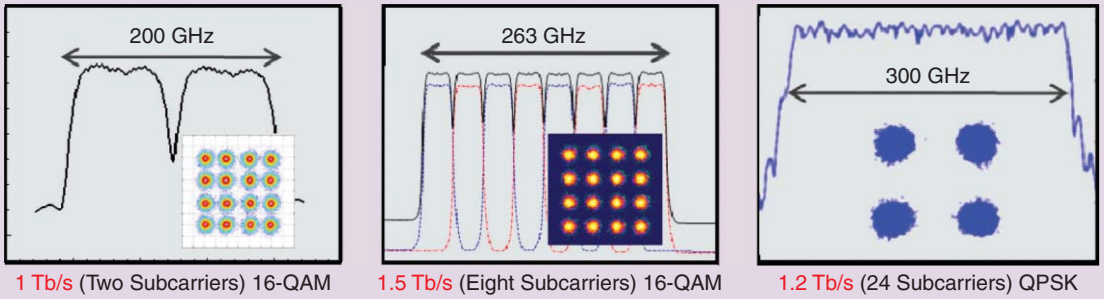
MODULATION FORMAT	SUPERCHANNEL DATA RATE	COMPOSITION	ISE (b/s/Hz)	REACH (km)	ISEDP (km × b/s/Hz)
SEAMLESS CO-OFDM (WITH FREQUENCY-LOCKED OPTICAL CARRIERS)					
PDM-QPSK [3]	1,200 Gb/s	24 × 50 Gb/s	3.74	7,200	26,928
PDM-16 QAM [5]	1,500 Gb/s	15 × 100 Gb/s	7.00	1,200	8,400
RGI-OFDM 16 QAM [10]	485 Gb/s	10 × 48.5 Gb/s	6.20	4,800	29,760
GUARD-BANDED CO-OFDM (WITH FREQUENCY-UNLOCKED CARRIERS)					
OFDM 16 QAM [6]	1,864 Gb/s	8 × 233 Gb/s	5.75	5,600	32,000
QUASI-NYQUIST–WDM (WITH FREQUENCY-UNLOCKED CARRIERS AND DIGITAL SPECTRAL SHAPING)					
PDM-32-64 QAM [11]	504 Gb/s	5 × 100.8 Gb/s	8.00	1,200	9,600
QUASI-NYQUIST–WDM (WITH FREQUENCY-UNLOCKED CARRIERS AND OPTICAL SPECTRAL SHAPING)					
PDM-16 QAM [7]	1,280 Gb/s	2× 640 Gb/s	5.00	3,200	16,000

simplicity, scalability, and performance. With the use of this "guard-banded OFDM," Liu et al. showed the transmission of a 1.5-Tb/s superchannel, consisting of eight PDM-OFDM 16-QAM signals, over 5,600 km of ULAF, achieving a record ISEDP of over 32,000 km·b/s/Hz for Tb/s-class superchannel transmission with more than 5 b/s/Hz net SE [6]. Using Nyquist–WDM and hybrid 32-QAM and 64-QAM, Zhou et al. demonstrated the transmission of five 400-Gb/s superchannels on a 50-GHz grid over 1,200-km ULAF with a WDM SE as high as 8 b/s/Hz [11]. With the use of 80-Gbaud modulation and detection, Raybon et al. demonstrated the transmission of a 1-Tb/s superchannel using only two optical carriers and two pairs of transmitter/receiver front ends [7]. There, close-to-Nyquist spectral shaping was achieved by optical filters instead of digital pulse shaping. Figure 4 shows spectra of representative Tb-class superchannels, where the tradeoff between the number and modulation rate of the subcarriers becomes evident.

OVERVIEW OF DSP BUILDING BLOCKS

With access to the full optical field information at both the transmitter and receiver, modern coherent optical communications has benefited from many powerful DSP techniques originally developed for wireless communications. However, data rates in optical communications are usually several orders of magnitude higher than in wireless communications, necessitating efficient DSP techniques tailored specifically to optical communications, indeed reflecting a statement by Haykin [12], "Signal processing is at its best when it successfully combines the unique ability of mathematics to generalize with both the insight and prior information gained from the underlying physics of the problem at hand." For optical communications in single-mode fiber, channel equalization can be efficiently realized by separating slowly varying transmission effects such as chromatic dispersion (CD) from rapidly varying effects such as polarization rotations and polarization-mode dispersion (PMD).

Figure 5(a) schematically shows a superchannel transmitter DSP architecture. Key signal processing steps include forward error correction (FEC) encoding, constellation mapping (MAP), fiber nonlinearity compensation (NLC), up-sampling (UpS), electronic dispersion precompensation (EDC), Nyquist prefiltering, and equalization of a nonideal transmitter frequency response through a static transmitter equalizer (SEQ). The four resulting electrical signals representing I and Q components of the baseband signal in both x and y polarization are then modulated onto a high-frequency carrier, generated by a reasonably



1 Tb/s (Two Subcarriers) 16-QAM
5.2 b/s/Hz
3,200-km Transmission [7]
(a)

1.5 Tb/s (Eight Subcarriers) 16-QAM
5.7 b/s/Hz
5,600-km Transmission [6]
(b)

1.2 Tb/s (24 Subcarriers) QPSK
3 b/s/Hz
7,200-km Transmission [3]
(c)

[FIG4] (a)–(c) Representative Tb-class superchannel spectra. Insets: representative recovered signal constellations.

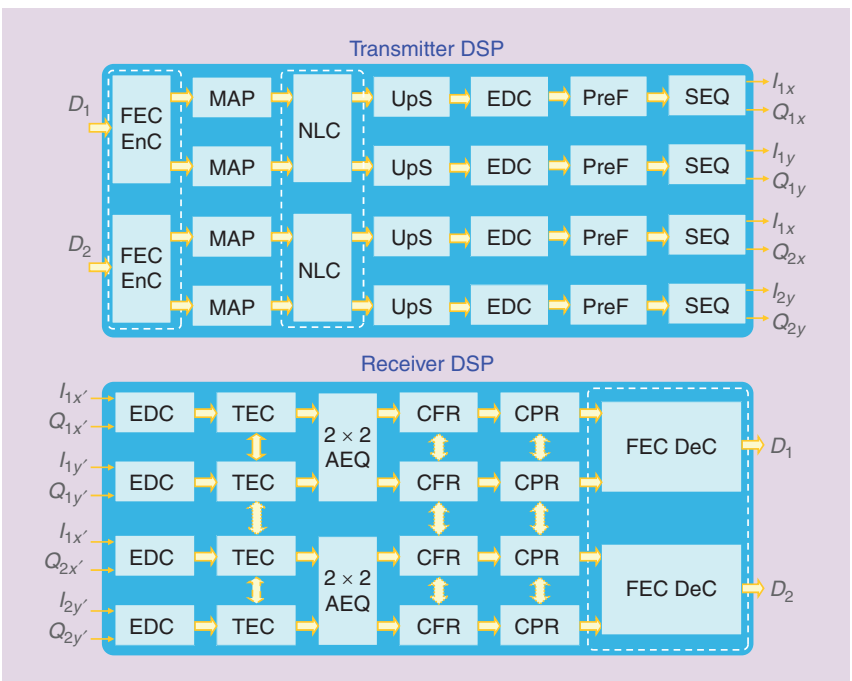
narrow-linewidth (~100-kHz) laser oscillating at around 193 THz. Note that some of the transmitter-side DSP modules shown in Figure 5, e.g., NLC and EDC, may bring valuable performance gains in long-haul optical transmission, but their DSP complexities need to be taken into consideration. In some cases, these modules can be made optionally available depending on the requirements on performance and power consumption.

Figure 5(b) shows the schematic architecture of a superchannel receiver DSP. After coherent O/E conversion, the DSP is presented with four electrical baseband signal representing I and Q components in x and y polarization. Key processing steps include EDC, timing error correction (TEC), adaptive 2×2 multiple-input-multiple-output (MIMO) equalization and polarization source separation compensating for polarization rotation and PMD, carrier frequency recovery (CFR), carrier phase recovery (CPR), and FEC decoding based on both soft decision (SD) and/or hard decision (HD). (Reviews of these techniques are given, e.g., in [1] and [13].) Typically, transmitter-side DSP has the potential advantages of 1) absence of optical noise and 2) efficient implementation operating on one sample per symbol. On the other hand, receiver-side DSP has the potential advantage of being able to adaptively and quickly react to dynamic channel changes. Note in this context that receiver-to-transmitter feedback is often problematic in optical systems: For a 1,000-km link, signal round-trip times are on the order of 10 ms, whereas channel dynamics can be in the 10-kHz range or above, owing to acoustically and mechanically induced vibrations of transmission fiber segments. As optical channels are usually well defined (in terms their frequency response), blind channel equalization and phase recovery has been widely used in optical coherent detection [13]. For fast channel equalization and/or sophisticated modulation formats, training symbols (TSs) can be inserted in the data symbol stream to aid channel equalization [14]. To enable the efficient implementation of high-performance FEC codes, it is desirable to recover the phase of the modulated symbols (as opposed to just the differential phase between symbols), which can be readily realized by pilot-assisted phase estimation (PA-PE) [15]. For superchannel transmission, some of the processes such as FEC and NLC may be performed jointly over multiple signals within the superchannel to improve the overall superchannel performance. We will discuss some key signal processing modules essential to superchannel transmission in the following sections.

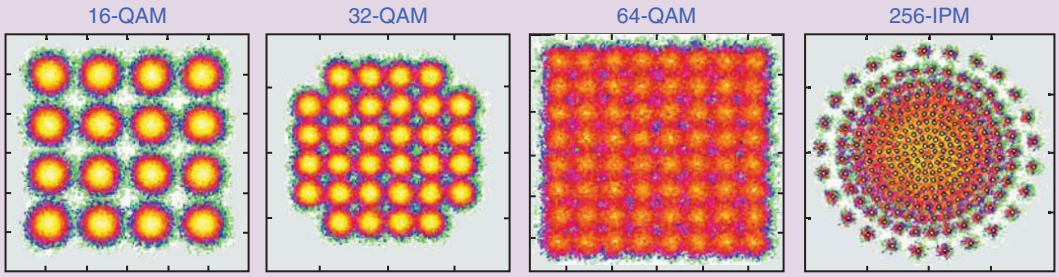
FOR SUPERCHANNEL TRANSMISSION, SOME OF THE PROCESSES SUCH AS FEC AND NLC MAY BE PERFORMED JOINTLY OVER MULTIPLE SIGNALS WITHIN THE SUPERCHANNEL TO IMPROVE THE OVERALL SUPERCHANNEL PERFORMANCE.

SPECTRAL SHAPING FOR NYQUIST-WDM SUPERCHANNELS

To enable the spectrally efficient construction of a superchannel, spectral shaping is needed at the transmitter to band-limit the optical spectrum of each superchannel subcarrier to be equal to or slightly higher than the modulation symbol rate. This is done by prefiltering the spectrum of a single-carrier signal by using a root raised cosine filter (RRC) with a given roll-off factor. Usually, a roll-off factor of about 0.1 is implementable at reasonable complexity and with negligible performance loss. Figure 4(b) shows the measured optical spectrum of a 1.5-Tb/s guard-banded OFDM superchannel [6]. It consists of eight spectrally shaped signals, closely packed at a net SE of 5.75 b/s/Hz using individual lasers without mutual frequency locking. The transmitter-side signal processing can also be used to mitigate the bandwidth limitation from DAC, modulator, modulator driver, and optical add/drop multiplexers (ROADMs), optical filtering elements used to optically route wavelength channels across a mesh network; c.f. Figure 1. The transmitter-side equalizer is usually implemented as a relatively static equalizer.



[FIG5] **A schematic architecture of a transceiver DSP for a superchannel having two optical carriers.**



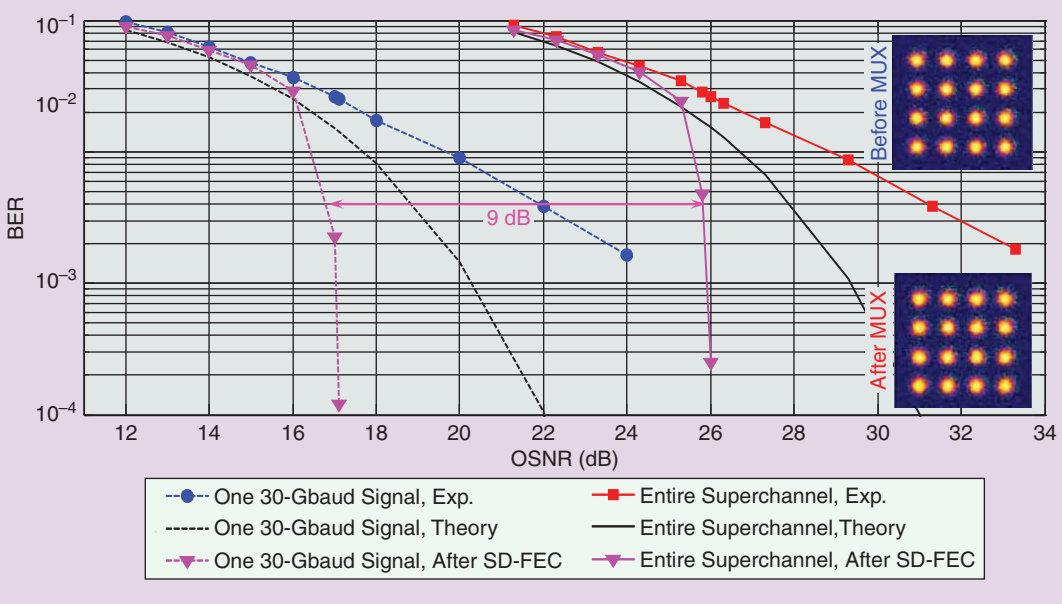
[FIG6] **Measured Tb/s-class signal constellations generated by software-defined transmitters.**

SOFTWARE-DEFINED MODULATION

With transmitter-site DSP and collaborative receiver-side DSP, the modulation format can be software defined to enable the optimization of system performance depending on the link conditions. Figure 6 shows some exemplary signal constellations generated in Tb/s-class superchannels by software-defined transmitters. In addition to conventional square QAM, advanced modulation formats such as iterative polar modulation (IPM) [16] and four-dimensional (4-D) modulation formats [17], [18] that jointly utilize the four orthogonal dimensions (two polarizations and two quadratures) can be realized. With optimized 4-D constellation shaping, the signal's tolerance to both linear noise and nonlinear fiber transmission impairments can be improved.

ADVANCED CODING

FEC is a powerful technique to improve transmission performance. Since soft information is available within a coherent DSP ASIC, an SD FEC can be readily implemented [19]. In a recent experiment [6], a low-density parity check (LDPC) code with a rate of 0.864 (16% coding overhead) was implemented for the inner SD-FEC, and a rate 0.935 (7% coding overhead) HD FEC code was assumed as the outer code to achieve the typical optical networking requirement of post-FEC bit-error ratios (BERs) on the order of 10^{-15}. Figure 7 shows the back-to-back BER performance of the 1.5-Tb/s superchannel based on offline digital signal processing. Remarkably, there is essentially no extra penalty due to guard-banded multiplexing (MUX), thanks to the well-confined optical spectrum of the OFDM signals making up the superchannel. For a typical 7%-overhead outer HD-FEC, the correction threshold is about 4×10^{-3} for a final BER of $< 10^{-15}$. For the output BER of the SD-FEC decoder to be below the correction threshold, the input BER of the SD-FEC needs to be $< 2.7 \times 10^{-2}$ when 15 decoder iterations are used. (Note that optical systems experiments rarely implement actual FEC but rather use the pre-FEC BER to assess transmission



[FIG7] **The measured back-to-back BER performance of the 1.5-Tb/s superchannel before and after the SD-FEC decoding. Insets: recovered subcarrier constellations before and after the guard-banded multiplexing (after [6]).**

quality, which implicitly assumes sufficient scrambling to ensure the statistical independence of errors, and in the case of SD-FEC also assumes additive white Gaussian noise. Since either premise may be violated under nonlinear optical transmission conditions, advanced experiments implement at least the inner SD-FEC decoder.) Figure 8 shows the measured BER (based on offline signal processing) of all eight 30-Gbaud PDM-OFDM-16QAM signals that make up the 1.5-Tb/s superchannel after 5,600-km transmission at a superchannel launch power (P_{in}) of 9 dBm. The BER after SD-FEC is below the outer HD-FEC threshold for each of the eight subcarriers.

More recently, a coded 231.5-Gb/s RGI-CO-OFDM signal, with constellation shaping through 256-IPM subcarrier modulation, was transmitted over a distance of 800 km in ULAF, achieving an ISE as high as 11.15 b/s/Hz [14]. This study shows the promise of combining coding with modulation and nonlinear compensation to achieve substantially higher optical transmission performance than using coding alone. With iterative decoding based on a maximum a posteriori decoder and a LDPC FEC, 30.58-Tb/s transmission over 7,230 km at an SE of 6.1 b/s/Hz has recently been demonstrated, achieving a record capacity-distance product of 221 Pb/s·km for a submarine transmission link [18].

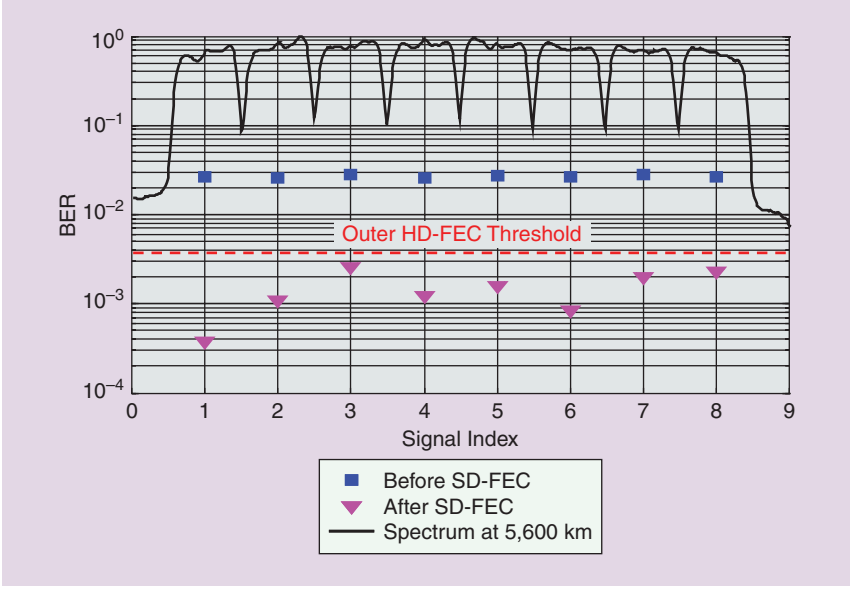
JOINT PROCESSING OF SUPERCHANNEL CONSTITUENTS

As the constituents of a superchannel share the same optical transmission path, they experience some common transmission properties. Hence, the monitoring of some transmission properties of the superchannel can be simplified by monitoring only one of the constituents, or more accurate measurements can be obtained by averaging certain channel estimates across superchannel constituents. For example, if the source lasers of the superchannel constituents are frequency-locked, as in the case of CO-OFDM, the frequency estimation of a superchannel can be simplified. As the signal fields of all constituents are available at the receiver, joint DSP can be applied to mitigate certain impairments, such as coherent crosstalk between adjacent signals [20].

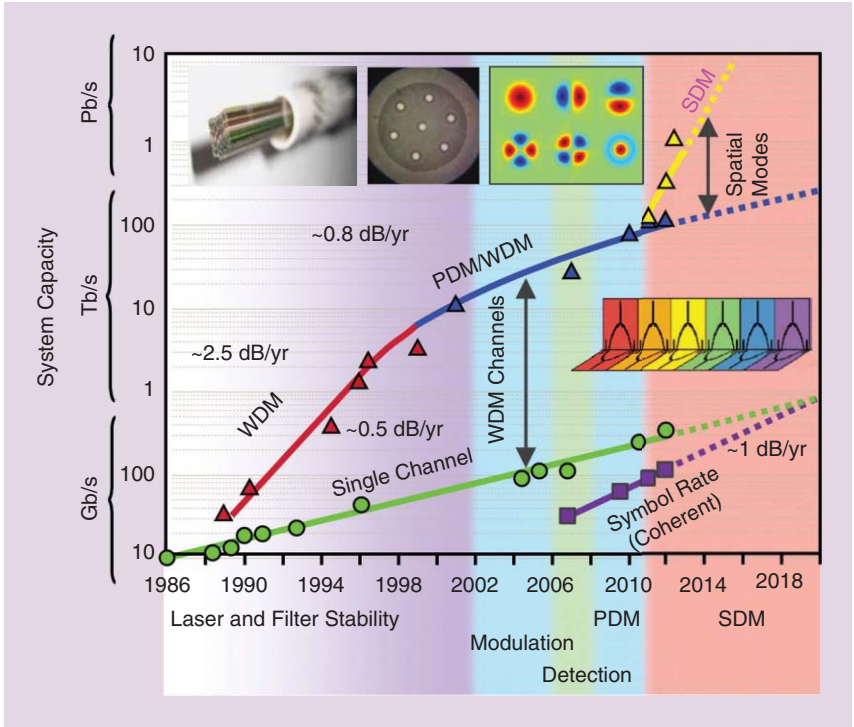
FIBER NONLINEARITY IS A MAJOR TRANSMISSION IMPAIRMENT IN OPTICAL FIBER COMMUNICATIONS.

Fiber nonlinearity is a major transmission impairment in optical fiber communications. In contrast to static nonlinearities encountered in some radio-frequency systems, fiber transmission suffers from the interplay of power-dependent phase and amplitude distortions within and across WDM channels, coupled with chromatic dispersion that induces correlation across many, sometimes even hundreds of symbols. Mathematically, fiber transmission in two polarizations is described by the coupled nonlinear Schrödinger equation (NLS). Digital backpropagation (DBP) as a means to numerically invert the NLS at the receiver [21], [22] has been introduced to mitigate signal-to-signal nonlinear interactions such as self-phase modulation (SPM) and cross-phase modulation (XPM). With the availability of the signal fields of all the constituents of a superchannel, joint DBP can be applied to mitigate both SPM and XPM impairments. XPM compensation using DBP via the coupled nonlinear Schrödinger equation has been experimentally demonstrated [22]. It was found that as XPM is phase-insensitive, the phase-locking among the interacting signals is not needed for XPM compensation. The DSP complexity of DBP based fiber nonlinearity compensation is much higher than that of a conventional coherent receiver. This calls for DSP-efficient fiber nonlinearity compensation schemes. Recently, novel fiber nonlinearity mitigation schemes based on the Volterra series [23] and phase-conjugated twin waves [24] have been introduced.

For DSP-efficient dispersion compensation of superchannels, subband equalizers [25] and filter-bank-based digital subbanding [26] have been proposed. In these approaches, a superchannel is divided into multiple subbands, which are individually dispersion-compensated to achieve higher DSP efficiency and easier parallelization. As briefly mentioned earlier, joint FEC using all the superchannel constituents may provide additional performance improvements. In [27], a coherent modem is implemented by coding across two PDM-QPSK channels spaced at 200 GHz and is found to reduce the transmission penalty due to PMD and polarization-dependent loss as compared to that without joint FEC processing. In a recent WDM transmission experiment,



[FIG8] The measured BER of all the constituents making up the superchannel after 5,600-km transmission (after [6]).



[FIG9] The evolution of optical transmission system capacity. Experimentally achieved single-channel bit rates are also shown. The possible future evolution of SDM is also shown (after [1]).

short-wavelength channels are worse performing that long-wavelength channels, and joint FEC processing using a long-wavelength channel and a short-wavelength channel has also been found to be beneficial [18].

FUTURE OUTLOOK

As today's WDM communication technology has already taken advantage of all degrees of freedom of a lightwave in a single-mode fiber, particularly frequency, polarization, amplitude, and phase, further multiplicative growth has to explore new degrees of freedom. Space-division multiplexing (SDM) is expected to further scale network capacities, using parallel strands of single-mode fiber, uncoupled or coupled cores of multicore fiber, or even individual modes of a few-mode fiber [28]. Figure 9 illustrates the capacity growth trend of optical fiber communication systems [1]. In SDM systems that induce coupling between the spatial modes, multiple-input, multiple-output (MIMO) DSP has been demonstrated to undo the mode coupling and enable penalty-free mode multiplexing [28].

Similar to the case of superchannel transmission, joint DSP can also be applied in the case of SDM transmission to efficiently compensate for transmission impairments that are common to the spatial modes involved. Joint DSP for CFR in such spatial superchannel transmission has recently been experimentally demonstrated [29].

DSP WILL CONTINUE TO PLAY A KEY ROLE IN THE FUTURE EVOLUTION OF OPTICAL NETWORKS TO SUPPORT THE EVER-INCREASING DEMAND FOR INTERNET TRAFFIC AND CLOUD COMPUTING.

One important goal of optical fiber communications is to support the communication capacity growth in a sustainable fashion, which means continuously lowering the cost per information bit. In addition to improving transmission capacity and/or performance for a given set of hardware components, this can be achieved by using advanced DSP to compensate for the shortcomings of lower-cost and highly integrated O/E components. For example, compact all-InP integrated laser and modulator assemblies and silicon photonics-based devices, when complemented with advanced DSP, are promising candidates to reduce the cost, size, and power consumption of future optical transport systems [30].

CONCLUSIONS

DSP has played an important role in supporting the recent capacity expansion of optical core networks based on coherent modulation and detection. To enable sustainable capacity growth, the cost per information bit needs to continuously decrease. Superchannels, with their improved system SE and their natural compatibility with large-scale integration, are expected to be well suited to meet this demand. DSP may see more emerging applications in superchannel transmission and in relaxing the physical requirements on optical components, and continue to play a key role in the future evolution of optical networks to support the ever-increasing demand for Internet traffic and cloud computing.

AUTHORS

Xiang Liu (xiang.liu@alcatel-lucent.com) is a distinguished member of technical staff at Bell Labs, Alcatel-Lucent. He received his Ph.D. degree in applied physics from Cornell University in 2000. Since joining Bell Labs, he has been working on high-speed optical communication technologies including advanced modulation formats, digital coherent detection, and fiber nonlinear impairment mitigation. He was recognized as a core member of the "100 Gb/s Coherent (Long Haul–High Capacity WDM Interface) Team" that received the 2010 Bell Labs President's Award. He has authored/coauthored over 250 journal and conference papers and holds 51 U.S. patents. He is a fellow of the Optical Society of America and an associated editor of *Optics Express*.

S. Chandrasekhar (sc@alcatel-lucent.com) is a distinguished member of technical staff at Bell Labs, Alcatel-Lucent. He received the Ph.D. degree in physics from the University of Bombay, India, in 1985, and joined Bell Labs, New Jersey in

1986. He was recognized as a member of the "100 Gb/s Coherent (Long Haul–High Capacity WDM Interface) Team" that received the 2010 Bell Labs President's Award. His current interests include coherent optical transmission for high SE transport and networking beyond 100 Gb/s. He is a Fellow of the IEEE and the OSA. He received the IEEE LEOS Engineering Achievement Award in 2000 and the Optical Society of America Engineering Excellence Award in 2004 for his contributions to optoelectronic integrated circuits and WDM systems.

Peter J. Winzer (peter.winzer@alcatel-lucent.com) heads the Optical Transmission Systems and Networks Research Department at Bell Labs, Alcatel-Lucent. He received his Ph.D. degree in electrical engineering from the Vienna University of Technology, Austria, and joined Bell Labs in 2000, focusing on hardware and architectural aspects of fiber optic transmission systems and networks. He has published over 300 journal and conference papers, holds over 30 U.S. patents, and is actively involved in technical and organizational tasks with the IEEE and the Optical Society of America, currently serving as editor-in-chief of *Journal of Lightwave Technology*. He is a distinguished member of technical staff at Bell Labs and a Fellow of the IEEE and the Optical Society of America.

REFERENCES

[1] P. J. Winzer, "High-spectral-efficiency optical modulation formats," *J. Lightwave Technol.*, vol. 30, no. 24, pp. 3824–3835, 2012.

[2] T. J. Xia, S. Gringeri, and M. Tomizawa, "High-capacity optical transport networks," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 170–178, 2012.

[3] S. Chandrasekhar, X. Liu, B. Zhu, and D. W. Peckham, "Transmission of a 1.2-Tb/s 24-carrier no-guard-interval coherent OFDM superchannel over 7200-km of ultra-large-area fiber," in *Proc. European Conf. Optical Communication (ECOC) 2009*, Sept., paper PD2.6.

[4] Y. Ma, Q. Yang, Y. Tang, S. Chen, and W. Shieh, "1-Tb/s single-channel coherent optical OFDM transmission over 600-km SSMF fiber with subwavelength bandwidth access," *Opt. Express*, vol. 17, no. 11, pp. 9421–9427, 2009.

[5] Y.-K. Huang, E. Ip, Z. Wang, M.-F. Huang, Y. Shao, and T. Wang, "Transmission of spectral efficient super-channels using all-optical OFDM and digital coherent receiver technologies," *J. Lightwave Technol.*, vol. 29, no. 24, pp. 3838–3844, 2011.

[6] X. Liu, S. Chandrasekhar, P. J. Winzer, T. Lotz, J. Carlson, J. Yang, G. Cheren, and S. Zederbaum , "1.5-Tb/s guard-banded superchannel transmission over 56x 100-km (5600-km) ULAF using 30-Gbaud pilot-free OFDM-16QAM signals with 5.75-b/s/Hz net spectral efficiency," in *Proc. European Conf. Optical Communication (ECOC) 2012*, Sept., post-deadline paper Th.3.C.5.

[7] G. Raybon, S.Randel, A.Adamiecki, P. J. Winzer, L. Salamanca, R. Urbanke, S. Chandrasekhar, A. Konczykowska, F. Jorge, J.-Y. Dupuy, L. L.Buhl, S.Draving, M. Grove, and K. Rush, "1-Tb/s dual-carrier 80-Gbaud PDM-16QAM WDM transmission at 5.2 b/s/Hz over 3200km," in *IEEE Photonics Conf. (IPC)*, 2012, paper PD1.2.

[8] G. Bosco, V. Curri, A. Carena, P. Poggiolini, and F. Forghieri, "On the performance of Nyquist-WDM terabit superchannels based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM subcarriers," *J. Lightwave Technol.*, vol. 29, no. 1, pp. 53–61, 2011.

[9] A. Sano, T. Kobayashi, S. Yamanaka, A. Matsuura, H. Kawakami, Y. Miyamoto, K. Ishihara, and H. Masuda, "102.3-Tb/s (224 x 548-Gb/s) C- and extended L-band all-Raman transmission over 240 km using PDM-64QAM single carrier FDM with digital pilot tone," in *Proc. Optical Fiber Communication Conf. and Expo. (OFC) and Nat. Fiber Optic Engineers Conf. (NFOEC) 2012*, paper PDP5C.3.

[10] X. Liu, S. Chandrasekhar, P. Winzer, B. Zhu, D.W. Peckham, S. Draving, J. Evangelista, N. Hoffman, C. J. Youn, Y.-H. Kwon, and E. S. Nam, "3x 485-Gb/s WDM transmission over 4800 km of ULAF and 12x 100-GHz WSSs using CO-OFDM and single coherent detection with 80-GS/s ADCs," in *Proc. Optical Fiber Communication Conf.*, paper JThA037.

[11] X. Zhou and L. E. Nelson, "400G WDM transmission on the 50 GHz grid for future optical networks," *J. Lightwave Technol.,* vol. 30, no. 24, pp. 3779–3792, 2012.

[12] S. Haykin, "Signal processing: Where physics and mathematics meet," *IEEE Signal Processing Mag.*, vol. 18, no. 4, pp. 6–7, 2001.

[13] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE J. Select. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, 2010.

[14] T. H. Lotz, X. Liu, S. Chandrasekhar, P. J. Winzer, H. Haunstein, S. Randel, S. Corteselli, B. Zhu, and D. W. Peckham, "Coded PDM-OFDM transmission with shaped 256-iterative-polar-modulation achieving 11.15-b/s/Hz intrachannel spectral efficiency and 800-km reach," *J. Lightwave Technol.*, vol. 31, no. 4, pp. 538–545, 2013.

[15] Q. Zhuge, M. Morsy-Osman, X. Xu, M. E. Mousa-Pasandi, M. Chagnon, Z. A. El-Sahn, and D. V. Plant, "Pilot-aided carrier phase recovery for M-QAM using superscalar parallelization based PLL," *Opt. Express*, vol. 20, no. 17, pp. 19599–19609, 2012.

[16] I. B. Djordjevic, H. G. Batshon, L. Xu, and T. Wang, "Coded polarization-multiplexed iterative polar modulation (PM-IPM) for beyond 400 Gb/s serial optical transmission," in *Proc. Optical Fiber Communication Conf. (OFC)/Nat. Fiber Optic Engineers Conf. (NFOEC)*, 2010, paper OMK2.

[17] M. Karlsson and E. Agrell, "Four-dimensional optimized constellations for coherent optical transmission systems," in *European Conf. Optical Communication (ECOC) 2010*, paper We8C3.

[18] H. Zhang, H. G. Batshon, D. Foursa, M. Mazurczyk, J. Cai, C. Davidson, A. Pilipetskii, G. Mohs, and N. Bergano, "30.58 Tb/s transmission over 7,230 km using PDM half 4D-16QAM coded modulation with 6.1 b/s/Hz spectral efficiency," in *Proc. Optical Fiber Communication Conf.*, 2013, paper OTu2B.3.

[19] F. Chang, K. Onohara, and T. Mizuochi, "Forward error correction for 100 G transport networks," *IEEE Commun. Mag.*, vol. 48, no. 3, pp. S48–S55, 2010.

[20] C. Liu, J. Pan, T. Detwiler, A. Stark, Y.-T. Hsueh, G.-K. Chang, and S. E. Ralph, "Super receiver design for superchannel coherent optical systems," in *Proc. SPIE 8284, Next-Generation Optical Communication: Components, Sub-Systems, and Systems*, 2012, p. 828405.

[21] E. Mateo, L. Zhu, and G. Li, "Impact of XPM and FWM on the digital implementation of impairment compensation for WDM transmission using backward propagation," *Opt. Express*, vol. 16, no. 20, pp. 16124–16137, 2008.

[22] E. Ip, Y. Huang, E. Mateo, Y. Aono, Y. Yano, T. Tajima, and T. Wang, "Interchannel nonlinearity compensation for 3 $\lambda \times 114$-Gb/s DP-8QAM using three synchronized sampling scopes," in *Proc. Optical Fiber Communication Conf. and Expo. (OFC) and Nat. Fiber Optic Engineers Conf. (NFOEC) 2012*, paper OM3A.6.

[23] G. Shulkind and M. Nazarathy, "Nonlinear digital back propagation compensator for coherent optical OFDM based on factorizing the volterra series transfer function," *Opt. Express*, vol. 21, no. 11, pp. 13145–13161, 2013.

[24] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Photon.*, vol. 7, pp. 560–568, July 2013.

[25] K.-P. Ho, "Subband equaliser for chromatic dispersion of optical fibre," *Electron. Lett.*, vol. 45, no. 24, pp. 1224–1226, 2009.

[26] A. Tolmachev and M. Nazarathy, "Filter-bank-based efficient transmission of reduced-guard-interval OFDM," *Opt. Express*, vol. 19, no. 26, pp. 370–384, 2011.

[27] J. Rahn, K. Croussore, G. Goldfarb, S. Kumar, M. Mitchell, V. Dominic, B. Taylor, Y. Park, V. Adavani, V. Shyamsundar, M. R. Patil, S. Yu, P. Freeman, A. Mathur, M. Ziari, and D. F. Welch, "Transmission improvement through dual-carrier FEC gain sharing," in *Proc. Optical Fiber Communication Conf., 2013,* paper OW1E.5.

[28] P. J. Winzer, "Optical networking beyond WDM," *IEEE Photon. J.*, vol. 4, no. 2, pp. 647–651, Apr. 2012.

[29] M. D. Feuer, L. E. Nelson, X. Zhou, S. L. Woodward, R. Isaac, B. Zhu, T. F. Taunay, M. Fishteyn, J. M. Fini, and M. F. Yan, "Joint digital signal processing receivers for spatial superchannels," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1957–1960, 2012.

[30] S. Chandrasekhar, X. Liu, P. J. Winzer, J. E. Simsarian, and R. A. Griffin, "Small-form-factor all-InP integrated laser vector modulator enables the generation and transmission of 256-Gb/s PDM-16QAM modulation format," in *Proc. Optical Fiber Communication Conf.*, 2013, post-deadline paper PDPB5.6.

[SP]

Sercan Ö. Arık, Joseph M. Kahn, and Keang-Po Ho

MIMO Signal Processing for Mode-Division Multiplexing

An overview of channel models and signal processing architectures

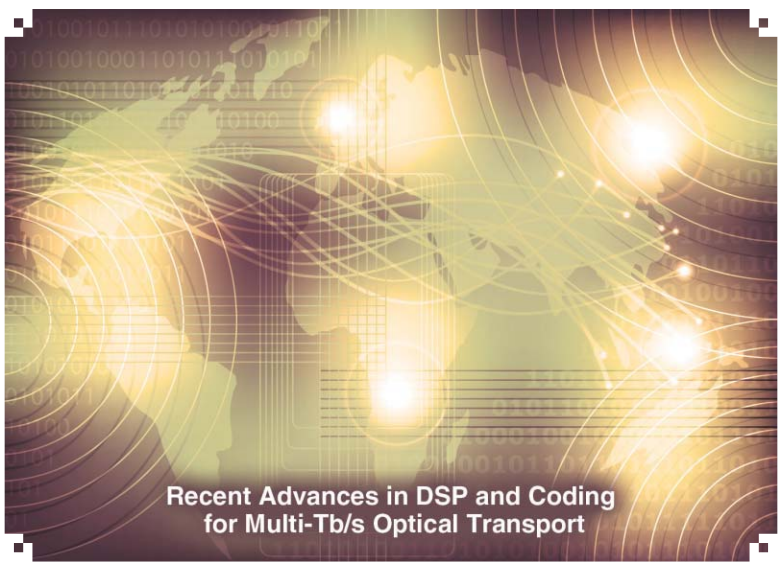


IMAGE LICENSED BY INGRAM PUBLISHING /JANMIKS

We present the fundamentals of multiple-input, multiple-output (MIMO) signal processing for mode-division multiplexing (MDM) in multimode fiber (MMF). As an introduction, we review current long-haul optical transmission systems and how continued traffic growth motivates study of new methods to increase transmission capacity per fiber. We describe the key characteristics of MIMO channels in MMF, contrasting these with wireless MIMO channels. We review MMF channel models, the statistics derived from them, and their implications for MDM system performance and complexity. We show that optimizing performance and complexity requires management of channel parameters—particularly group delay (GD) spread and mode-dependent loss and gain—by design of transmission fibers and optical amplifiers, and by control of mode coupling along the link. We describe a family of fibers optimized for low GD spread, which decreases with an increasing number of modes. We compare the performance and complexity of candidate MIMO signal processing architectures in a representative long-haul system design, and show that programmable frequency-domain equalization (FDE) of chromatic dispersion (CD) and adaptive FDE of modal dispersion (MD) is an attractive combination. We review two major algorithms for adaptive FDE of MD—least mean

Digital Object Identifier 10.1109/MSP.2013.2290804
Date of publication: 12 February 2014

squares (LMS) and recursive least squares (RLS)—and analyze their complexity, throughput efficiency, and convergence time. We demonstrate that, with careful physical link design and judicious choice of signal processing architectures, it is possible to overcome MIMO signal processing challenges in MDM systems.

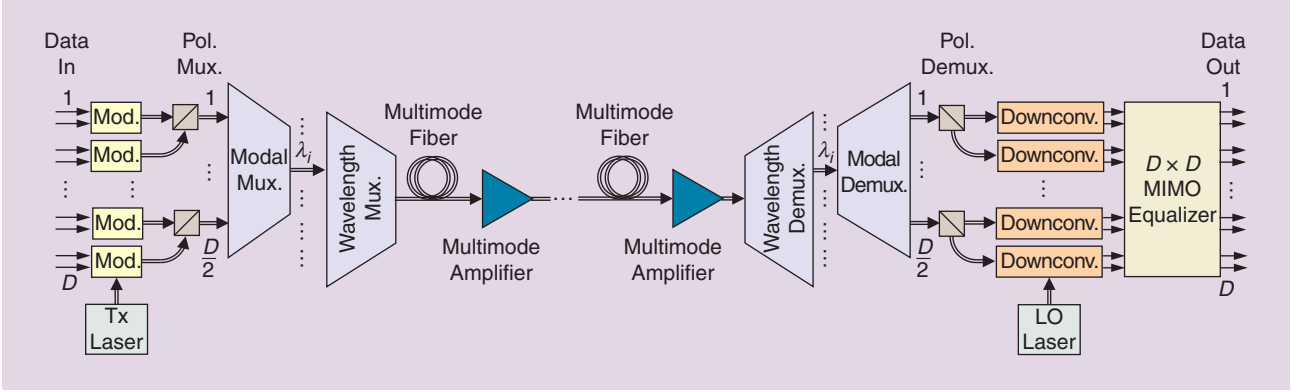
INTRODUCTION

Spurred by high-definition video streaming, multimedia file sharing, cloud computing, mobile networking, online gaming, and other information technologies, worldwide data traffic is growing at a rate estimated to exceed 50% annually [1]. Long-haul optical fiber networks form the backbone of the Internet, and scaling to higher throughput and lower cost per bit is essential to ensure continued growth of information technologies. For four decades, optical networks have exploited significant progress in physical technologies, including low-loss single-mode fibers (SMFs), high-performance lasers and modulators, and low-noise optical amplifiers. In parallel, communication methods have evolved to multiplex information in various physical dimensions in SMFs, including time, frequency, quadrature phase and polarization, a trend hastened in recent years by coherent optical receivers using large-scale digital circuits for signal processing and error-correction decoding. The transmission capacity per fiber, after decades of exponential growth sustained by these technologies [2], is now approaching fundamental information-theoretic limits imposed by optical amplifier noise and by the nonlinear response of the silica fiber medium [3]. Digital signal processing methods can mitigate nonlinear effects [4], enabling denser constellations and increasing transmission capacity, but at the cost of high complexity and power consumption.

WITH CAREFUL PHYSICAL LINK DESIGN AND JUDICIOUS CHOICE OF SIGNAL PROCESSING ARCHITECTURES, IT IS POSSIBLE TO OVERCOME MIMO SIGNAL PROCESSING CHALLENGES IN MDM SYSTEMS.

New types of fibers, which provide more spatial dimensions, provide intriguing options for increasing transmission capacity per fiber. Multicore fibers provide one option [2]. By incorporating multiple cores in a single strand of glass, they provide a proportional increase in the number of spatial dimensions. The plurality of modes propagating in MMF provides another option for increasing the number of spatial dimensions [2]. A waveguide mode is defined as a pattern of the optical electric field that propagates without changing, apart from an amplitude change and phase shift. Different waveguide modes are mutually orthogonal, so they provide independent spatial dimensions for data transmission. Although propagation over long distances causes coupling between signals in different modes, received signals can be separated by MIMO processing. Throughout this article, we use D to denote the total number of modes, including spatial and polarization dimensions. A "single-mode" fiber allows propagation of one spatial mode in two polarizations, a total of $D = 2$ modes. A 2×2 MIMO transmission in SMF is also known as polarization-division multiplexing (PDM). By increasing the core diameter and adjusting the refractive index profile, an MMF with circular cross section can be designed to support $D = 6, 10, 12, 16, 20, 24, 30, \ldots$ modes. Fibers supporting these small values of D are sometimes called "few-mode fibers" to distinguish them from the MMFs used in short-range data communications, which support up to hundreds of modes.

Figure 1 shows the general structure of a long-haul MDM transmission system. At the transmitter, at each wavelength, D signals are modulated with data streams and mapped into two polarizations and $D/2$ spatial modes, a fixed transmit basis comprising D orthogonal modes. Signals from up to about 100 transmitters at different wavelengths are combined by a wavelength multiplexer, as in a conventional SMF system. As signals propagate through multiple spans of fiber, loss is compensated by optical amplifiers designed to amplify multiple modes at multiple wavelengths. At the receiving end, different wavelengths are demultiplexed. The signal at each wavelength is separated into $D/2$ spatial modes and two polarizations, i.e., projected onto a fixed receive basis of D modes. The signal in each mode is mixed with a local oscillator, downconverting it to baseband. After sampling, digital signal processing implements several essential functions, which



[FIG1] A long-haul optical transmission system using mode-division multiplexing.

[TABLE 1] A COMPARISON OF MIMO CHANNELS IN WIRELESS SYSTEMS AND IN MDM OPTICAL SYSTEMS.

ATTRIBUTE	WIRELESS	OPTICAL
CHANNEL INPUTS AND OUTPUTS	SIGNALS AT ANTENNAS	COMPLEX AMPLITUDES IN MODES
NUMBER OF INPUTS OR OUTPUTS	2, 3, 4, 5, 6,…	2, 6, 10, 12, 16, 20, 24, 30,…
CARRIER FREQUENCY	0.8–6 GHz (MOST PHONE AND Wi-Fi BANDS)	185–196 THz (OPTICAL AMPLIFIER BANDS)
SYMBOL RATE	10 s OF Mbaud	10 s OF Gbaud
ADDITIVE NOISE	THERMAL NOISE (CLASSICAL)	SPONTANEOUS EMISSION (QUANTUM)
INTERFERENCE AND DISTORTION	COCHANNEL INTERFERENCE (LINEAR)	INTRACHANNEL DISTORTION (NONLINEAR) INTERCHANNEL INTERFERENCE (NONLINEAR)
DISPERSION	MULTIPATH PROPAGATION	FREQUENCY-DEPENDENT GROUP INDEX (CD) MODE-DEPENDENT GROUP INDEX (MD)
DELAY SPREAD (rms)	10 s OF ns (INDOORS) 100 s OF ns (OUTDOORS)	100 s OF ns (LONG HAUL)
FADING	MULTIPATH PROPAGATION	OPTICAL AMPLIFIER MODE-DEPENDENT GAIN
AMPLITUDE VARIATIONS IN CHANNEL MATRIX	LARGE (e.g., I.I.D. RAYLEIGH)	SMALL (e.g., NEAR-UNITARY)
ORIGIN OF CHANNEL VARIATIONS	TERMINAL MOTION (DOPPLER)	FIBER MOTION (MODE COUPLING)
CHANNEL VARIATION TIME SCALE	~1 ms (DRIVING OR WALKING SPEEDS)	10 TO 100 s OF μs (VIBRATION OR IMPACT)
DIVERSITY	MAY REQUIRE SPACE-TIME CODING	INHERENT FREQUENCY DIVERSITY

include automatic gain control, timing recovery, carrier recovery, and MIMO equalization that compensates for CD, MD, and mode coupling, separating the D multiplexed data signals [5].

Making long-haul MDM systems a practical reality requires development of new fibers and other optical components supporting multiple spatial modes, including modal (de)multiplexers, wavelength (de) multiplexers, wavelength-selective switches, and inline optical amplifiers. Implementing transceivers supporting multiple spatial modes while reducing the cost, size, and power consumption per information bit will require extensive photonic and electronic integration. Likewise, it will require efficient, high-performance MIMO signal processing architectures, which are the subject of this article.

MAKING LONG-HAUL MDM SYSTEMS A PRACTICAL REALITY REQUIRES DEVELOPMENT OF NEW FIBERS AND OTHER OPTICAL COMPONENTS SUPPORTING MULTIPLE SPATIAL MODES.

COMPARISON OF WIRELESS AND OPTICAL MIMO TRANSMISSION

Over the past 15 years, MIMO wireless transmission has progressed from its information-theoretic foundations [6] to recent commercialization. MDM, as an optical form of MIMO transmission, shares some similarities with wireless MIMO transmission. But optical and wireless MIMO channels exhibit some fundamental differences. A comparison of these two types of MIMO channels is given in Table 1.

Wireless systems use lower carrier frequencies, where bandwidth is limited and regulated. Optical systems use far higher carrier frequencies, where bandwidth is limited mainly by the gain spectra of optical amplifiers. In wireless, the dominant noise is thermal, since the carrier photon energy is far less than thermal energy, while in dense user environments, linear cochannel interference from other users can become the dominant "noise." In optical systems, where the carrier photon energy is far greater than thermal energy, the dominant noise is spontaneous emission from inline optical amplifiers. In glass fibers, the refractive index is weakly dependent on the local light intensity. This so-called Kerr nonlinearity causes self-phase modulation, a form of intrachannel distortion, and also cross-phase modulation and four-wave mixing, which are forms of adjacent-channel interference. In current systems, these nonlinear effects are treated as "noise." It is possible to mitigate them partially by digital backpropagation [4], which would allow some increase in spectral efficiency, but at the cost of increased signal processing complexity, which precludes its implementation in current hardware technologies.

In wireless MIMO systems, multipath propagation causes linear distortion, with a delay spread depending mainly on the environment. Multipath also causes fading, which can cause the signal energy in a narrow band to vary by tens of decibels. These wireless channel properties vary randomly at rates depending on the speed of transceiver or scatterer motion. In a narrowband MIMO system, fading causes the instantaneous capacity to fluctuate, potentially causing an outage, and if a system does not provide sufficient frequency or time diversity, then space-time coding may be required to achieve reliable operation [6].

In optical MIMO systems, linear distortion arises from two mechanisms. First, the group refractive index of a mode is frequency dependent, an effect called CD that is similar in SMF and MMF. At frequencies employed in long-haul systems, CD arises from the inherent material dispersion of glass and from frequency-dependent confinement of the mode in the waveguide, an opposing contribution that is typically smaller in magnitude. Second, the group refractive index is slightly different for different modes. In SMF, with only $D = 2$ polarization modes, this is called polarization-mode dispersion (PMD) [7] and causes a delay spread of

only 10s of ps. In MMF, the effect is called MD and can cause a delay spread up to 100s of ns. In optical MIMO systems, fading arises because different modes may experience different gains in optical amplifiers (or different losses in fibers), effects collectively called mode-dependent loss (MDL) [8]. In SMF $D = 2$, this effect is called polarization-dependent loss (PDL). If managed carefully, root-mean-square (rms) MDL fluctuations are expected to be smaller than 10 dB, so the channel matrix remains near-unitary. Both MD and MDL may vary randomly over time and frequency because of random coupling between modes caused by perturbations of the fiber, including index imperfections and bends [9]. In response to mechanical perturbations of the fiber, optical MIMO channels are expected to fluctuate on the time scale of 10 s–100 s of μs [10], [11]. In optical MIMO systems, the product of delay spread and symbol rate far exceeds unity, yielding efficient frequency diversity that reduces capacity fluctuations [12]. As a result, all D spatial dimensions can be exploited for multiplexing, and none need be sacrificed for diversity.

A final comparison concerns the choice of single- versus multicarrier modulation. The latter is often implemented as orthogonal frequency-division multiplexing (OFDM). Comparison between the two approaches is a perennial topic in wireless research and has motivated study of multicarrier techniques for optical PDM and MDM [13]. In wireless communications, when single- and multicarrier systems are optimally designed, they achieve similar performance and complexity. By contrast, in long-haul optical systems, single-carrier modulation is less impacted by fiber nonlinearity than multicarrier modulation [14], and single-carrier has been used in all commercial systems to date. Nevertheless, the single-carrier FDE methods described below can be adapted to OFDM with only minor modifications.

MDM CHANNEL MODEL AND STATISTICS

In this section, we describe the modeling of MDM channels in MMF, emphasizing the effect of mode coupling on the channel GD spread and fading statistics. We provide an MDM system example to illustrate how the GD spread affects MIMO signal processing complexity and performance.

MATRIX PROPAGATION MODEL

A signal at frequency Ω propagating along the z-axis in a MMF can be represented as an electric field vector $\mathbf{E}(x, y, z, \Omega) = \sum_{i=1}^{D} A_i(z, \Omega)\mathbf{E}_i(x, y, \Omega)$, where the $\mathbf{E}_i(x, y, \Omega)$, $i = 1, \ldots, D$, are the orthonormal waveguide modes of an ideal unperturbed fiber. Since these modes are fixed, we can represent the propagating signal by a $D \times 1$ vector of complex amplitudes $\mathbf{A}(z, \Omega) = [A_1(z, \Omega) \ \cdots \ A_D(z, \Omega)]^T$, which we write as $\mathbf{A}(\Omega)$. Ignoring noise, linear propagation through the fiber can be described by

$$\mathbf{A}^{(\text{out})}(\Omega) = \mathbf{M}(\Omega)\mathbf{A}^{(\text{in})}(\Omega), \tag{1}$$

where $\mathbf{A}^{(\text{in})}(\Omega)$ and $\mathbf{A}^{(\text{out})}(\Omega)$ represent the input and output and $\mathbf{M}(\Omega)$ is a propagation operator described by a $D \times D$ matrix. In SMF ($D = 2$), the $\mathbf{A}(z, \Omega)$ and $\mathbf{M}(\Omega)$ are the Jones vector and Jones matrix [7].

Because of perturbations inducing mode coupling, signals evolve randomly as they propagate through a fiber, and their complex envelopes remain correlated over a characteristic correlation length. A fiber can be modeled as a concatenation of multiple sections [9], each of length roughly equal to the correlation length. This is a generalization of the multisection model used in modeling PMD and PDL in SMF [15]. As shown in Figure 2, a long-haul system is composed of K_{amp} spans, each comprising a fiber of length L_{amp}, followed by an amplifier to compensate for the mode-averaged loss of the fiber. Each span is subdivided into K_{sec} sections, each of length L_{sec}. The overall system has $K_{\text{tot}} = K_{\text{amp}} K_{\text{sec}}$ sections and total length $L_{\text{tot}} = K_{\text{amp}} L_{\text{amp}} = K_{\text{amp}} K_{\text{sec}} L_{\text{sec}}$.

The propagation operator can be represented as [9]

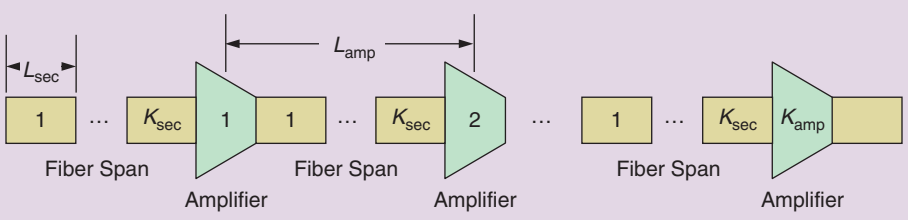
$$\mathbf{M}_{\text{tot}}(\Omega) = \exp\left(-\frac{j}{2}\Omega^2 \bar{\beta}_2 L_{\text{tot}}\right) \cdot \mathbf{M}(\Omega). \tag{2}$$

The exponential factor represents mode-averaged propagation, where $\bar{\beta}_2$ represents mode-averaged CD per unit length (for simplicity, we neglect the mode-averaged GD). The matrix $\mathbf{M}(\Omega)$ represents mode-dependent effects, including MDL, MD, and crosstalk caused by mode coupling. It can be written as a product over the K_{amp} spans

$$\mathbf{M}(\Omega) = \prod_{k=1}^{K_{\text{amp}}} \mathbf{M}^{(k)}(\Omega), \tag{3}$$

where $\mathbf{M}^{(k)}(\Omega)$ represents propagation in the kth span. Assuming MDL in the transmission fibers is negligible compared to that in the amplifiers [5], it can be written

$$\mathbf{M}^{(k)}(\Omega) = \text{diag}\left[\exp\left(\frac{g_1^{(k)}}{2}\right) \ \cdots \ \exp\left(\frac{g_D^{(k)}}{2}\right)\right] \cdot \prod_{l=1}^{K_{\text{sec}}} \mathbf{V}^{(k,l)} \mathbf{\Lambda}(\Omega) \mathbf{U}^{(k,l)^H}. \tag{4}$$



[FIG2] The multisection model of a long-haul MDM system.

The first factor in $\mathbf{M}^{(k)}(\Omega)$ describes uncoupled MDL in the kth amplifier. The $g_i^{(k)}$, $i = 1,\ldots,D$, are the uncoupled modal gains measured in log power gain units (proportional to decibels). They satisfy $\sum_{i=1}^{D} g_i^{(k)} = 0$ and have rms spread σ_g (assumed to be the same for all amplifiers). The second factor in $\mathbf{M}^{(k)}(\Omega)$ describes MD and mode coupling in the kth span and is a product over the K_{sec} sections. The matrix $\mathbf{\Lambda}(\Omega)$ describes uncoupled MD in each section (assumed to be the same for all sections)

$$\mathbf{\Lambda}(\Omega) = \text{diag}[\exp(-j\Omega\tau_1) \ \ldots \ \exp(-j\Omega\tau_D)], \tag{5}$$

where the τ_i, $i = 1,\ldots,D$, are the uncoupled modal GDs, which satisfy $\sum_{i=1}^{D} \tau_i = 0$ and have rms spread σ_τ. The matrices $\mathbf{V}^{(k,l)}$ and $\mathbf{U}^{(k,l)}$ are frequency-independent unitary matrices representing the random mode coupling in the lth section of the kth span. The symbol ^{H} denotes Hermitian conjugate.

Mode coupling in fibers may be induced by random or intentional index perturbations, bends or stresses. Many random perturbations have low-pass longitudinal power spectra, so they strongly couple modes having nearly equal propagation constants, but weakly couple modes having highly unequal propagation constants [9]. When mode coupling occurs only between nearly degenerate modes in the same mode group, which is the weak-coupling regime, $\mathbf{V}^{(k,l)}$ and $\mathbf{U}^{(k,l)}$ are block unitary matrices. When mode coupling occurs with approximately equal strength between all modes, $\mathbf{V}^{(k,l)}$ and $\mathbf{U}^{(k,l)}$ are fully random unitary matrices. Further assuming the correlation length is short relative to the system length, so the number of independent sections is large, this corresponds to the strong-coupling regime. In this regime, as shown below, the GD spread and the MDL spread are reduced, and the statistics of MD and MDL approach limiting distributions. Strong mode coupling is highly beneficial in long-haul MDM systems.

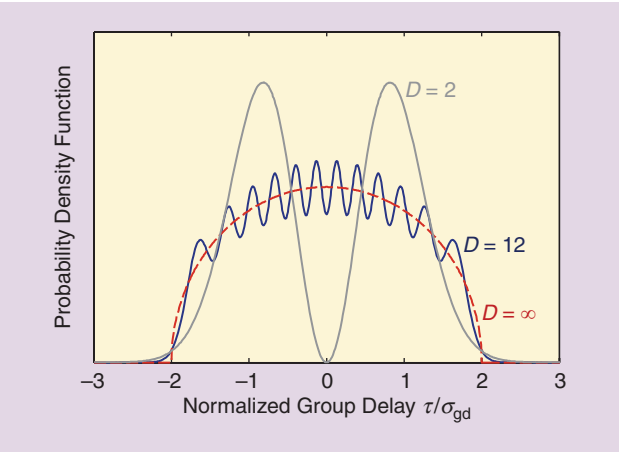
MODAL DISPERSION

Having described the channel model, we discuss the effect of mode coupling on the statistics of the modal GDs, which governs MIMO equalizer complexity. We ignore MDL for simplicity. Given the operator $\mathbf{M}(\Omega)$ describing MD and mode coupling, a GD operator is defined as $\mathbf{G} = j(\partial\mathbf{M}(\Omega)/\partial\Omega)\mathbf{M}(\Omega)^H$ [16], whose eigenvalues are the coupled GDs $(\tau_1^{(t)},\ldots,\tau_D^{(t)})$, ordered as $\tau_1^{(t)} \le \tau_2^{(t)} \le \cdots \le \tau_D^{(t)}$.

Assume that in the absence of coupling, the fiber has peak-to-peak (p-p) and rms GD spread per unit length $\Delta\beta_{1,\text{pp}}$ and $\Delta\beta_{1,\text{rms}}$. Then in each section, the uncoupled GDs τ_i, $i = 1,\ldots,D$ in (5) have p-p spread $\Delta\beta_{1,\text{pp}}L_{\text{sec}}$ and rms spread $\sigma_\tau = \Delta\beta_{1,\text{rms}}L_{\text{sec}}$.

In the weak-coupling regime, the p-p coupled GD spread of the system is $\tau_D^{(t)} - \tau_1^{(t)} = \Delta\beta_{1,\text{pp}}K_{\text{sec}}L_{\text{sec}} = \Delta\beta_{1,\text{pp}}L_{\text{tot}}$, a deterministic quantity that scales linearly with the total number of sections K_{tot} or the total fiber length L_{tot}.

In the strong-coupling regime, which assumes all modes couple equally and $K_{\text{tot}} >> 1$, the coupled GDs $\left(\tau_1^{(t)},\ldots,\tau_D^{(t)}\right)$ are random variables. The rms coupled GD spread of the system is $\sigma_{\text{gd}} = \sqrt{K_{\text{tot}}}\cdot\sigma_\tau = \Delta\beta_{1,\text{rms}}\sqrt{L_{\text{sec}}L_{\text{tot}}}$, which scales with the square root of K_{tot} or L_{tot} [9]. By the central limit theorem, the GD operator becomes a zero-trace Gaussian unitary ensemble. The statistics of the coupled GDs are given by the statistics of the eigenvalues of the ensemble, and depend only on D and σ_{gd} [9]. Figure 3 shows the probability density function (pdf) of the coupled GDs, normalized by σ_{gd}, for several values of D. When $D = 2$, the pdf is Maxwellian, which is known from the study of PMD [17]. For any D, the pdf peaks at D values of GD. In the limit $D \to \infty$, the pdf approaches a semicircle with a finite support $4\sigma_{\text{gd}}$. The presence of MDL is expected to change the GD statistics slightly, as noted for $D = 2$ [18]. The p-p coupled GD spread $\left(\tau_D^{(t)} - \tau_1^{(t)}\right)$ determines the temporal memory required in an equalizer for compensating MD [19], [20], as discussed below.



[FIG3] **Pdfs of normalized coupled GDs for $D = 2$, 12, and ∞ modes.**

MODE-DEPENDENT GAIN AND LOSS

We now discuss the impact of mode coupling on the statistics of MDL, which is the fading distribution governing the channel capacity of MDM systems [8]. Given the operator $\mathbf{M}(\Omega)$ including MDL, MD and mode coupling, an MDL operator is defined as $\mathbf{M}(\Omega)\mathbf{M}(\Omega)^H$. Its eigenvalues are the system's coupled MDL values $(g_1^{(t)}(\Omega),\ldots,g_D^{(t)}(\Omega))$, ordered as $g_1^{(t)}(\Omega) \ge g_2^{(t)}(\Omega) \ge \cdots \ge g_D^{(t)}(\Omega)$, measured in log power gain units. These are the logarithms of the spatial subchannel gains that determine MIMO channel capacity [8].

In the weak-coupling regime, the system's p-p MDL spread $g_1^{(t)}(\Omega) - g_D^{(t)}(\Omega)$ is a deterministic quantity that scales linearly with the number of amplifiers K_{amp}.

In the strong-coupling regime, which assumes all modes couple equally and the number of MDL sources (amplifiers) is large, $K_{\text{amp}} >> 1$, the system's coupled gains $(g_1^{(t)}(\Omega),\ldots,g_D^{(t)}(\Omega))$ are random variables. The statistics of coupled MDL depend only on the number of modes D and on the rms accumulated MDL $\xi = \sqrt{K_{\text{amp}}}\cdot\sigma_g$ [8], which scales with the square root of K_{amp}. Equivalently, these statistics depend on the rms overall MDL $\sigma_{\text{mdl}} = \xi\sqrt{1 + \xi^2/(12(1 - D^{-2}))}$ [18]. In the low-to-moderate MDL range that is of practical interest, the logarithm of the MDL operator is approximately a zero-trace Gaussian unitary ensemble

[8], so the statistics of MDL, normalized by σ_{mdl}, are described by the same distributions as the normalized GDs, shown in Figure 3. Although MDL can potentially make optical amplifier noise spatially nonwhite, this noise becomes spatially white in the strong-coupling regime with a large number of noise sources $K_{\text{amp}} >> 1$ [8].

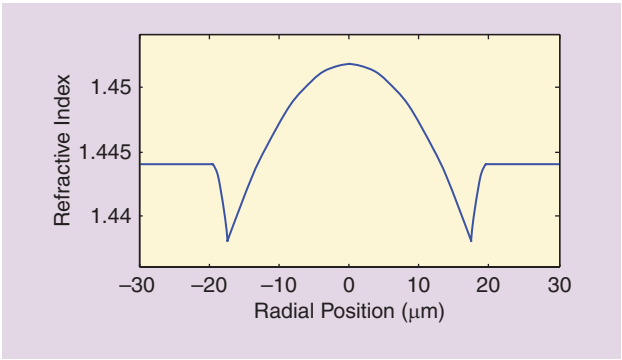
MANAGING AN MDM SYSTEM'S END-TO-END GD SPREAD IS CRUCIAL IN CONTROLLING THE COMPLEXITY AND PERFORMANCE OF MIMO SIGNAL PROCESSING.

In long-haul MDM systems, round-trip propagation delay may be tens of ms, and reliable channel state information (CSI) is not available at the transmitter. At any one frequency Ω, given a realization of the subchannel gains $(g_1^{(t)}(\Omega), \ldots, g_D^{(t)}(\Omega))$, the capacity (per unit frequency) is $C = \sum_{i=1}^{D} \log_2\left(1 + \rho \exp(g_i^{(t)}(\Omega))\right)$, where ρ is the mode-averaged transmitted signal power divided by the mode-averaged received noise power [8]. As the rms accumulated MDL increases, the average capacity decreases, and the variance of capacity increases, increasing the probability of outage [8]. Fortunately, the subchannel gains $(g_1^{(t)}(\Omega), \ldots, g_D^{(t)}(\Omega))$ vary over frequency, with a correlation bandwidth of approximately $1/\sigma_{\text{gd}}$. As long as the symbol rate R_s is high enough that $R_s \sigma_{\text{gd}} >> 1$ (a condition easily satisfied in practice), efficient frequency diversity reduces the outage probability and the outage capacity approaches the average capacity [12].

DELAY SPREAD: MANAGEMENT AND IMPACT

Managing an MDM system's end-to-end GD spread is crucial in controlling the complexity and performance of MIMO signal processing. Here, we review strategies for GD spread management and present an MDM system example to help illustrate how the GD spread affects MIMO signal processing.

Several early MDM experiments, e.g., [21], used fibers supporting two mode groups ($D = 6$), where a low GD spread can be realized by choosing a core radius at which the GD-versus-radius curves for the two mode groups intersect [22]. This approach does not scale easily beyond two mode groups, since the curves for different pairs of modes intersect at different radii. An alternate approach minimizes GD spread by interconnecting different fiber types in which lower- and higher-order modes exhibit an opposite ordering of GDs [23]. However, this approach may be difficult to scale to several mode groups, since specific lengths of several fiber types with specific GD properties may be required. Also, mode coupling may affect the GD compensation obtained. Hence, we adopt the approach described in [22], using fibers with low uncoupled GD spread and relying on strong mode coupling, induced by splices or other perturbations, to further reduce the GD spread [9].



[FIG4] **A fiber design with graded-index core and graded depressed cladding for $D = 12$ modes.**

We consider a family of fibers with graded-index core and graded depressed cladding (GIGDC), which have the desirable properties that the rms uncoupled GD spread $\Delta\beta_{1,\text{rms}}$ decreases with an increasing number of modes D [24]. The index profile for $D = 12$ is shown in Figure 4. Fabrication techniques for these kinds of fibers are discussed in [25] and [26].

To compensate for CD, a digital equalizer must have a duration, measured in samples, given by [27]

$$N_{\text{CD}} = \left\lceil 2\pi |\tilde{\beta}_2| L_{\text{tot}} (r_{\text{os}} R_s)^2 \right\rceil, \tag{6}$$

where r_{os} is the oversampling rate, and $\lceil x \rceil$ denotes the ceiling function. Note that N_{CD} scales linearly with the fiber length L_{tot} and quadratically with the symbol rate R_s.

To compensate for MD and modal crosstalk, a MIMO equalizer must have a duration sufficient to span the p-p coupled GD spread $\left(\tau_D^{(t)} - \tau_1^{(t)}\right)$. Measured in samples, the required duration is $N_{\text{MD}} = \lceil (\tau_D^{(t)} - \tau_1^{(t)}) r_{\text{os}} R_s \rceil$, which scales linearly with R_s. In the weak-coupling regime, the p-p coupled GD spread $\left(\tau_D^{(t)} - \tau_1^{(t)}\right)$ is deterministic, as noted above, and the required equalizer duration, measured in samples, is

$$N_{\text{MD}} = \lceil \Delta\beta_{1,\text{pp}} K L_{\text{sec}} r_{\text{os}} R_s \rceil = \lceil \Delta\beta_{1,\text{pp}} L_{\text{tot}} r_{\text{os}} R_s \rceil, \tag{7}$$

which scales linearly with the number of sections K_{tot} or the fiber length L_{tot}. In the strong-coupling regime, $\left(\tau_D^{(t)} - \tau_1^{(t)}\right)$ is random. As in [19] and [20], we define $u_D(p)$ such that an equalizer duration $\sigma_{\text{gd}} u_D(p)$ is longer than $\left(\tau_D^{(t)} - \tau_1^{(t)}\right)$ with probability $1 - p$. For typical values of D and $p \sim 10^{-4}$ to 10^{-6}, $u_D(p) \sim 4$ to 5, as expected from Figure 3. The required equalizer duration, measured in samples, is

$$N_{\text{MD}} = \left\lceil \sqrt{K_{\text{tot}}}\, \sigma_\tau u_D(p) r_{\text{os}} R_s \right\rceil = \left\lceil \Delta\beta_{1,\text{rms}} \sqrt{L_{\text{sec}} L_{\text{tot}}}\, u_D(p) r_{\text{os}} R_s \right\rceil, \tag{8}$$

which scales with the square-root of K_{tot} or L_{tot}.

To compare the equalizer memory length requirements (6)–(8), we consider a long-haul transmission system with 20 fiber spans, each of length $L_{\text{amp}} = 100$ km, assuming $R_s = 32$ Gbaud, $r_{\text{os}} = 2$ and $p = 10^{-5}$. Figure 5 shows values of N_{CD} and N_{MD} for GIGDC fibers supporting $D = 6$, 12, 20, or 30 modes, as a function of K_{sec} or $L_{\text{sec}} = L_{\text{amp}}/K_{\text{sec}}$, comparing

these to N_{CD} for standard SMF. In the weak-coupling regime, values of N_{MD} for the GIGDC fibers are up to an order of magnitude larger than the values of N_{CD}. In the strong-coupling regime, as K_{sec} increases from 10^0 to 10^3 and L_{sec} decreases from 10^2 to 10^{-1} km, N_{MD} becomes up to two or three orders of magnitude smaller than the values of N_{CD}. As shown in [19] and discussed below, to counter the complexity associated with the $D \times D$ matrix structure of the MIMO equalizer, it is desirable to have N_{MD} at least an order of magnitude smaller than N_{CD}, which requires strong-mode coupling, even with these optimized fibers.

Finally, for comparison purposes, we evaluate the memory length of PMD in standard SMF, denoted by N_{PMD}. Assuming $\Delta\beta_{1,rms}\sqrt{L_{sec}} = 0.1\ \text{ps}/\sqrt{\text{km}}$ and $u_D(p) = 5$, (8) yields $N_{PMD} = 2$, which is far smaller than N_{CD}.

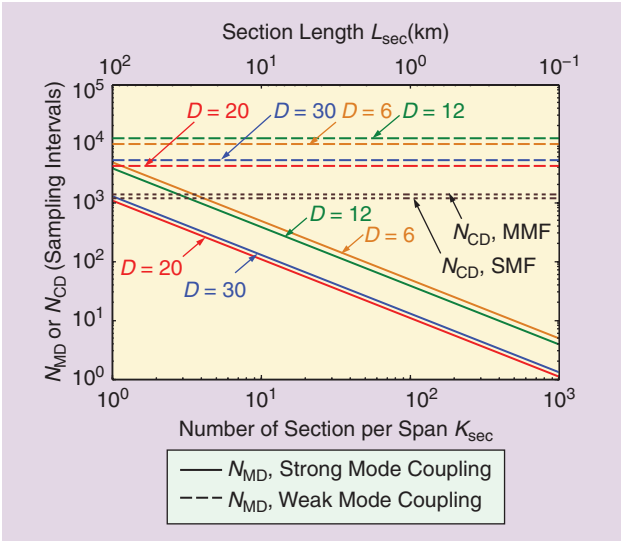
RECEIVER SIGNAL PROCESSING

ARCHITECTURES

Receivers for PDM in SMF ($D = 2$) employ homodyne (or intradyne) downconversion and high-speed analog-to-digital conversion, followed by fully digital processing [14], [27]. Digital equalization compensates for linear channel effects, including CD, PMD, and polarization crosstalk. Other functions implemented digitally may include automatic gain control, timing recovery, carrier recovery, symbol demapping, and error-correction decoding. In receivers for MDM in MMF ($D > 2$), most of these latter functions may be performed as in PDM receivers with straightforward modifications. However, compensation of linear channel effects, especially MD and modal crosstalk, will require different architectures than in PDM receivers to achieve similar adaptation speed and complexity per information bit. The changes from PDM to MDM receivers are necessitated by the increased MIMO dimensionality ($D \times D$ versus 2×2) and because MD in MMF has a delay spread far larger than PMD in SMF ($N_{MD} >> N_{PMD}$).

Several general considerations apply equally to PDM or MDM receivers:

- An equalizer for single-carrier modulation may use either FDE or time-domain equalization (TDE). FDE can handle long delay spreads (large N_{CD} or N_{MD}) with lower computational complexity than TDE by using the fast Fourier transform (FFT) for efficient implementation of convolution.
- CD is nearly fixed, changing very slightly with temperature variations, so it can be compensated by a programmable (but not adaptive) equalizer. PMD (or MD) can change as fast as the 10–100 μs time scale [10], [11] and so requires an adaptive $D \times D$ MIMO equalizer.
- Equalization of CD and MD (PMD in the special case $D = 2$) can be performed together or separately. Figure 6(a) shows a combined $D \times D$ equalizer for CD and MD (or PMD). Figure 6(b) shows a set of D equalizers for CD, followed by a $D \times D$ equalizer for MD (or PMD).

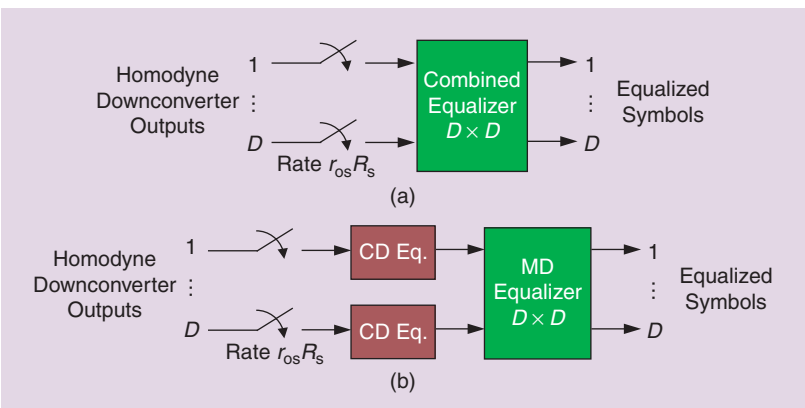


[FIG5] The memory length (in sampling intervals) required to compensate CD or MD, versus number of sections per span K_{sec} or section length L_{sec}, for a long-haul transmission with 20 spans of GIGDC fiber, each 100-km long, with symbol rate $R_s = 32$ Gbaud and oversampling ratio $r_{os} = 2$.

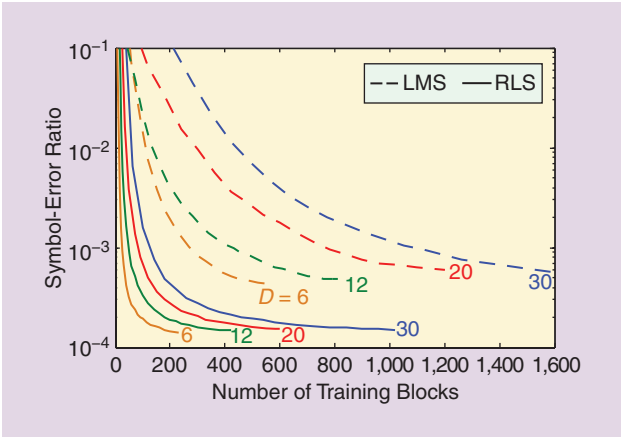
These considerations, along with differences in delay spread and MIMO dimensionality, provide the rationale for choosing different equalizer architectures for PDM and MDM receivers.

Most receivers for PDM in SMF ($D = 2$) use the architecture of Figure 6(b) with a hybrid of FDE and TDE. Since N_{CD} is large, two programmable FDEs compensate CD efficiently. Since N_{PMD} is small, an adaptive 2×2 TDE compensates PMD and polarization crosstalk efficiently and facilitates fast adaptation [28].

In MDM systems, both N_{CD} and N_{MD} are large, so TDE is not viable for either CD or MD, but the optimal FDE architecture is not obvious. Combined equalization of CD and MD using an adaptive $D \times D$ FDE, as in Figure 6(a), has the lowest computational complexity for equalization (not considering adaptation) [19]. But



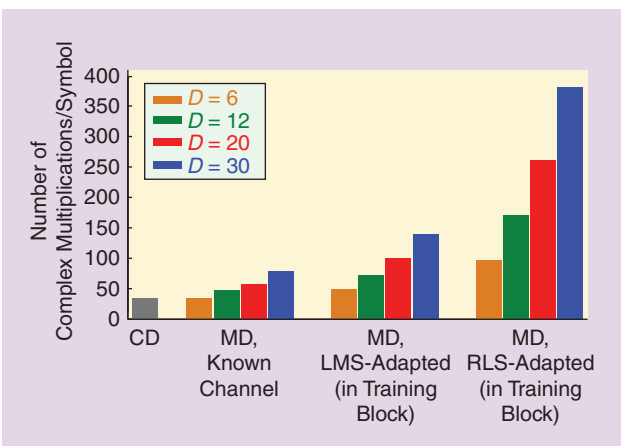
[FIG6] Single-carrier MIMO equalizer structures: (a) a combined equalizer for CD and MD (or PMD) and (b) separate equalizers.



[FIG7] The average symbol-error ratio versus the number of training blocks for $D = 6$, 12, 20, and 30 modes with RLS (solid line) and LMS (dashed line) algorithms (SNR = 10.5 dB, no MDL, $\kappa = 0.999$, and $\mu = 1.5 \cdot 10^{-5}$).

the FFT block length N_{FFT}, which should scale in proportion to $N_{\text{CD}} + N_{\text{MD}}$, becomes long, slowing down adaptation and impeding a single-chip realization. The separate equalizer structure of Figure 6(b) represents a more practical alternative. CD is compensated by a set of D programmable FDEs with FFT block length N_{FFT} proportional to N_{CD}, while MD and modal crosstalk (and any residual CD) are compensated by an adaptive $D \times D$ FDE with FFT block length N_{FFT} dependent on N_{MD} and D. While Figure 6(b) requires a slightly higher complexity than Figure 6(a) for equalization alone [19], minimizing the FFT block length in the adaptive $D \times D$ FDE facilitates fast adaptation and a single-chip realization.

There are two well-known approaches for FFT-based FDE for single-carrier modulation. Block convolution, e.g., overlap-save, avoids any overhead associated with a cyclic prefix, but complicates the realization of an adaptive FDE, requiring additional FFTs to enforce time-domain gradient constraints [29]. An alternate approach, adopted here, is to prepend a cyclic prefix of length N_{CP}



[FIG8] The number of complex multiplications/symbol for various frequency-domain equalization structures for $D = 6$, 12, 20, and 30 modes.

to each block of $N_{\text{FFT}}/r_{\text{os}}$ symbols before transmission. The cyclic prefix length N_{CP}, must be no shorter than the channel delay spread (N_{CD}, N_{MD} or $N_{\text{CD}} + N_{\text{MD}}$). When using a cyclic prefix, linear filtering by the channel corresponds to a multiplicative relationship between the FFTs of input and output signal blocks, which simplifies realization of FFT-based FDE. While cyclic prefix transmission decreases complexity, it reduces system throughput and average-power efficiency, as quantified by an efficiency parameter $\eta_{\text{CP}} = N_{\text{FFT}}/(N_{\text{FFT}} + N_{\text{CP}})$. Given a channel delay spread defining N_{CP}, η_{CP} is maximized by choosing $N_{\text{FFT}} >> N_{\text{CP}}$. Single-carrier modulation with FDE aided by a cyclic prefix or unique word has become popular in wireless systems [30].

ADAPTIVE EQUALIZATION

Assuming the architecture of Figure 6(b), we discuss adaptive frequency-domain implementation of the $D \times D$ equalizer for compensating MD, comparing the LMS and RLS algorithms. The linear MIMO channel $\mathbf{M}(\Omega)$ corresponds to a $D \times D$ matrix $\mathbf{M}[k]$ in the discrete frequency domain. The equalizer is described by a $D \times D$ matrix $\mathbf{W}[k]$. Since a typical channel matrix $\mathbf{M}[k]$ is close to unitary, either LMS or RLS can yield a $\mathbf{W}[k]$ close to the inverse of $\mathbf{M}[k]$. Both LMS and RLS are designed to iteratively minimize the squared error at all discrete frequencies. LMS is a stochastic gradient descent minimization using instantaneous error estimates [29], described by an update equation for the equalizer at frequency k

$$\mathbf{W}[k] \leftarrow \mathbf{W}[k] + \mu(\tilde{\mathbf{x}}[k] - \mathbf{W}[k]\tilde{\mathbf{y}}[k])\tilde{\mathbf{y}}[k]^H, \tag{9}$$

where $\tilde{\mathbf{x}}[k] = [X_1[k] \ \ldots \ X_D[k]]^T$ represents an FFT of sampled known training or detected data blocks, $\tilde{\mathbf{y}}[k] = [Y_1[k] \ \ldots \ Y_D[k]]^T$ represents an FFT of the sampled received signals, and μ is the step size. RLS involves iterative minimization of an exponentially weighted cost function [29], described by update equations

$$\mathbf{W}[k] \leftarrow \mathbf{W}[k] + (\tilde{\mathbf{x}}[k] - \mathbf{W}[k]\tilde{\mathbf{y}}[k])\tilde{\mathbf{y}}[k]^H(\mathbf{R}[k]\kappa^{-1}), \tag{10}$$

$$\mathbf{R}[k] \leftarrow (\mathbf{R}[k]\kappa^{-1}) - \frac{(\mathbf{R}[k]\kappa^{-1})\tilde{\mathbf{y}}[k]\tilde{\mathbf{y}}[k]^H(\mathbf{R}[k]\kappa^{-1})}{1 + \tilde{\mathbf{y}}[k]^H(\mathbf{R}[k]\kappa^{-1})\tilde{\mathbf{y}}[k]}, \tag{11}$$

where $\mathbf{R}[k]$ is an tracked inverse time-averaged weighted correlation matrix [24] and κ is a forgetting factor.

The discrete frequencies for FDE are determined by the symbol rate R_s, oversampling rate r_{os}, and FFT block length N_{FFT}, while the cyclic prefix length must be at least equal to the MD memory length N_{MD} [24]. The choice of N_{FFT} has major consequences on the system, as it affects the adaptation time, cyclic prefix efficiency, and complexity. Given n_{tr}, the number of training blocks required until convergence, the adaptation time is $T_{\text{adapt}} = n_{\text{tr}}(N_{\text{FFT}} + N_{\text{MD}})T_s/r_{\text{os}}$, so fast adaptation favors small N_{FFT}. On the other hand, a high cyclic prefix efficiency η_{CP} favors a large N_{FFT}. Overall computational complexity, including equalization and adaptation, is hardly affected by the choice of N_{FFT} over the range yielding acceptable T_{adapt} and η_{CP} [24].

To study adaptive FDE techniques, we consider the long-haul system described above, using 2,000 km (20 spans, each 100-km long) of GIGDC fiber, assume strong-mode coupling with section length $L_{sec} = 1$ km, neglect MDL, and assume symbol rate $R_s = 32$ Gbaud and oversampling rate $r_{os} = 2$. Initially, with the goal of achieving fast adaptation, we assume $N_{FFT} = 2^{11}, 2^{11}, 2^9, 2^9$ for $D = 6$, 12, 20, and 30, respectively, which yields a fairly low cyclic prefix efficiency $\eta_{CP} \sim 80\%$. Later, we consider increasing N_{FFT} to increase η_{CP} toward 100%. We assume quaternary phase-shift keying and do not include error-correction coding. We quantify signal-to-noise ratio (SNR) by the transmitted signal power per mode divided by the noise variance per mode at the receiver, choosing SNR = 10.5 dB. We present Monte Carlo simulations averaged over random channel and symbol realizations.

Figure 7 illustrates equalizer convergence in terms of symbol-error ratio (SER) versus training block number. Although N_{MD} decreases with increasing D for these GIGDC fibers, the equalizer dimensionality scales as D^2, so the number of training blocks required for convergence increases with increasing D. For RLS, the number of training blocks needed for convergence is roughly proportional to D (the curves nearly overlap when scaled by D^{-1}), and the knee and flat-SER regions of the adaptation curves occur at $\sim 6D$ and $\sim 15D$, respectively. For LMS, the knees occur at $\sim 20D$ to $\sim 30D$, while the flat-SER regime starts at $\sim 50D$ to $\sim 70D$. Choosing the number of training blocks n_{tr} as the onset of the flat-SER region, the estimated adaptation time T_{adapt} is ~ 3–$5\,\mu$s for RLS and ~ 15–$25\,\mu$s for LMS. These should be compared with 25 μs, the experimentally estimated time scale for mode-coupling dynamics [10]. While LMS appears to converge fast enough, it suffers from low cyclic prefix efficiency $\eta_{CP} \sim 80\%$ and higher asymptotic SERs. On the other hand, the convergence of RLS appears to be faster than required, so N_{FFT} can be quadrupled, yielding $\eta_{CP} \sim 95\%$. Moreover, RLS yields lower asymptotic SERs than LMS.

Another important advantage of RLS over LMS is its robustness for ill-conditioned channel matrices caused by MDL. In the presence of MDL, RLS converges reliably, albeit to a higher asymptotic SER than in the absence of MDL. By contrast, LMS must use a reduced step size to ensure convergence, which slows down convergence, and its asymptotic SER is degraded more than that of RLS [24]. Another advantage is that as the SNR increases, the asymptotic SER for RLS decreases more rapidly than that for LMS [24].

The major drawback of RLS is the high adaptation complexity per step arising from the matrix operations in (10) and (11). Figure 8 compares computational complexities, in terms of complex multiplications per data symbol in a training block, for FDEs adapted using LMS or RLS. This complexity scales sublinearly with D for both algorithms, and is 2.0 to 2.7 times higher for RLS than LMS for a given value of D. However, RLS converges to a low SER using ~ 4–5 times fewer training blocks, so the overall complexity for initial adaptation to an unknown channel is lower for RLS than for LMS. For reference, Figure 8 also shows complexities for equalizing CD (using overlap-save [19]) and MD for a known channel (using a cyclic prefix [24]), assuming optimized block lengths.

DISCUSSION

We have considered adaptation to an initially unknown channel, which is readily modeled without making assumptions about channel dynamics. Continuously tracking a channel in decision-directed mode is critically important. Our simulations indicate that continuous RLS adaptation would be effective, but at the cost of potentially prohibitive complexity and power consumption. Alternatives for tracking include RLS adaptation at periodic intervals or continuous LMS adaptation. Evaluating tracking performance requires models for channel dynamics, such as one proposed for fibers with two mode groups ($D = 6$) [31]. Such models should be extended to larger D and validated experimentally.

CONTINUOUS RLS ADAPTATION WOULD BE EFFECTIVE, BUT AT THE COST OF POTENTIALLY PROHIBITIVE COMPLEXITY AND POWER CONSUMPTION.

GD spread is a key factor determining the complexity, adaptation speed and cyclic prefix efficiency of MIMO equalizers for MDM systems. To manage GD spread, we have proposed fibers with low uncoupled GD spread in conjunction with strong mode coupling described by a section length $L_{sec} = 1$ km. As noted in [22], [26], manufacturing process variations may increase the uncoupled GD spread beyond its ideal value, while splices between fiber sections may be expected to lead to section lengths L_{sec} of order 5 km. Obtaining the low coupled GD spread assumed here may require intentionally perturbing the fiber in some way analogous to the "spinning" used to reduce PMD in SMF [32]. Fiber designs for reduced GD spread and methods to enhance mode coupling without increasing loss and MDL are important topics for future study.

Given the present uncertainty about mode-coupling dynamics and coupled GD spreads achievable in long-haul fibers, the MIMO FDE design parameters and performance and efficiency metrics provided here should be considered illustrative examples more than precise determinations.

In this article, we have focused on MIMO equalization, which is one of several important signal processing functions in MDM systems. Optimization of other functions, such as carrier recovery, and the integration and hardware implementation of the various functions, are important topics for future works.

ACKNOWLEDGMENTS

We are grateful to the guest editors for inviting us to write this tutorial article and to the anonymous reviewers for their helpful suggestions. This research was supported by National Science

Foundation grant number ECCS-1101905, Corning, Inc., and a Stanford Graduate Fellowship.

AUTHORS

Sercan Ö. Arık (soarik@stanford.edu) received the B.S. degree in 2011 from Bilkent University (Ankara, Turkey) and the M.S. degree in 2013 from Stanford University in electrical engineering. He is currently working toward his Ph.D. degree at Stanford University. He worked at EPFL Signal Processing Labs (Lausanne, Switzerland) in 2010, Google (Mountain View, California) in 2012, and Mitsubishi Electric Research Labs (Cambridge, Massachusetts) in 2013. His main current research interests include multimode optical communications; advanced modulation, coding, and digital signal processing techniques for fiber-optic systems; and optical signal processing. He is a Student Member of the IEEE.

Joseph M. Kahn (jmk@ee.stanford.edu) is a professor of electrical engineering at Stanford University. His research addresses communication and imaging through optical fibers, including modulation, detection, signal processing, and spatial multiplexing. He received A.B. and Ph.D. degrees in physics from the University of California, Berkeley, in 1981 and 1986. From 1987 to 1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, in Holmdel, New Jersey. He was on the faculty of the Electrical Engineering Department at the University of California, Berkeley, from 1990 to 2003. In 2000, he cofounded StrataLight Communications, which was acquired in 2009 by Opnext, Inc. He received the National Science Foundation Presidential Young Investigator Award in 1991 and is a Fellow of the IEEE.

Keang-Po Ho (kpho@ieee.org) received his Ph.D. degree from the University of California, Berkeley, in 1995. He is currently with Silicon Image as a senior manager for the Baseband Algorithm Group. In Silicon Image (SiBEAM before acquisition), he invented and implemented wireless HD video area network using 60-GHz antenna array, 3.8-Gb/s OFDM modem, and beam search. He was the chief technology officer and cofounder of StrataLight Communications. He had also been with IBM, Bellcore, CUHK, and National Taiwan University. He has authored over 200 journal and conference articles, 15 issued patents, and one book. He is a Senior Member of the IEEE.

REFERENCES

[1] R. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Tech. J.*, vol. 14, no. 4, pp. 3–9, 2010.

[2] E. Ip, P. Ji, E. Mateo, Y. Huang, L. Xu, D. Qian, N. Bai, and T. Wang, "100G and beyond transmission technologies for evolving optical networks and relevant physical-layer issues," *Proc. IEEE*, vol. 100, no. 5, pp. 1065–1078, May 2012.

[3] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightwave Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[4] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *J. Lightwave Technol.*, vol. 28, no. 6, pp. 939–951, Mar. 2010.

[5] S. Randel, R. Ryf, C. Schmidt, M. A. Mestre, R.-J. Essiambre, and P. J. Winzer, "MIMO processing for space-division multiplexed transmission," in *Proc. Advanced Photonics Congr.*, 2012, paper SpW3B.4.

[6] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[7] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Natl. Acad. Sci.*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000.

[8] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Opt. Express*, vol. 19, no. 17, pp. 16612–16635, Aug. 2011.

[9] K.-P. Ho and J. M. Kahn, "Mode coupling and its impact on spatially multiplexed systems," in *Optical Fiber Telecommunications VI*, I. P. Kaminow, T. Li, and A. E. Willner, Eds. Amsterdam, The Netherlands: Elsevier, pp. 491–568, 2013.

[10] X. Chen, J. He, A. Li, J. Ye, and W. Shieh, "Characterization of dynamic evolution of channel matrix in two-mode fibers." in *Proc. Optical Fiber Communication Conf. and Expo./Nat. Fiber Optic Engineers Conf. (OFC/NFOEC)*, Anaheim, CA, Mar. 2013, paper OM2C.3.

[11] P. Krummrich and K. Kotten, "Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems," in *Proc. Optical Fibers Communication Conf.*, 2004, pp. 1–3.

[12] K.-P. Ho and J. M. Kahn, "Frequency diversity in mode-division multiplexing systems," *J. Lightwave Technol.*, vol. 29, no. 24, pp. 3719–3726, Dec. 2011.

[13] W. Shieh, H. Bao, and Y. Tang, "Coherent optical OFDM: Theory and design," *Opt. Express*, vol. 16, no. 2, pp. 841–859, Jan. 2008.

[14] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Express*, vol. 16, no. 2, pp. 753–791, Jan. 2008.

[15] E. Ip and J. M. Kahn, "Fiber impairment compensation using coherent detection and digital signal processing," *J. Lightwave Technol.*, vol. 28, no. 4, pp. 502–519, Feb. 2010.

[16] S. Fan and J. M. Kahn, "Principal modes in multimode waveguides," *Opt. Lett.*, vol. 30, no. 2, pp. 135–137, Jan. 2005.

[17] G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single-mode fibers," *J. Lightwave Technol.*, vol. 9, no. 11, pp. 1439–1456, Nov. 1991.

[18] R. Feced, S. J. Savory, and A. Hadjifotiou, "Interaction between polarization mode dispersion and polarization dependent losses in optical communication links," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 20, no. 3, pp. 424–433, Mar. 2003.

[19] S. O. Arik, D. Askarov, and J. M. Kahn, "Effect of mode coupling on signal processing complexity in mode-division multiplexing," *J. Lightwave Technol.*, vol. 31, no. 3, pp. 423–431, Feb. 2013.

[20] K.-P. Ho and J. M. Kahn, "Delay-spread distribution for multimode fiber with strong mode coupling," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1906–1909, Nov. 2012.

[21] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R.-J. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6×6 MIMO processing," *J. Lightwave Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 2012.

[22] D. Peckham, Y. Sun, A. McCurdy, and R. Lingle, "Few-mode fiber technology for spatial multiplexing," in *Optical Fiber Telecommunications VI A*, I. P. Kaminow, T. Li, and A. E. Willner, Eds. Amsterdam, The Netherlands: Elsevier, pp. 283–319, 2013.

[23] R. Ryf, M. A. Mestre, A. H. Gnauck, S. Randel, C. Schmidt, R.-J. Essiambre, P. J. Winzer, R. Delbeu, P. Pupalaikis, A. Sureka, Y. Sun, X. Jiang, D. W. Peckham. A. McCurdy, and R. Lingle, "Low-loss mode coupler for mode-multiplexed transmission in few-mode fiber," in *Proc. Optical Fiber Communication Conf.*, Los Angeles, CA, Mar. 2012, paper PDP5B.5.

[24] S. O. Arik, D. Askarov, and J. M. Kahn, "Adaptive frequency-domain equalization in mode-division multiplexing systems," submitted for publication.

[25] D. Đonlagic, "Opportunities to enhance multimode fiber links by application of overfilled launch," *J. Lightwave Technol.*, vol. 23, no. 11, pp. 3526–3540, Nov. 2005.

[26] L. G. Nielsen, Y. Sun, J. W. Nicholson, D. Jakobsen, K. G. Jespersen, R. Lingle, and B. Palsdottir, "Few mode transmission fiber with low DGD, low mode coupling, and low loss." *J. Lightwave Technol*, vol. 30, no. 23, pp. 3693–3698, Dec. 2012.

[27] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE Select. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Oct. 2010.

[28] B. Spinnler, "Equalizer design and complexity for digital coherent receivers," *IEEE J. Quantum Electron.*, vol. 16, no. 5, pp. 1180–1192, Oct. 2010.

[29] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[30] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.

[31] N. Bai and G. Li, "Adaptive frequency-domain equalization for mode-division multiplexed transmission," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1918–1921, Nov. 2012.

[32] D. A. Nolan and M. J. Li, "Fiber spin-profile designs for producing fibers with low polarization mode dispersion," *Opt. Lett.*, vol. 23, no. 21, pp. 1659–1661, Nov. 1998.

[SP]

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

Xiang Zhou

Efficient Clock and Carrier Recovery Algorithms for Single-Carrier Coherent Optical Systems

A systematic review on challenges and recent progress

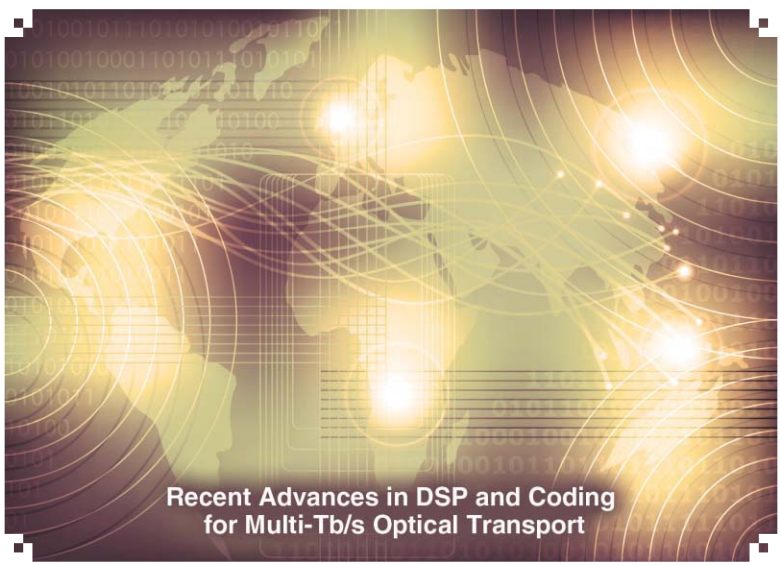


IMAGE LICENSED BY INGRAM PUBLISHING /JANMIKS

This article presents a systematic review on the challenges and recent progress of timing and carrier synchronization techniques for high-speed optical transmission systems using single-carrier-based coherent optical modulation formats.

INTRODUCTION

To increase the capacity of optical networks in a more cost-effective way, spectrally efficient signal modulation techniques, such as quadrature phase-shift keying (QPSK), have recently been introduced for 100 Gb/s-class optical transmission. To further increase the transport spectral efficiency (SE) for future 400 Gb/s and beyond optical transmission systems, a higher-order optical modulation format, such as M-ary quadrature amplitude modulation (M-QAM) (16-QAM, 32-QAM, 64-QAM, etc.) has also received significant attentions in recent years.

Optimal reception of these spectrally efficient and multidimensional modulation formats, however, requires the receiver to be able to access to the full field information of the transmitted optical signal, which mandates the use of coherent detection, where the received optical signal is mixed with a narrow-line width CW

Digital Object Identifier 10.1109/MSP.2013.2281071
Date of publication: 12 February 2014

local oscillator (LO) before photodetection. In such a coherent system, the frequency and phase offset between the transmitter laser and the LO, which usually change over time, have to be estimated and removed at the receiver side (i.e., the carrier recovery) to correctly demodulate the received optical signal.

In the early coherent optical communication experiments (1980s to early 1990s), optical phase-locked loop (PLL)-based methods were used to perform carrier recovery. But this type of optical method has proven to be too complex for practical implementation. With the recent advancement of a high-speed electrical processing technique, digital signal processing (DSP)-based methods, traditionally only used for lower speed wireless/baseband systems, have been introduced for high-speed optical systems [1] and quickly established themselves as the most promising solution to address this challenge and the well-known polarization mode dispersion (PMD) problem [2]. Such a digital coherent detection technique has enjoyed a great success in enabling spectrally efficient 100 Gb/s transmission using polarization-multiplexed (PM) QPSK.

Digital carrier recovery in a high-speed optical system is more challenging than a lower-speed wireless/baseband system, especially for using higher-order modulation formats, because

- unlike the wireless system in which the frequency and phase changes (compared to the symbol rate) are relatively similar and slow, the characteristics of frequency and phase offsets in a high-speed optical system are very different: frequency change is relatively slow but the range can be very large (may be up to 5 GHz) while the phase noise typically varies at a much higher speed (relative to the symbol rate) than a wireless system
- high-speed optical systems demand a high-degree of parallel processing that make feedback-based digital carrier phase recovery algorithms (widely used in wireless/baseband communication systems) less effective for high-speed optical systems
- high-speed optical systems are much more constrained on the complementary metal–oxide–semiconductor capability and therefore the requirement on the hardware efficiency is higher.

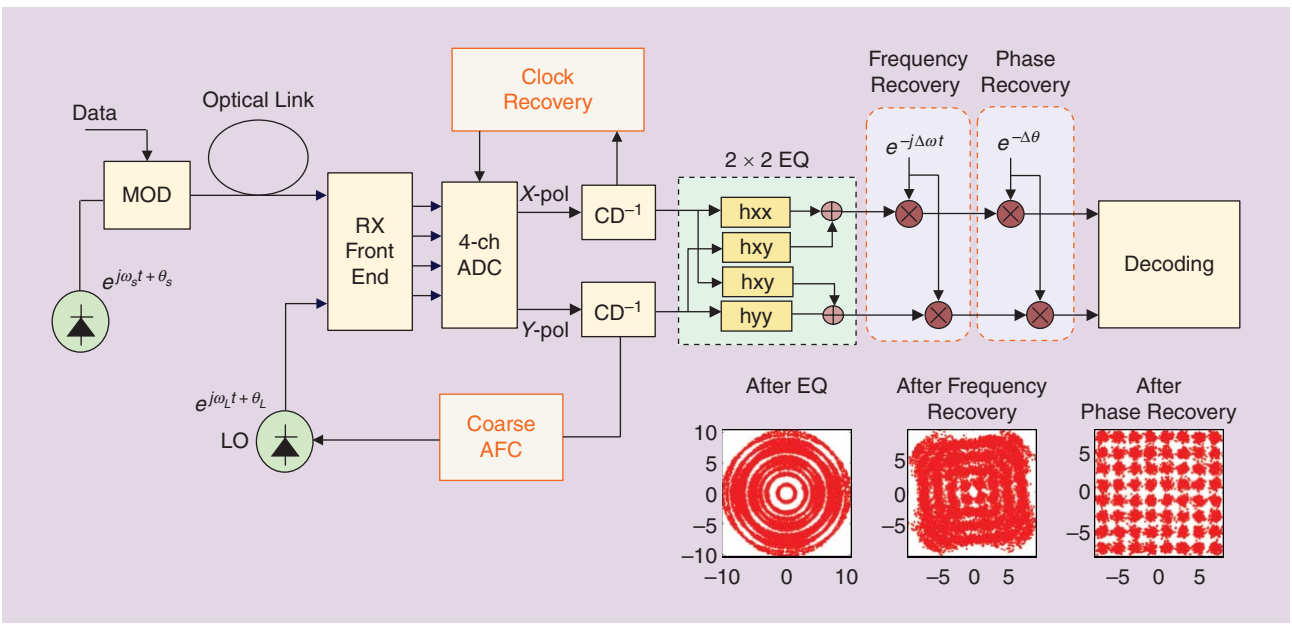
OVER THE PAST FEW YEARS, SIGNIFICANT PROGRESS HAS BEEN MADE TOWARD DEVELOPING IMPAIRMENT-TOLERANT CLOCK RECOVERY AND HARDWARE-EFFICIENT CARRIER RECOVERY ALGORITHMS.

In addition to carrier recovery, accurate timing synchronization is also vital for a coherent receiver. Because fiber chromatic dispersion (CD) and PMD will cause significant distortion to the signal, how to recover the clock in the presence of significant channel distortion is another challenging issue in high-speed optical transmission systems. The recent introduction of Nyquist pulse shaping further complicates the design of clock recovery circuits.

Over the past few years, significant progress has been made toward developing impairment-tolerant clock recovery and hardware-efficient carrier recovery algorithms. This article intends to present a systematic review on these algorithms as well as the remaining challenges.

CONCEPT OF DIGITAL CLOCK AND CARRIER RECOVERY

The basic concept of digital clock and carrier recovery for a high-speed coherent optical transmission system is illustrated in Figure 1, where a square 64-QAM is used as an example. In



[FIG1] A schematic illustration of digital clock and carrier recovery for a high-speed optical system using a square 64-QAM modulation format as an example. ADC: analog to digital converter; MOD: modulator; RX: receiver; CD: chromatic dispersion; EQ: equalizer.

Figure 1, the CD compensator is basically a static (or slowly adaptable) equalizer (one for each polarization) typically implemented in the frequency domain while the 2×2 butterfly-configured equalizer is a fast adaptive equalizer usually implemented in the time domain for fast polarization tracking and PMD compensation. The digital clock recovery is typically performed prior to the carrier recovery: it can be performed right after CD compensation (as shown in Figure 1) but it may also be implemented after the 2×2 equalizer. The recovered clock can be directly used as the clock of the analog-to-digital converter (ADC) to achieve synchronous sampling, but it may also be used to resample the CD-compensated data samples while a free-running clock is used for the ADC (i.e., asynchronous sampling). Note that the finite-impulse-response (FIR) filter-based adaptive equalizer also has the capability to track the timing phase error, but its capability to track the sampling clock frequency-offset is limited unless very sophisticated algorithms are used to decouple the impact of sampling clock offset and PMD and dynamically adjust the filter weight distribution [3]. So it is preferable to have an independent clock recovery unit.

Due to the different characteristics of carrier frequency and phase offset, an independent carrier frequency recovery circuit is usually needed before the phase recovery. Because the frequency of a free running laser can drift up to ± 2.5 GHz end of its life, and without controlling the LO frequency, up to 5 GHz extra electrical bandwidth is needed to preserve all the needed signal components for the following processing. To avoid such an increased receiver bandwidth requirement, a coarse automatic frequency control (AFC) circuit maybe introduced prior to the nominal frequency recovery unit. The use of a coarse AFC could reduce the frequency offset (FO) from several GHz to be within several hundred megahertz range. Coarse frequency control at a resolution >100 MHz can be easily realized by using some classic frequency domain-based methods, for example, by simply monitoring the spectrum asymmetry of the digitized signal during the frequency-domain-based dispersion compensation process. It can also be realized by using a time-domain-based inphase/quadrature I/Q correlation method using pre-equalized signal as was reported in [4].

The nominal (fine) frequency and phase recovery are realized after AFC and equalization. The frequency recovery unit estimates the residual FO ($\Delta\omega$) by using the equalized signals to an accuracy in the MHz range, and then removes this offset from the signal. The phase recovery unit then estimates the combined phase noise ($\Delta\theta$) from the LO and the signal source and then removes it from the signal. The phase-recovered signal is then sent to the decoding unit.

IMPAIRMENT-TOLERABLE CLOCK RECOVERY

Two types of clock recovery methods have been proposed for high-speed coherent optical transmission systems. The first one is a feedback-based algorithm where the timing phase error signal is detected from sampled signals, which are then used to control voltage-controlled oscillator or to drive a digital interpolator [5]–[9]. The second one is a feedforward-based method, where the timing offset is estimated from blocks of samples, and then adjusted by using digital interpolation [10], [11]. The two methods may be used together to improve channel distortion tolerance while maintaining good jitter tolerance but at the expense of increased implementation complexity [12].

FEEDBACK-BASED CLOCK RECOVERY

A hardware-efficient feedback clock recovery scheme capable of tolerating large CD and PMD was first proposed in [5] by introducing a new frequency-domain-based timing phase error detector (TEPD), which is derived from the classic discrete-time Gardner's phase detector [6]. For a complex input sequence $x(n)$ sampled at two samples per symbol, the classic Gardner's phase error detector is given by

$$\tau_{\text{err}} = \sum_{n=0}^{N/2-1} [x(2n-1) - x(2n+1)] \cdot x^*(2n), \tag{1}$$

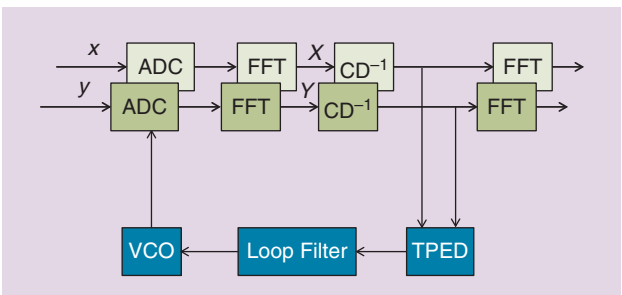
where $*$ denotes complex conjugation and N denotes the sample numbers used for each timing error estimation. The frequency domain expression of (1) can be given by

$$\tau_{\text{err}} = \sum_{k=0}^{N/2-1} \text{Im}\{X(k) \cdot X^*(k+N/2)\}, \tag{2}$$

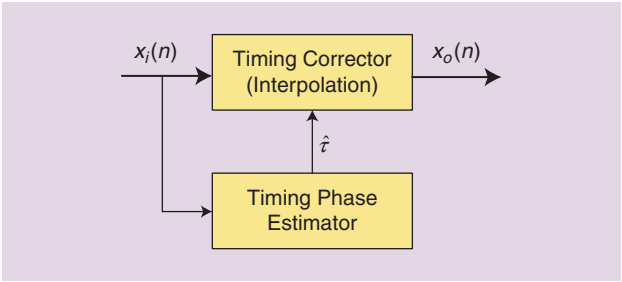
where $X(k)$ is the fast Fourier transform (FFT) of $x(n)$. As can be seen from Figure 2, using a frequency-domain phase error detector enables us to detect the timing phase error using CD-compensated frequency-domain signals. To be able to tolerate large PMD, oversampled signals from both X- and Y-polarizations have to be considered. In accordance with [5], a modified timing phase error can be given by

$$\tau_{\text{err}} = \sum_{k=0}^{N/2-1} \text{Im}\left\{[X(k) + Y(k) \cdot e^{j\Phi_u}] \cdot \left[X^*\left(k+\frac{N}{2}\right) + Y\left(k+\frac{N}{2}\right) \cdot e^{j\Phi_L}\right]\right\}, \tag{3}$$

where a linear combination of received signals from both polarizations are included. Note that two additional rotation phase angles, Φ_u and Φ_L, are introduced to the positive and negative frequency components, respectively. It has been shown that, by choosing appropriate values of Φ_u and Φ_L, the weakened clock tone due to first-order PMD effects can be partly restored (because changing the relative phase between the X and Y signals in the frequency domain is equivalent to adjusting the relative delay in the time



[FIG2] An illustration of a CD and PMD-tolerant frequency-domain clock extraction method [10]. TPED: timing phase error detector.



[FIG3] **Feedforward clock recovery functional blocks.**

domain). In the real world in which the PMD changes over time, the optimal value of Φ_u and Φ_L can be tracked by using an extra gradient-based feedback loop.

For normal return-to-zero (RZ) or nonreturn-to-zero (NRZ) pulse-shaped signals, a sampling rate at two samples per symbol is usually good enough for a high-fidelity clock extraction (assume with optimal receiver bandwidth). For a raised-cosine-pulse-shaped signal with a low roll-off factor (i.e., the Nyquist-shaped signal), however, the phase detector presented in (3) could fail because the clock tone for a Nyquist-shaped signal is very weak (and completely disappears for an ideal Nyquist-shaped signal) [7]. Recently, a modified Gardner phase detector algorithm has been proposed to address this issue [8], [9]. For this method, the phase error detector is given by

$$\tau_{\text{err}} = \sum_{n=1}^{N/2} [P(2n+1) - P(2n-1)] \cdot P(2n), \tag{4}$$

where P is the power of the received signal (assume two samples per symbols). It has been shown that such a modified timing error detector greatly improves the performance for a Nyquist-shaped signal. Alternatively, the standard deviation of the amplitude of the received signal can also be used as a measure of timing phase error [7] for Nyquist-shaped signals. For these modified algorithms aiming at Nyquist-shaped signals, more work needs to be done to understand their tolerance toward channel distortion.

FEEDFORWARD-BASED CLOCK RECOVERY

The feedforward clock recovery diagram is shown in Figure 3. Unlike the feedback scheme requiring the generation of timing error signal for each symbol, the feedforward schemes operate on blocks of samples. The well-known square timing recovery [10] is the most widely used feedforward algorithm that directly computes the phase of the clock line of the signal power spectrum. The estimated timing offset is given by

$$\tau(nL_0) = -\frac{T}{2\pi}\arg\left\{\sum_{n=0}^{ML_0-1} |x_i(nT)|^2 e^{-j2\pi n/M}\right\}, \tag{5}$$

where M is the oversampling number, L_0 denotes the number of symbols used for timing offset estimation, and T is the symbol duration. Because square nonlinear operation will double the signal bandwidth, four samples per symbol is typically required. To reduce the required sampling rate, additional modulation and filtering techniques have to be used to downshift the signal spectral components that carry the clock information (i.e., the roll-off passband portion) to the baseband [11]. Furthermore, a feedforward algorithm requires a digital timing corrector (i.e., interpolation) operating on a symbol by symbol basis. So a feedforward-based method is inherently less hardware efficient than feedback-based algorithms. But a feedforward-based method can achieve much faster clock recovery, which may be useful for future burst-switched systems.

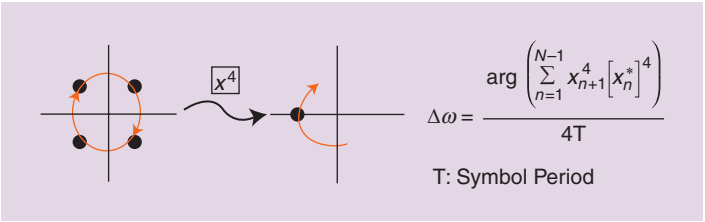
In [12], a new channel distortion tolerable clock recovery method has been proposed, where the timing signal is extracted after the 2×2 adaptive equalizer (i.e., after the correction of channel distortions) by using the combination of feedback- and feedforward-based methods. Due to the extended feedback delay, purely feedback-based methods are susceptible to high-frequency timing jitter. So a feedforward path is introduced to improve the overall jitter tolerance. Although this method is very robust against channel distortion, the implementation complexity is quite high.

In summary, extracting the clock prior to the 2×2 adaptive equalizer using Gardner or modified Gardner phase detectors through a feedback loop can be easier to implement than a feedforward-based timing recovery circuit. But more work needs to be done to understand their channel distortion tolerance performance. Extracting the clock after the 2×2 adaptive equalizer by using joint feedback and feedforward methods greatly improves the channel distortion tolerance, but the implementation complexity can be high.

HARDWARE-EFFICIENT FREQUENCY RECOVERY

For high-speed optical communication systems, blind frequency recovery is preferable because sending a training signal will decrease the achievable transport SE. For blind frequency recovery, the challenging part is how to perform robust and accurately FO estimation for arbitrary modulation formats. Over the past few years, much effort has been made to develop efficient FO estimation algorithms that work for parallel high-speed optical systems. These algorithms can be largely divided into three categories. The first types of methods use the well-known Mth power algorithm [13] to remove the data modulation and then extract FO through either time-domain-based differential phase method or frequency domain-based FFT method. This type of algorithm works well for M-PSK modulated signals but is much less effective for high-order QAM-modulated systems. The second category is a blind frequency search (BFS) algorithm, where the carrier FO is scanned over certain frequency region and the optimal carrier FO is determined by using a minimum phase/Euclidean distance method [14]. This method can be used for any modulation formats and the required number of symbols for reliable FO estimate is much less than the Mth-power-based method. The third method uses a starting training sequence to get an initial FO estimate and then use the recovered phase angle from the following phase recovery unit to track the FO change in a blind way [15]. The third method has the lowest implementation complexity and also works for arbitrary modulation formats. Moreover, the tolerable frequency-offset can be up to half of the baud rate. It should be noted that, in principle simultaneous FO estimate and phase

recovery can be achieved by using a single-stage second-order PLL [16] or some of its variants [17]. But this type of methods generally performs poor for high-speed optical transmission systems requiring high-degree of parallel processing (assume using common symbol-wise time-interleaved parallel processing architecture) because the correct estimation of the carrier FO relies on correct phase estimation, which is very sensitive to the feedback delay.



[FIG4] **An illustration of the fourth-power carrier FO estimation using a time-domain-based differential phase (or phase increment) method.**

Mth-POWER-BASED BLIND FREQUENCY OFFSET ESTIMATION METHOD

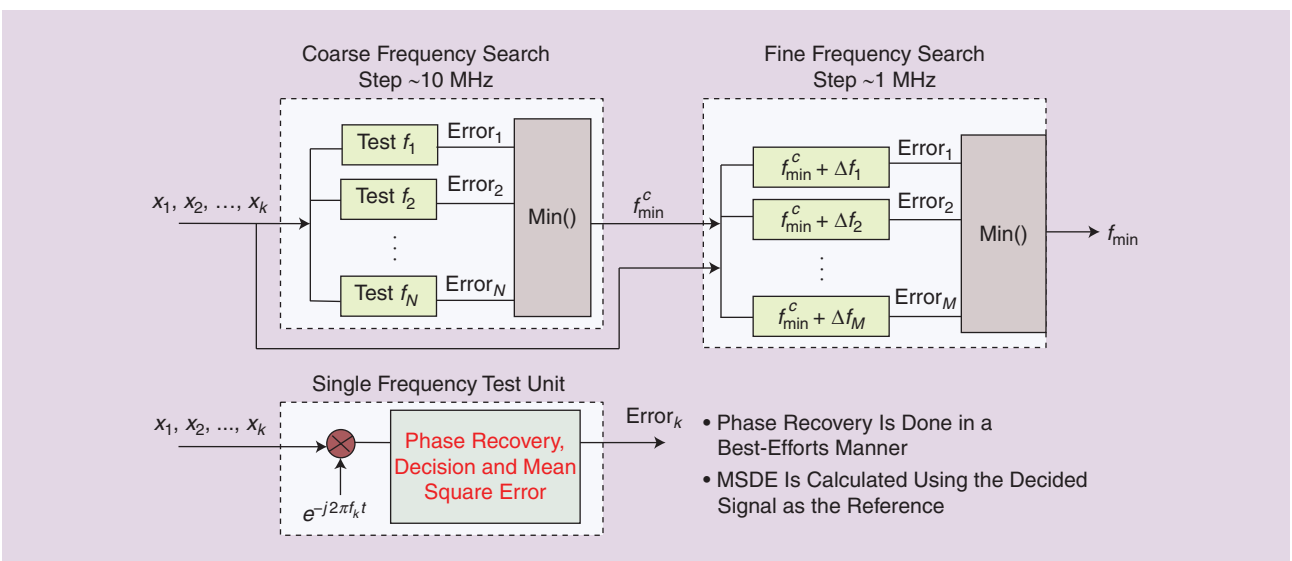
Figure 4 shows a schematic illustration of a time-domain differential phase-based FO estimation method, where the well-known Mth-power algorithm [13] is used to remove the data modulation. Here QPSK and the fourth power algorithm are used as an example. This time-domain differential phase algorithm [18] can be easily extended to higher-order MPSK signals. But the performance for high-order QAMs is poor because, for a higher-order QAM, only a small portion of the constellation points having equal phase spacing is usable for FO extraction. Even for an MPSK-modulated signal, such a time-domain-based FO estimation method can become unstable when operating at a low optical-to-signal noise ratio (OSNR) region (with relatively large phase noise).

The FFT-based method can also be used to extract the FO from the data-erased signals, because the phase angle of a data-erased signal will exhibit a FFT peak at M times the frequency offset. FFT-based methods can achieve much more reliable FO estimation than the time-domain-based method (by using the same number of symbols) but the implementation complexity is much higher, especially for higher-order QAMs, where tens of thousands of symbols may have to be used for a reliable and accurate FO estimate [19]. Furthermore, a single FFT operation can only determine the magnitude of the FO. To get the sign of the FO, an additional FFT may have to be used. Some efforts have been made to simplify FFT-based methods by using, for example, constellation-partitioning-based methods [20]. For a high-order square QAM, it has been shown that [20], by using only the outmost four constellation points combined with the use of linear interpolation and downsampling-based methods, the implementation complexity can be greatly reduced. But the overall complexity is still high as compared to QPSK signals.

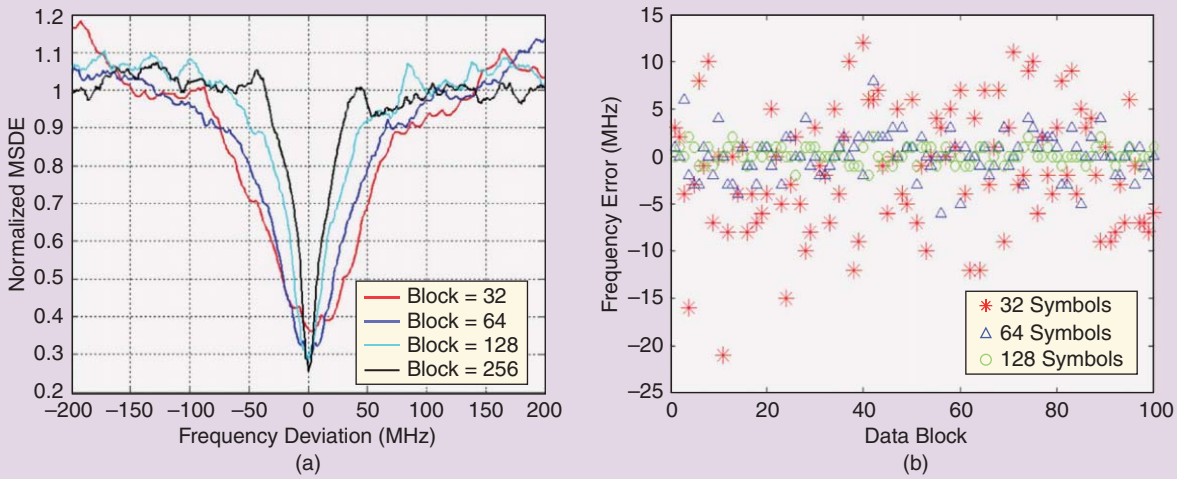
BLIND FREQUENCY SEARCH ALGORITHM

The BFS algorithm was first proposed and demonstrated in [14] as a universal carrier FO estimation method, where minimum mean square distance error (MSDE, in terms of phase or Euclidean distance) is used as the frequency selection criteria. For this method, the FO is first scanned at a coarse step size (~10 MHz) and then at a fine step size of (~1 MHz), and the optimal FO is the one that gives the minimum MSDE (see Figure 5). For each trial frequency, the carrier phase is first recovered (with best efforts) by using blind phase estimation algorithms, and decisions made following this phase estimation are then approximated as the reference signals for MSDE calculation.

Figure 6(a) and (b) show the simulated results for a 38 Gbaud 64-QAM system with laser line width = 100 kHz and OSNR = 25 dB. Figure 6(a) shows how the normalized MSDE varies with the frequency deviation (the difference between the trial FO and the actual FO) and the number of symbols used for FO estimation, while Figure 6(b) shows the FO estimate error



[FIG5] **The operation principle for the proposed blind frequency search algorithm.**

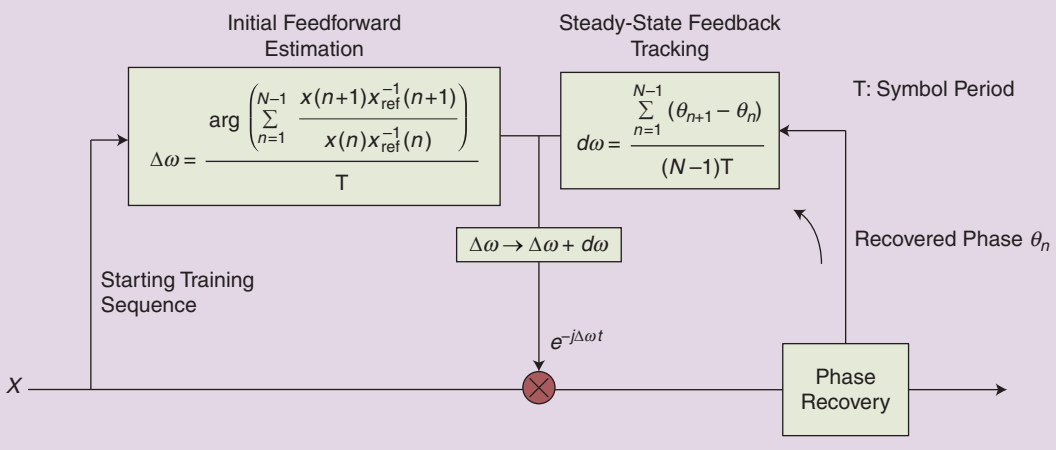


[FIG6] Simulated results for a 38-Gbaud 64-QAM system with laser line width = 100 kHz and OSNR = 25 dB by using the proposed BFS method. (a) The impact of frequency deviation and block lengths on MSDE and (b) The FO estimate error distribution by using different block lengths for the FO estimation.

distributions by using several different data block lengths. One can see that, BFS can reliable estimate FO to be within 20 MHz by using only 32 symbols. Increasing the number of symbols to 128 can improve the FO estimation accuracy to be within 3 MHz, which is good enough for most of the applications even for a modulation order as high as 64 QAM. Because BFS requires a much smaller number of symbols for FO estimation (as compared to the previous Mth-power-based differential phase or FFT-based methods), it can be used to achieve very fast carrier frequency recovery if it is implemented with a parallel processing architecture as is shown in Figure 5. For a typical coherent receiver where the carrier frequency varies much more slowly than the symbol rate, BFS may also be implemented with a sequence or partial sequence processing architecture (such as the central processing unit in a computer) to reduce the implementation complexity.

TRAINING-INITIATED FEEDBACK FREQUENCY OFFSET TRACKING METHOD

In [15], a much simpler FO estimation method that also works for arbitrary modulation format has been proposed and demonstrated. For this method, the initial signal-LO FO is estimated by using a starting training sequence in a feedforward manner and then the FO change is tracked through a feedback configuration using the recovered carrier phases from the following phase recovery stage (typically using a feedforward configuration) as is shown in Figure 7. Note that, unlike the fast-changing phase noise that cannot tolerate extended feedback delay in high-speed optical systems requiring a high-degree of parallel processing, carrier frequency typically varies much more slowly and thus it can be tracked by using a feedback-based architecture. The advantages of this new method are that 1) it is applicable to arbitrary QAM, 2) its implementation complexity is very low because it



[FIG7] A schematic illustration of the proposed training-initiated frequency recovery algorithm. T: symbol period, ($\Delta\omega$): estimated carrier frequency offset.

requires significantly fewer complex multiplications than previous described algorithms, and 3) the tolerable frequency offset can be very large (up to half of the symbol rate). As compared to the frequency recovery method based on second-order PLL, where both the frequency and phase recovery relies on a feedback mechanism, here only the FO tracking uses a feedback configuration while the phase recovery is achieved in a feedforward manner, as will be discussed in the following section.

In summary, the Mth-power-based time-domain differential phase method (a feedforward method) is a hardware-efficient blind FO estimation method for MPSK-modulated signals operating at high OSNR with relatively small phase noise. For more general QAM-modulated signals, however, the training-initiated feedback-based method presented earlier can achieve much reliable performance with even lower implementation complexity. The BFS method presented in the "Blind Frequency Search Algorithm" section has the potential to achieve much faster FO estimation at the expense of higher implementation complexity.

TO REDUCE THE IMPLEMENTATION COMPLEXITY OF THE SINGLE-STAGE BPS ALGORITHM, A NEW MULTISTAGE HYBRID BPS AND ML PHASE RECOVERY ALGORITHM HAS RECENTLY BEEN PROPOSED.

HARDWARE-EFFICIENT PHASE RECOVERY ALGORITHM

As mentioned in the "Introduction" section, if the common symbol-by-symbol time-interleaved parallel processing architecture is utilized, feedback-based phase recovery algorithms (widely used for lower-speed wireless/baseband systems) are generally not suitable for high-speed coherent optical systems due to its poor phase noise tolerance, especially for higher-order QAM systems. Using a block-by-block time-interleaved architecture (i.e., the so-called super scalar parallelization) [21] in principle can reduce the feedback delay. For this method, however, not only does extra overhead have to be introduced within each data block to start the PLL operation on a block-by-block basis, substantial memory/buffer units are also needed to realize block-by-block parallelization, which not only complicate the circuit design, but also introduce considerable latencies.

This article will focus on several promising feedforward phase recovery algorithms that were proposed or demonstrated over the past few years. These include the blind phase recovery algorithms such as the classic Mth-power-based algorithm [13], the single-stage blind phase search (BPS) algorithm [22], the multiple-stage hybrid BPS, and maximum likelihood (ML) estimation algorithm [23], as well as the multistage hybrid PLL and ML algorithm [24]. Just recently, a training-assisted two-stage phase recovery algorithm [15] has also been proposed to address the well-known cyclic phase slipping problem to remove the need for differential coding and decoding that is typically required for using the blind phase recovery algorithms. In the following we will discuss in detail on these algorithms, including the pros and cons of each algorithm.

SINGLE-STAGE BLIND PHASE RECOVERY ALGORITHMS

Mth-POWER-BASED ALGORITHMS

For this algorithm, the signal is first raised to the Mth-power to erase data modulation and the carrier phase is extracted from other noise sources through a low-pass phase-averaging filter with the assumption that the carrier phase does not change (or with negligible change) over multiple symbol periods. This method was originally proposed for M-PSK systems because their data modulation can be completely removed by simply doing an Mth-power operation [13]. Later this algorithm was extended to QAM-modulated systems (such as 16 QAM) with some modifications, for example, by introducing constellation partitioning and transformation-based techniques [25], [26]. But even with these modifications, the phase noise tolerance still decreases for high-order QAMs because only a small portion of the signal constellation points having equal phase spacing are useful for the phase estimation.

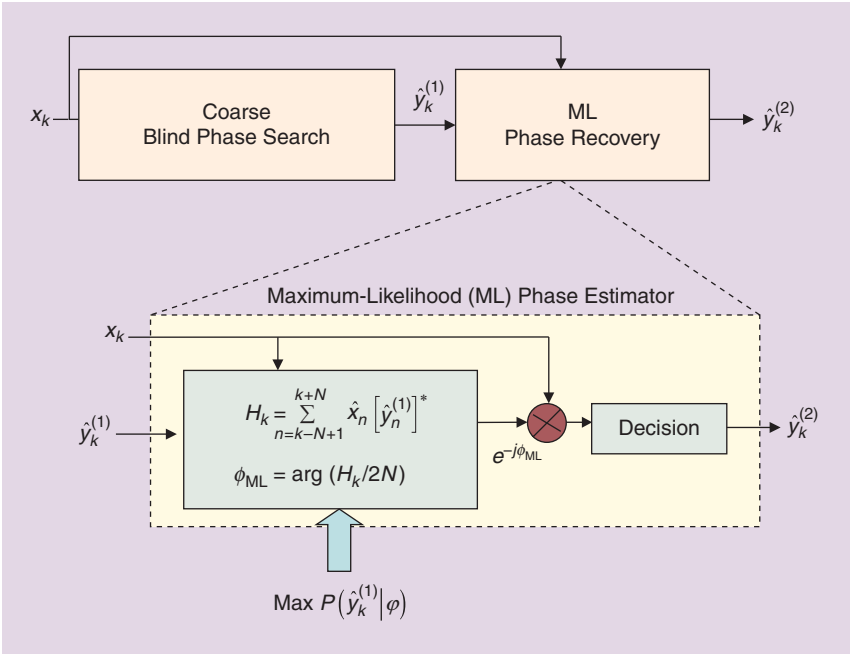
BPS ALGORITHM

The BPS algorithm was first introduced for coherent optical transmission systems in [22]. For this algorithm, the carrier phase is scanned over a limited phase range ($[0, \pi/2]$ for a square QAM) at fixed or variable phase increments, and the decisions made following each trial phase is approximated as the reference signal for MSDE calculation. The optimal phase is the one that gives the minimum MSDE. This method can achieve nearly optimum line width tolerance (in high OSNR region and assume no signal distortion) for any QAM formats but the implementation complexity increases with the modulation order, and can be very high for a high-order QAM. For example, to achieve a close to optimal performance, the required number of trial phases need to be greater than 16 for 16 QAM, and greater than 64 for 64 QAM [22]. As a comparison, the implementation complexity for a QPSK-modulated system using fourth-power-based phase recovery algorithm is only equivalent to test two phase angles with BPS method.

MULTISTAGE BLIND PHASE RECOVERY ALGORITHMS

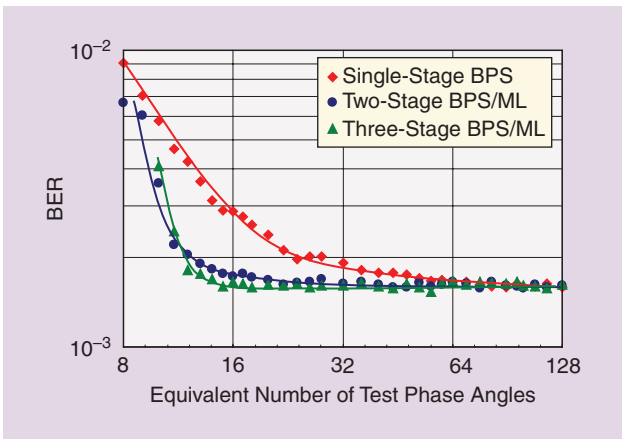
MULTISTAGE HYBRID BLIND PHASE SEARCH AND MAXIMUM LIKELIHOOD ALGORITHM

To reduce the implementation complexity of the single-stage BPS algorithm, a new multistage hybrid BPS and ML phase recovery algorithm has recently been proposed [23]. For this method, a coarse BPS method is used in the first stage to find a rough location of the optimal phase angle. Decisions made following this rough phase estimation are then used for more accurate phase estimation through an ML phase estimator [27] in the second stage as is shown in Figure 8. To further improve the accuracy of phase estimation, one can add more than one ML phase estimation stage to approach the optimal phase angle iteratively.



[FIG8] A schematic illustration of the proposed multistage hybrid BPS and ML phase recovery algorithm. (Only the two stage is shown.)

The effectiveness of the proposed method has been verified by numerical simulation using a 38 Gbaud square PDM-64QAM system as is shown in Figure 9, where the required equivalent number of test phase angle (ENTPA) using three different feed-forward carrier phase recovery scenarios, the single-stage BPS method and the proposed two-stage and three-stage hybrid BPS and ML phase recovery schemes has been displayed. For these investigations, the laser phase noise for both the signal source and the LO is assumed to be 100 kHz. The received OSNR in 0.1-nm noise bandwidth and for a single polarization is 28 dB. From Figure 9 one can find that, to achieve a performance that is close to the optimum, the single-stage BPS method needs to test about 64 different phase angles, while the proposed three-stage hybrid BPS/ML algorithm only needs to equivalently test 18 different phase angles (14 test phase angles used in the first coarse BPS stage plus two cascaded ML phase estimation stages), resulting in a reduction of computational effort by more than a factor of three.



[FIG9] Simulated results on the required ENTPA for square 64-QAM using three different feedforward carrier phase recovery scenarios.

MULTISTAGE HYBRID PHASE-LOCKED LOOP AND MAXIMUM LIKELIHOOD ALGORITHM

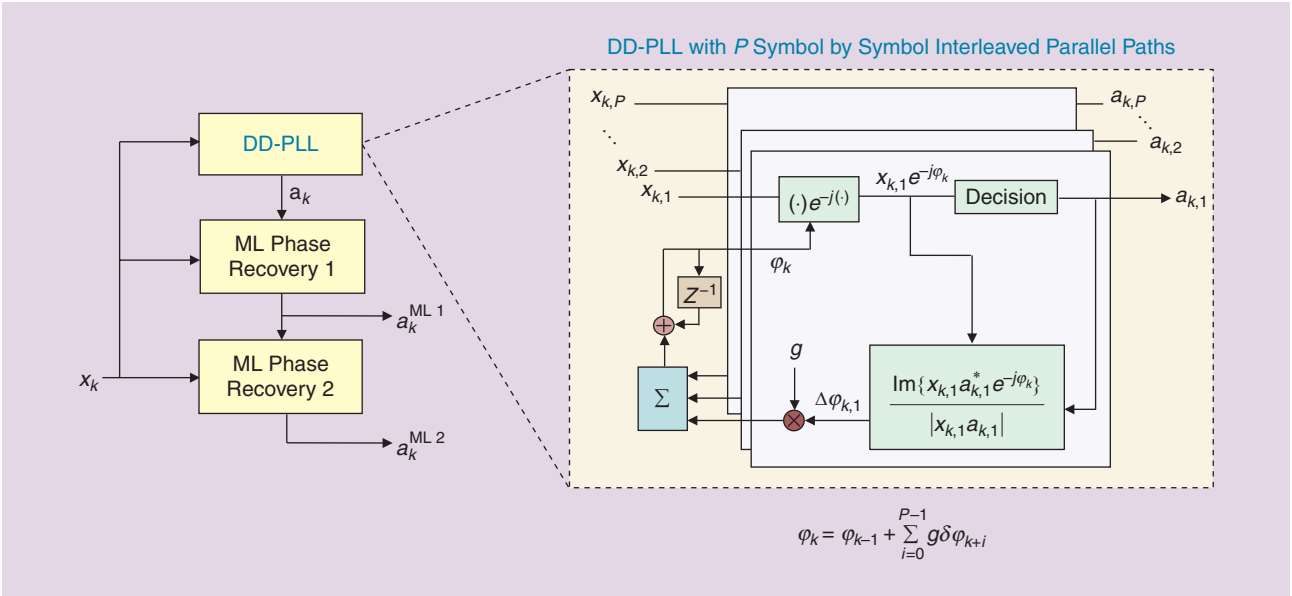
To further reduce the implementation complexity, one can use the classic decision-directed PLL for coarse phase recovery and one or two ML estimators for phase fine-tuning [24] as is shown in Figure 10. Such a multistage configuration can reduce the implementation complexity by more than one order of magnitude as compared to the single-stage BPS method for a 64-QAM system, but at the cost of slightly reduced laser phase noise tolerance when the degree of parallelization is high.

The effectiveness of the proposed multistage PLL and ML phase recovery algorithm has been tested in a 9.4 Gbaud 64 QAM (single polarization) back-to-back experiment as is shown in Figure 11(a) and (b). Figure 11(a) shows the impact of parallel processing on the proposed algorithms for a constant 23 dB OSNR. One can observe that the proposed algorithm with two ML estimators can achieve the same bit-error ratio (BER) performance as the BPS method for the symbol-by symbol interleaved parallel path P up to 20. The BER performance versus OSNR level for $P = 16$ is given in Figure 11(b). The proposed algorithm can achieve performance similar to the BPS method for a wide range of OSNR levels with BER ranging from 2×10^{-2} to close to 10^{-4}.

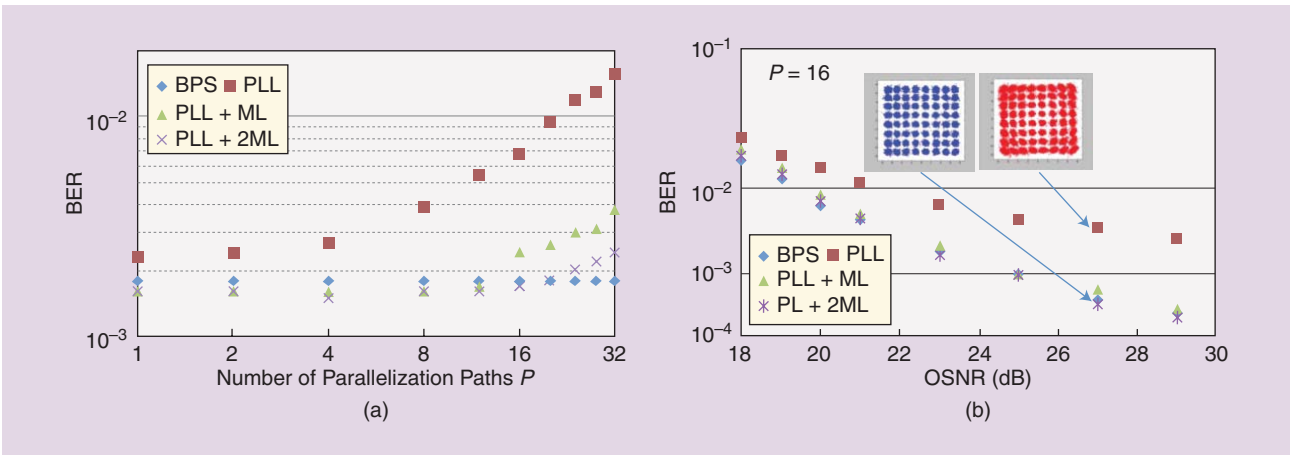
It should be noted that, for the above discussed multistage phase recovery algorithms, one can also use Mth-power-based algorithms as the coarse recovery and/or using a constrained BPS algorithm in the fine phase recovery stage.

TRAINING-ASSISTED PHASE RECOVERY ALGORITHM

For the blind phase recovery algorithms discussed above, differential coding and decoding has been used to mitigate the detrimental cyclic phase slipping problem. But the use of differential coding and decoding will result in up to 1 dB (depending on the used modulation formats and system operating points) sensitivity penalty for a system using a high coding gain forward error correction (FEC). To remove the need of differential coding and decoding and improve the receiver sensitivity, a new training-assisted two-stage phase recovery algorithm has recently been proposed and demonstrated [15]. This new algorithm could effectively mitigate the cyclic slipping problem by using a relatively small training overhead ($\cong$ 2%). The functional block illustration of the proposed phase recovery algorithm is shown in Figure 12. For this new method, training symbols that are known at the receiver are



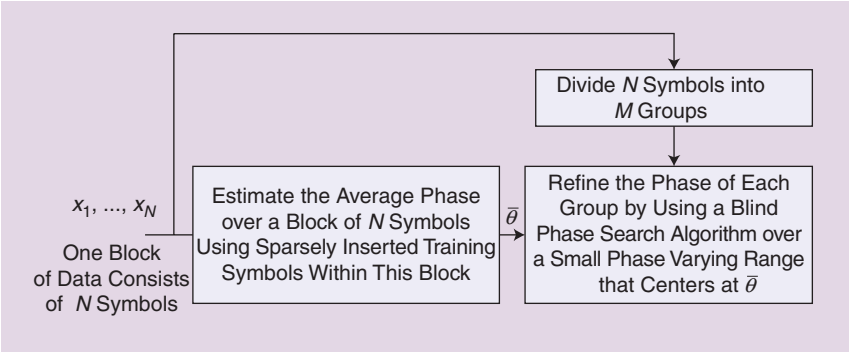
[FIG10] A schematic illustration of the proposed multistage hybrid PLL/ML phase estimator. P: symbol by symbol interleaved parallel paths. DD-PLL: decision-directed PLL.



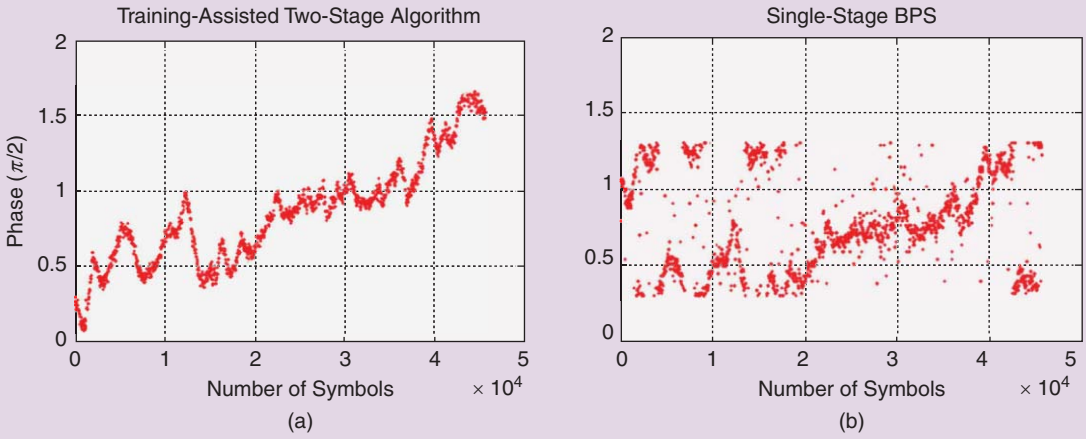
[FIG11] Experimental results for a 9.4-Gbaud 64-QAM system, where (a) shows the BER performance for different degrees of parallelization with 23-dB OSNR and (b) shows the BER performance at different OSNR levels. PLL delay D = 5.

periodically inserted into the data stream to assist in the phase recovery. To reduce the training overhead, training symbols are only sparsely inserted at the transmitter. At the receiver, the received data is processed block by block, where each block consists of at least two training symbols. For each block of data, first the average phase over this block is estimated by using the inserted training symbols through, for an example, the well-known ML-based phase estimation method [27]. Second, each block is divided into multiple groups, and then the phase of each group is refined by using a BPS algorithm [22] over a small phase-varying range that is centered at the average phase estimated through the training symbols.

The validity of this new method has been verified by a 400 Gb/s experiment using a time-domain hybrid 32–64 QAM modulation format [15]. It has been found that this new



[FIG12] A functional block illustration for the proposed phase recovery algorithm.



[FIG13] Recovered carrier phases using two different algorithms: training-assisted two-stage algorithm and single-stage BPS algorithm, for a 400 Gb/s WDM transmission experiment using a time-domain hybrid PDM 32–64 QAM modulation format.

two-stage algorithm is very robust against cyclic phase slipping or phase jumping problem as is shown Figure 13, where the recovered carrier phases using two different algorithms: the proposed training-assisted two-stage algorithm and the conventional BPS are displayed for a back-to-back measurement with OSNR = 24.2 dB (corresponding to a BER 2e-2 for using training-assisted algorithm). One can see that there was no phase jump when using the proposed training-assisted two-stage algorithm, whereas the phase-jump problem was severe (due to low OSNR, nonideal equalization, and imperfect signal constellation in this experiment) when using the conventional single-stage BPS algorithm, which mandates the use of differential coding.

This new two-stage algorithm can achieve comparable or even better (in the low OSNR region) phase noise tolerance than the single-stage BPS method with much lower implementation complexity. The required about 2% training overhead can be further reduced by exploring joint phase recovery over two orthogonal polarization states for current polarization-multiplexed transmission systems or joint phase recovery over multiple spatial channels for future space division multiplexed systems [28]. Joint phase recovery over multiple spatial channels can also be explored to improve the line width tolerance.

As a summary, Table 1 gives a brief comparison of the aforementioned phase recovery algorithms in terms of the achievable hardware efficiency, line width tolerance, and several other metrics.

CONCLUSIONS AND DISCUSSIONS

For regular RZ or NRZ pulse-shaped coherent optical transmission systems, CD and PMD tolerable clock recovery can be realized by using a feedback-based synchronous sampling method through the use of a newly proposed frequency-domain timing error detector, which is derived from the classic Gardner timing error detector operating at two samples per symbol. Clock recovery can also be realized by using feedforward-based methods, but the implementation complexity increases due to the need for a higher sampling rate (or alternatively with additional modulation and filtering operation) as well as the need for an extra digital interpolator. For a Nyquist-shaped optical system, two new algorithms have been proposed: one is based on a modified Gardner timing error detector and the other uses the standard deviation of the signal amplitude to measure the timing error. But more work need to be done to understand their channel distortion tolerance performance.

Significant progress has been made toward developing hardware-efficient carrier frequency and phase recovery algorithms that wok for arbitrary QAMs. Among the various developed frequency recovery algorithms, the training-initialized feedback blind frequency recovery algorithm is probably the best choice for the common continuous-mode coherent receiver, mostly due to its low implementation complexity and large frequency offset tolerance (up to half of the symbol rate). For the carrier phase recovery algorithm, the newly proposed training-assisted two-stage

[TABLE 1] A COMPARISON OF THE DISCUSSED PHASE RECOVERY ALGORITHMS.

	HARDWARE COMPLEXITY	LINE WIDTH TOLERANCE	APPLICABLE MODULATION	TRAINING OVERHEAD	DIFFERENTIAL CODING
*M*th-POWER	LOW	HIGH	MPSK	NO	YES
BPS	HIGH	HIGH	ANY QAMS	NO	YES
BPS/ML	MODERATE	HIGH	ANY QAMS	NO	YES
PLL/ML	LOW	MODERATE	ANY QAMS	NO	YES
TRAINING ASSISTED	MODERATE	HIGH	ANY QAMS	≤2%	NO

algorithm holds great promise for coherent optical systems employing high-order coherent modulation formats, because this method could eliminate the need for differential coding and decoding (thus improve the receiver sensitivity), and also has relatively low implementation complexity.

So far, most of the developed carrier recovery algorithms are optimized based on the additive circular Gaussian noise assumption (i.e., the laser phase noise is only corrupted by additive circular Gaussian noise) and has assumed ideal equalization, but in a real optical communication system there exist not only Gaussian noise, but also nonlinear phase and amplitude noise, as well as the interaction between the equalizer and the phase noise [29]. How to optimize the phase recovery algorithm for these realistic nonlinear optical communication systems with imperfect equalization is still a significant challenge and also an interesting topic for future research. In addition, with the development of spectral-efficiency agile modulation formats [15], the modulation format-independent DSP platform [30] is also becoming an interesting and challenging topic. Much more work can be expected toward this direction in the coming years.

AUTHOR

Xiang Zhou (zhoux@ieee.org) received his Ph.D. degree in electrical engineering from the Beijing University of Posts and Telecommunications in 1999. From 1999 to 2001, he was with Nangang Technological University, Singapore, as a research fellow, where he conducted research on optical CDMA and wideband Raman amplification. He was with AT&T Labs-Research, Middletown, New Jersey, from 2001 to 2013, working on various aspects of long-haul optical transmission and photonic networking technologies, including Raman amplification, polarization- and reflection-related impairments, optical power transients control (in dynamic optical networks), advanced modulation formats, and digital signal processing for high-speed transmission (100 Gb/s, 400 Gb/s, and beyond). He recently joined Google Inc. and turned his interest to data center interconnection problems. He has authorized/coauthored more than 100 peer-reviewed journals and conferences publication and holds more than 30 U.S. patents. He currently serves as an associate editor of *Optics Express* and is on the technical program committee for several IEEE and Optical Society of America (OSA) conferences, including the Optical Fiber Communication Conference. He is member of OSA and a Senior Member of the IEEE.

REFERENCES

[1] M. G. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb. 2004.

[2] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Opt. Exp.*, vol. 16, no. 2, pp. 873–879, 2008.

[3] K. Kikuchi, "Clock recovering characteristics of adaptive finite-impulse-response filters in digital coherent optical receivers," *Opt. Exp.*, vol. 19, no. 6, pp. 5611–5619, Mar. 2011.

[4] Z. Tao, H. Zhang, A. Isomura, L. Li, T. Hoshida, and J. C. Rasmussen, "Simple, robust, and wide-range frequency offset monitor for automatic frequency control in digital coherent receivers," presented at the European Conf. Optical Communications 2007, Paper 03.5.7.

[5] H. Sun and K.-T. Wu, "A novel dispersion and PMD tolerant clock phase detector for coherent transmission systems," presented at the Optical Fiber Communication Conf. 2011, Los Angeles, CA, Mar. 2011, Paper OMJ4.

[6] F. M. Gardner, "A BPSK/QPSK timing-error detector for sampled receivers," *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 423–429, May 1986.

[7] R. Schmogrow, M. Winter, M. Meyer, D. Hillerkuss, A. Ludwig, B. Nebendahl, S. Ben-Ezra, J. Meyer, M. Dreschmann, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold, "Real-time Nyquist pulse generation beyond 100 Gbit/s and its relation to OFDM," *Opt. Exp.*, vol. 20, no. 1, pp. 317–337, Dec. 2012.

[8] M. Yan, Z. Tao, L. Dou, L. Li, Y. Zhao, T. Hoshida, and J. C. Rasmussen, "Digital clock recovery algorithm for Nyquist signal," presented at the Optical Fiber Communication Conf. 2013, Paper OTu2I.7.

[9] N. Stojanovic, C. Xie, Y. Zhao, B. Mao, and N. G. Gonzalez, "Modified gardner phase detector for Nyquist coherent optical transmission systems," presented at the Optical Fiber Communication Conf. 2013, Paper JTh2A.50.

[10] M. Oerder and H. Meyr, "Digital filter and square timing recovery," *IEEE Trans. Commun.*, vol. 36, no. 5, pp. 605–612, Mar. 1988.

[11] W. Zhu, Y. Yan, M.O. Ahmad, and M. N. S. Swamy, "Feedforward symbol timing recovery technique using two samples per symbol," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 11, pp. 2490–2500, Nov. 2005.

[12] C. R. S. Fludger, T. Duthel, P. Hermann, and T. Kupfer, "Jitter tolerant clock recovery for coherent optical receivers," presented at the Optical Fiber Communication Conf. 2013, Paper OTh1F.3.

[13] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 4, pp. 543–551, July 1983.

[14] X. Zhou, J. Yu, M. -F. Huang, Y. Shao, T. Wang, L. E. Nelson, P. D. Magill, M. Birk, P. I. Borel, D. W. Peckham, and R. Lingle, "64-Tb/s, 8 b/s/Hz, PDM-36QAM transmission over 320 km using both pre- and post-transmission digital signal processing," *J. Lightwave Technol.*, vol. 29, no. 4, pp. 571–577, Feb. 2011.

[15] X. Zhou, L. E. Nelson, P. Magill, R. Isaac, B. Zhu, D. W. Peckham, P. Borel, and K. Carlson, "High spectral efficiency 400 Gb/s transmission using PDM time-domain hybrid 32-64QAM and training-asssisted carrier recovery," *J. Lightwave Technol.*, vol. 31, no. 7, pp. 999–1005, Apr. 2013.

[16] I. Fatadin, D. Ives, and S. J. Savory, "Compensation of frequency offset for differentially encoded 16- and 64-QAM in the presence of laser phase noise," *IEEE Photon. Technol. Lett.*, vol. 22, no. 32, pp. 176–179, Feb. 2010.

[17] I. Tselniker, N. Sigron, and M. Nazarathy, "Joint phase noise and frequency offset estimation and mitigation for optically coherent QAM based on adaptive multisymbol delay detection," *Opt. Exp.*, vol. 20, no. 10, pp. 10944–10962, 2012.

[18] A. Leven, N. Kaneda, U. V. Koc, and Y. K. Chen, "Frequency estimation in intradyne reception," *IEEE Photon.* Technol. Lett., vol. 19, no. 6, pp. 366–368, Mar. 2007.

[19] M. Selmi, Y. Jaouen, and P. Ciblat, "Accurate digital frequency offset estimator for coherent PolMux QAM transmission systems," presented at the European Conf. Optical Communications 2009, Paper P. 3.08.

[20] Y. Sun and X. Zhou, "Blind carrier frequency recovery methods for coherent receivers using QAM modulation formats," U.S. patent application 13/109853, May 17, 2011.

[21] A. Tolmachev, I. Tselniker, M. Meltsin, I. Sigron, and M. Nazarathy, "Efficient multiplier-free FPGA demonstration of polar-domain multi-symbol-delay-detector (MSDD) for high performance phase recovery of 16-QAM," presented at the Optical Fiber Communication Conf., 2013, Paper OM2C.8.

[22] T. Pfau, S. Hoffmann, and R. Noé., "Hardware-efficient coherent digital receiver concept with feed-forward carrier recovery for M-QAM constellations." *J. Lightwave Technol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[23] X. Zhou, "An improved feed-forward carrier recovery algorithm for coherent receiver with M-QAM modulation format," *IEEE Photon. Technol. Lett.*, vol. 22, no. 14, pp. 1051–1053, July 2010.

[24] X. Zhou and Y. Sun, "Low-complexity, blind phase recovery for coherent receivers using QAM modulation," presented at the Optical Fiber Communication Conf., 2011, Los Angeles, CA, 2011, Paper OMJ3.

[25] Y. Cao, S. Yu, Y. Chen, Y. Gao, W. Gu, and Y. Ji, "Modified frequency and phase estimation for M-QAM optical coherent detection," presented at the European Conf. Optical Communications 2010, Paper We.7.A.1.

[26] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning," *IEEE Photon. Technol. Lett.*, vol. 22, no. 9, pp. 631–633, May 2010.

[27] J. G. Proakis, *Digital Communication*, 3rd ed. New York: McGraw-Hill, 1995, ch. 6.

[28] M. D. Feuer, L. E. Nelson, X. Zhou, S. L. Woodward, R. Isaac, B. Zhu, T. F. Taunay, M. Fishteyn, J. F. Fini, and M. F. Yan, "Joint digital signal processing receivers for spatial superchannels, " *IEEE Photon.* Technol. Lett., vol. 24, no. 21, pp. 1957–1960, 2012.

[29] W. Shieh and K. P. Ho, "Equalization-enhanced phase noise for coherent-detection systems using electronic digital signal processing," *Opt. Exp.*, vol. 16, no. 20, pp. 15718–15727, Sept. 2008.

[30] Q. Zhuge, X. Xu, M. M. Osman, M. Chagnon, M. Qiu, and D. V. Plant, "Time domain hybrid QAM based rate-adaptive optical transmissions using high speed DACs," presented at the Optical Fiber Communication Conf. 2013, Paper OTh4E.6.

[SP]

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

Liang B. Du, Danish Rafique, Antonio Napoli, Bernhard Spinnler,
Andrew D. Ellis, Maxim Kuschnerov, and Arthur J. Lowery

Digital Fiber Nonlinearity Compensation

Toward 1-Tb/s transport

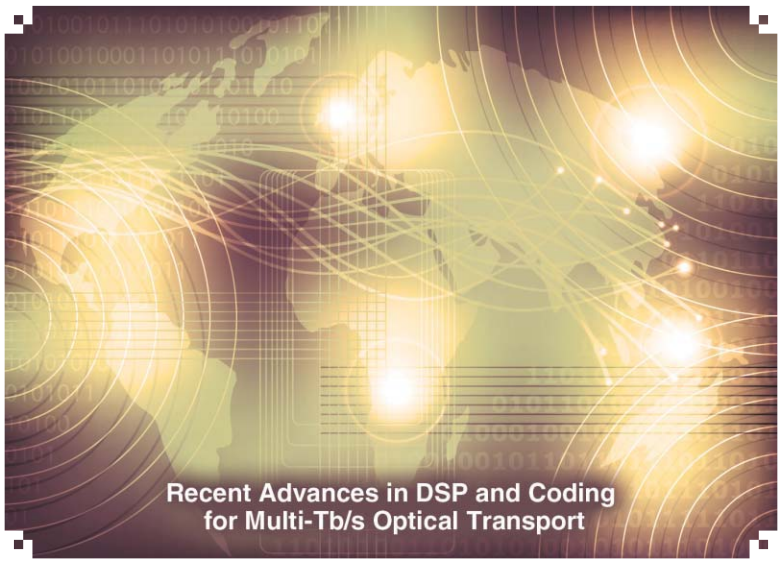


IMAGE LICENSED BY INGRAM PUBLISHING /JANMIKS

The world is connected by a core network of long-haul optical communication systems that link countries and continents, enabling long-distance phone calls, data-center communications, and the Internet. The demands on information rates have been constantly driven up by applications such as online gaming, high-definition video, and cloud computing. All over the world, end-user connection speeds are being increased by replacing conventional digital subscriber line (DSL) and asymmetric DSL (ADSL) with fiber to the home. Clearly, the capacity of the core network must also increase proportionally.

In the 1980s and 1990s, speeds in the core network were pushed forward by technologies such as external modulation and the erbium doped fiber amplifier (EDFA), which supported wavelength division multiplexing (WDM)—transmission of information using different colors. In the last decade, commercial systems have adopted coherent optical receivers and digital signal processing (DSP) to transmit 100 Gb/s per wavelength on over 80 wavelengths, enabling transmission capacity of ≥ 8 Tb/s per fiber. However, even with this impressive technological development, such system capacities will soon be insufficient to meet the demands of consumers.

The data rate in an optical transmission system is currently limited by amplified spontaneous emission (ASE), which determines the minimum power launched into each fiber span, and the interplay between chromatic dispersion (CD) and Kerr fiber nonlinearity, which limits the maximum launch power [1]. To increase the data rate of current-generation coherent systems, fiber nonlinearity compensation is required to enable higher launch powers, thereby providing enough optical signal-to-noise ratio (OSNR) to support larger constellation sizes [2]. Although

Digital Object Identifier 10.1109/MSP.2013.2288110
Date of publication: 12 February 2014

1053-5888/14/$31.00©2014IEEE

fiber nonlinearity and CD are both deterministic effects, their interplay along the length of the fiber is difficult to undo with optical components or DSP.

In this article, we describe the nonlinear fiber channel and its impact on optical communication systems. We then introduce two promising fiber nonlinearity compensation techniques: inverse fiber models [2]–[6] and pilot-aided phase compensation [7], [8], which may be useful in transmission systems operating at 400 Gb/s, 1 Tb/s, and beyond. We review the fundamental limitations of each method and illustrate their effectiveness, both for single-carrier and multicarrier systems. Finally, the practicality of various approaches is compared, in terms of the hardware requirements for digital nonlinear compensation.

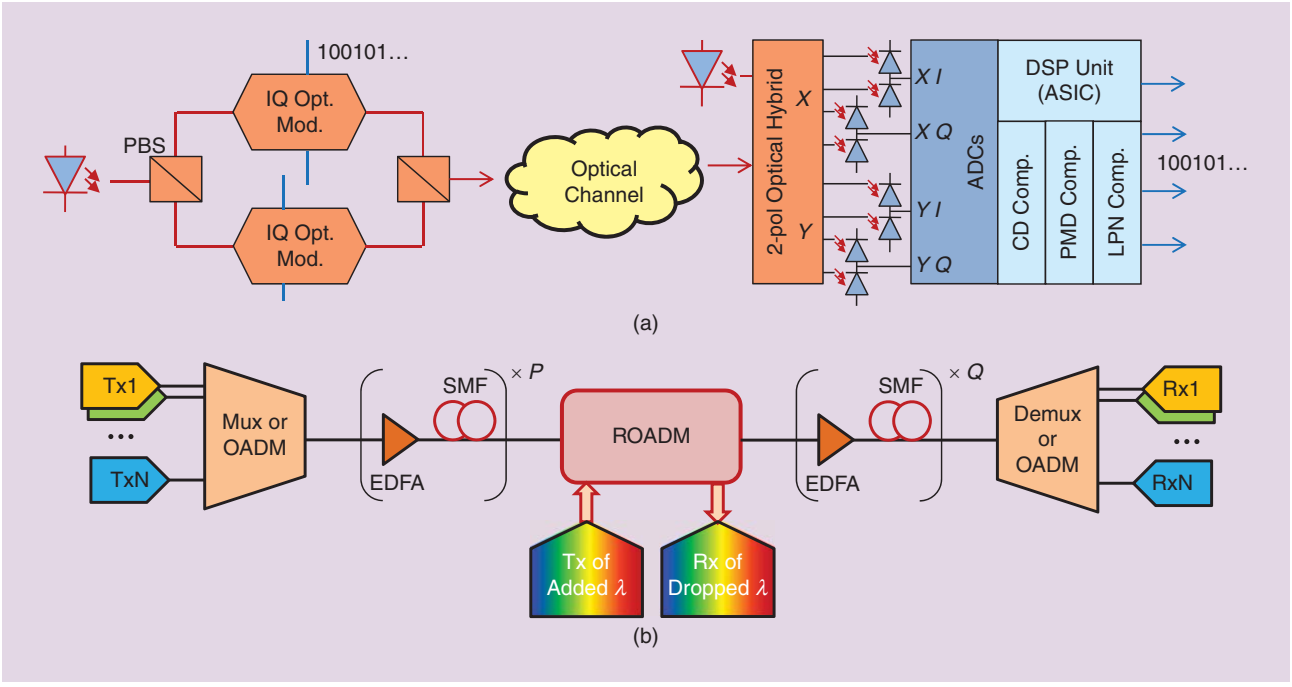
IN THIS ARTICLE, WE DESCRIBE THE NONLINEAR FIBER CHANNEL AND ITS IMPACT ON OPTICAL COMMUNICATION SYSTEMS.

REVIEW OF OPTICAL COMMUNICATIONS

BRIEF HISTORY

Optical fibers were first proposed by Nobel Laureate Kao as a suitable medium for transmitting growing telecommunications traffic at the time. He argued that the low loss and the large transmission bandwidth made them ideal for long-distance point-to-point communications. Systems using single-mode fibers (SMFs) became prominent in the 1980s, offering increased bandwidth and repeater spacing compared with copper lines. These systems operated at tens to hundreds of megabits per second. Technological breakthroughs enabled 960 Gb/s in a single fiber by 1999. These breakthroughs included dispersion engineered fibers to overcome pulse spreading from CD, wavelength selective multiplexers to combine multiple wavelengths in a single fiber (commonly referred as WDM), EDFAs to amplify multiple wavelengths, and external modulators to enable symbol rates of >10 Gbaud. Binary information was coded on the intensity of the light and detected with photodiodes, which produce an electrical current proportional to the intensity. These systems were referred to as *nonreturn-to-zero (NRZ)* or *intensity-modulated direct-detection systems*.

In 2005, Nortel (now Ciena) developed a 10 Gb/s NRZ system that used DSP to compensate for the CD in over 1,000 km of standard-SMF (S-SMF) by precompensating for CD at the transmitter. The high computational power required was provided by parallelization and application-specific integrated circuits (ASICs). Soon after, they introduced 40 Gb/s coherent optical quadrature-phase-shift keying (CO-QPSK) transmission [9]. This system used a coherent optical receiver that linearly mapped the optical field to the electrical field, enabling detection of phase modulation. Additionally, the receivers were polarization diverse, which allowed the two independently modulated polarizations of light to be detected. Signals were generated using integrated I-Q optical modulators with polarization beam splitters (PBSs)—a technique known as polarization multiplexing (POLMUX). Using only 10 Gbaud, 40 Gb/s was achieved. DSP was a key enabler of this technology as it compensated for CD, polarization mode dispersion (PMD), and laser phase noise (LPN), as illustrated in Figure 1(a). This technological breakthrough was followed by



[FIG1] (a) The coherent optical transmitter and receiver for current generation 40 Gb/s and 100 Gb/s systems. (b) The conceptual optical link with a single ROADM.

the introduction of 100 Gb/s CO-QPSK systems, which were commercialized a few years after the initial introduction of the first coherent 40 Gb/s chip set. CD compensation using DSP is commonly referred to as electronic dispersion compensation (EDC), contrasting it to in-line optical CD compensation.

OPTICAL LINKS

Figure 1(b) illustrates a typical optical transmission system. At each node, reconfigurable optical add-drop multiplexers (ROADMs) add and drop wavelength channels. The gain bandwidth of EDFAs is between 1,530 to 1,565 nm (191.56 THz to 195.95 THz) and is known as the C-band, which can support ~100 channels uniformly spaced at 50 GHz. Often, as a rule of thumb, the amplifier spacing is assumed to be ~70–80 km (~50 miles). However, in reality, amplifiers are placed wherever it is most convenient from cost and network architecture viewpoints.

IN ADDITION TO LINEAR IMPAIRMENTS, OPTICAL FIBERS ARE ALSO INHERENTLY NONLINEAR.

SMF only allows one mode to propagate, which prevents multipath interference; this mode supports two orthogonal polarizations. Typically SMF has 0.2–0.3 dB/km loss. The loss of each fiber span is compensated with an EDFA; this generates noise in the form of ASE, which is typically modeled as additive white Gaussian noise (AWGN) to the optical field. ASE is proportional to an amplifier's gain, so sufficiently high powers (around 1 mW) are launched into each span to provide enough OSNR at the end of the link.

Most terrestrial links are $<1{,}500$ km (a few are $>2{,}500$ km) and are often in mesh networks where neighboring wavelength channels may change at ROADMs. Submarine links can be up to 9,000-km long, have regularly spaced amplifiers, and typically have no in-line ROADMs.

LINEAR FIBER IMPAIRMENTS

SMF is not a perfect transmission medium: it induces different distortions than those common in copper or wireless systems. The greatest source of pulse spreading is caused by CD [3], which causes different spectral components of a signal to propagate at different velocities, giving a quadratic phase response [6]

$$\theta(\omega) = \frac{\beta_2}{2} L \cdot (\Delta\omega)^2, \tag{1}$$

where $\theta(\omega)$ is the phase at a given angular frequency, L is the length of the fiber, β_2 is the group velocity dispersion parameter, and $\Delta\omega$ is the difference in angular frequency from the center frequency of the channel.

Prior to DSP, CD was typically mitigated using dispersion compensating fiber (DCF)—fiber engineered to provide a dispersion opposing the CD of the transmission fiber. The loss of DCF often requires additional amplifiers. The locations of the DCF were determined by optimizing the performance of WDM systems; such communication networks are usually referred to as *dispersion managed* (*DM*) systems. Optical systems may also use dispersion engineered fibers for transmission, like dispersion shifted fibers (DSFs) (CD of ~0 ps/nm/km) or nonzero DSF (NZ-DSF) (CD of ~4 ps/nm/km). Compared with the most commonly deployed fiber—S-SMF (CD of ~17 ps/nm/km)—these lower dispersion fibers allow longer links without DCFs. However, the distortion due to fiber nonlinearity in WDM systems is reduced by the presence of CD because of the phase-array effect [6], [10]. Therefore, current systems are moving toward non-DM (NDM) S-SMF links and EDC for optimal performance.

SMF can support a pair of orthogonal polarizations, so two independent signals can be transmitted along a single SMF. Imperfections in the fiber manufacturing process cause light in the two principle polarizations to travel at slightly different speeds; this is called PMD. The pulse spreading caused by PMD is referred to as *differential group delay* (*DGD*). Scattering in the fiber randomly changes the signal's state of polarization (SOP) along the fiber on a millisecond timescale, thus a signal on any initial SOP travels through a random combination of fast and slow paths to reach its destination. The random scattering of SOP along the fiber causes the mean DGD of links to be proportional to the square root of the length. Current-generation fiber links have DGDs typically <24 ps. This is considerably less than their accumulated CD, which can be in the order of nanoseconds.

FIBER NONLINEARITY

In addition to linear impairments, optical fibers are also inherently nonlinear. This nonlinearity is caused by the optical fiber's refractive index being slightly dependent on the optical power propagating within the fiber, called the Kerr effect. Therefore, the signal imposes an instantaneous phase retardation proportional to the instantaneous signal power given by [6]:

$$\phi_X(t) = \gamma L_{\text{eff}} \left[|E_X(t)|^2 + \frac{2}{3} |E_Y(t)|^2 \right], \tag{2a}$$

$$\phi_Y(t) = \gamma L_{\text{eff}} \left[|E_Y(t)|^2 + \frac{2}{3} |E_X(t)|^2 \right], \tag{2b}$$

where $\phi_X(t)$ and $\phi_Y(t)$ are the instantaneous phase shifts of the signal in the X and Y polarizations respectively, $E_X(t)$ and $E_Y(t)$ are the field of the optical signal in two orthogonal SOPs normalized to the square root of the optical power, γ is the nonlinearity parameter, and L_{eff} is the effective length [6], which is given by

$$L_{\text{eff}} = \frac{1 - \exp(-\alpha L)}{\alpha}, \tag{3}$$

where α is the attenuation coefficient in nepers/m.

The combined effects of CD, PMD, attenuation, and fiber nonlinearity can be described by [11]

$$\frac{\partial E_X(z,t)}{\partial z} = \left(-\frac{j}{2}\beta_2\frac{\partial^2}{\partial t^2} - \frac{\alpha}{2} + j\gamma\left[|E_X|^2 + \frac{2}{3}|E_Y|^2\right]\right)E_X(z,t), \tag{4a}$$

$$\frac{\partial E_Y(z,t)}{\partial z} = \left(-\frac{j}{2}\beta_2\frac{\partial^2}{\partial t^2} - \frac{\alpha}{2} + j\gamma\left[|E_Y|^2 + \frac{2}{3}|E_X|^2\right]\right)E_Y(z,t). \tag{4b}$$

The first term on the right describes the effects of CD, the second the loss of power, and the final the nonlinear phase shifts. This is commonly known as the nonlinear Schrödinger equation (NLSE).

To use these equations for nonlinearity compensation, the SOP must be accurately tracked along the link. This is almost impossible. However, the SOP tends to scatter over a very short distance, which allows the powers in the two polarizations to be averaged. This is described by the Manakov equations [11]:

$$\frac{\partial E_{X,Y}(z,t)}{\partial z} = \left(-\frac{j}{2}\beta_2\frac{\partial^2}{\partial t^2} - \frac{\alpha}{2} + j\gamma\frac{8}{9}\left[|E_Y|^2 + |E_X|^2\right]\right) \times E_{X,Y}(z,t). \tag{5}$$

There are currently no closed-form solutions to the NLSE. However, numerical simulations can be used to optimize optical systems. Most commercial software, such as VPItransmissionMaker, use the split-step Fourier method (SSFM) to model the nonlinear fiber. Each step contains a linear and a nonlinear section [6], [12]. A large number of short steps are required to obtain accurate simulation results. This is a very computationally intensive process and real-time simulations at Gb/s rates are well beyond the computational power available today.

DSP OF CURRENT-GENERATION COHERENT SYSTEMS

The important functional blocks of current-generation 100 Gb/s coherent systems are shown in Figure 1(a). To maximize the spectral efficiency, these systems use both orthogonal polarizations in the fiber, and signal on both the in-phase (I) and quadrature (Q) dimensions. After accounting for the overheads due to framers and forward error correction (FEC), the required gross bit rate is around 112 Gb/s. Since these systems employ four independent dimensions, a symbol rate of 28 Gbaud is used. Typically, these systems operate on a 50-GHz WDM grid, achieving the spectral efficiency of ~2.2 bits/s/Hz and occupy around 60% of the spectrum.

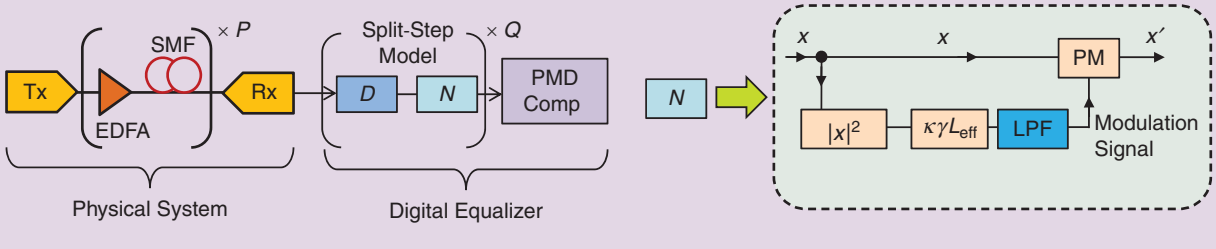
At the receiver, DSP is used to compensate for the impairments of the physical channel. The bulk of the CD in an NDM link is compensated by a frequency-domain equalizer, while the residual CD and PMD are equalized by using a 2×2 multiple-input, multiple-output (MIMO) butterfly structure, realized by means of a finite impulse response (FIR) filter. This can be realized either by blind- or training-based methods. Laser phase noise is then digitally tracked and compensated to obtain the transmitted constellations. To reduce the computational complexity of the equalizer, fibers are assumed to be linear (i.e., nonlinear mitigation is not present). The maximum reach of these systems is limited by the ASE generated within the EDFA stages and the interplay between CD and fiber nonlinearity [10], [13].

INVERSE MODEL NONLINEARITY COMPENSATION

NONLINEAR DIGITAL BACKWARD PROPAGATION

The NLSE governs the propagation of signals in optical fiber [see (4)]. If the polarities of loss, dispersion, PMD, and nonlinearity parameters are reversed, the same equation can be used to model a fictitious fiber with exactly opposite characteristics compared to the real fiber used for transmission. The distortions produced by the fictitious fiber will cancel the ones of the real fiber [14]. Although such a fiber does not exist, this concept paved the way for digital backward propagation (DBP), which digitally solves the reversed NLSE.

DBP was first proposed in 2005. A digital inverse link model was placed before the transmitter to launch a predistorted signal into the physical transmission link [15]. The predistortion is removed by the physical link so the received signal is almost perfect. Since the SOP along the link is almost impossible to track, the Manakov equations, shown in (5), were used. In 2008, an inverse nonlinear model was used after a coherent receiver [3]–[5], similar to that shown in Figure 2. The addition of the low-pass filter (LPF) was later proposed as explained in the section "Filtered DBP." The nonlinear sections are memoryless, and are therefore computationally trivial. However, they separate the linear sections, which are FIR filters. Because the linear sections are typically implemented in the frequency domain using fast Fourier transforms, a linear step representing a "long" section of fiber requires only slightly more computational effort than a "short" section. Therefore, the number of



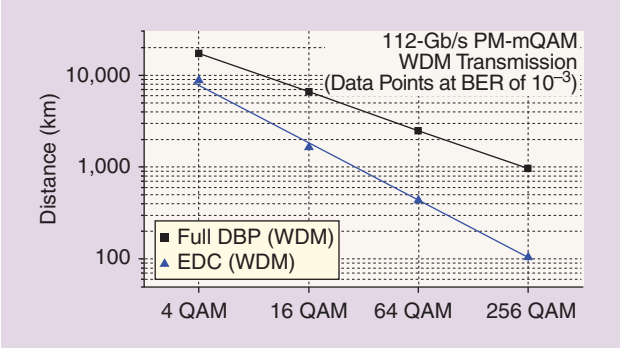
[FIG2] **An optical system with DBP: dispersion section (*D*), nonlinear section (*N*), number of amplified spans in the link (*P*), number of steps used in DBP (*Q*), scaling factor to be optimized κ, and phase modulator (PM).**

steps determines the total computational effort (discussed in the section "Complexity of Nonlinearity Compensation"). Most proof-of-concept demonstrations of fiber nonlinearity compensation have used offline signal processing; therefore, the processing time per bit was not an issue.

The upper bound in the performance enhancement of DBP was investigated by compensating for intrachannel and interchannel nonlinear fiber impairments in systems using different modulation formats [2]. In this work, numerous short steps were used. Although this is impractical in a real system because of the excessive computational power required, it established the fundamental limits of DBP. A nine-channel WDM system was considered. The power and modulation format was the same for all channels. The bit rate for each channel was fixed to 112 Gb/s; the baud rate was changed accordingly and the baud rate to channel spacing ratio was fixed to 0.56. Figure 3 shows full-band DBP significantly increases the reach for all the modulation formats tested. For QPSK, the reach is improved by a factor of three. This benefit becomes even greater for higher-order modulation formats. Counterintuitively, full-band DBP does not support unlimited transmission distances because intermixing of the ASE and signal eventually limit the performance [16]. Since the ASE component is random, nonlinear products originating from ASE cannot be compensated, even with the most detailed model.

Even though full-band DBP gives significant advantages, the computational cost is greatly increased because the maximum step size is inversely related to the total bandwidth of the WDM signal. Additionally, full-band DBP will only work for point-to-point links and places additional requirements on the receiver architecture [11], as explained in the section "Nonlinearity Compensation in 1-Tb/s Systems." Single-channel DBP (SC-DBP) has been more widely researched because it is typically considered to be more realistic in the near- to mid-term. Since a single-wavelength signal is back-propagated, SC-DBP only compensates for intrachannel nonlinearity, e.g., self-phase modulation (SPM). However, an interesting tradeoff between performance and complexity may be derived by only back-propagating limited WDM bandwidth, such as DBP of an entire superchannel (further discussed in the section "Digital Nonlinearity Compensation in Next-Generation Systems"), rather than full-band DBP.

AN INTERESTING TRADEOFF BETWEEN PERFORMANCE AND COMPLEXITY MAY BE DERIVED BY ONLY DIGITALLY BACK-PROPAGATING LIMITED WDM BANDWIDTH.



[FIG3] **The maximum transmission distance at optimal power achieving a BER of 10^{-3} for a WDM system.**

INTENSITY-DRIVEN PHASE MODULATION

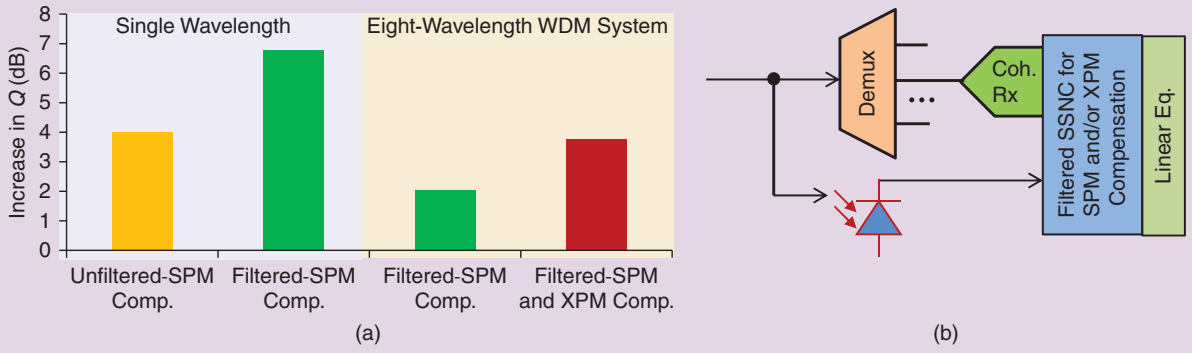
If CD is ignored, then only a single nonlinear step is required to fully mitigate deterministic nonlinear fiber impairments, thereby making it computationally trivial. In 2002, Xu and Liu successfully compensated for SPM in a DM link by phase modulating a differential PSK signal by its own intensity [17]. This system would perfectly compensate dispersionless nonlinear fiber. In 2007, digital implementations were proposed for coherent optical systems [18]. Neglecting CD means only the memoryless nonlinear section is required, which significantly simplifies the hardware complexity. Single-step nonlinearity compensation (SSNC) will still work in dispersive links if the CD in the link is sufficiently low, such as NZ-DSF links and DM S-SMF links [19], [20]. In these low-dispersion links, the phase shifts induced by CD causes the four-wave mixing (FWM) efficiency to be lower for larger frequency separations [10], [19]; the nonlinear phase shifts produced by high-frequency intensity fluctuations are therefore weaker than those produced by low-frequency intensity fluctuations. Inserting an LPF in the nonlinear section, as shown in Figure 2, was shown to improve SSNC [19]. Figure 4(a) shows the increase in signal quality, Q [18], at nonlinearity-limited powers from unfiltered- and filtered-SSNC for 100 Gb/s coherent optical orthogonal frequency division multiplexing (CO-OFDM) in a typical DM link.

Figure 4(a) shows that in a WDM system, the benefit from SPM compensation alone is greatly reduced because XPM is the dominant source of nonlinear distortion. XPM can be compensated using filtered SSNC by detecting the low-frequency intensity fluctuations of a band of WDM channels using a photodiode before the demultiplexer, as shown by Figure 4(b). The photodiode output can then be scaled and digitally filtered for optimal XPM compensation [20].

FOLDED DBP

Another possible DSP technique is folded DBP [21], where the symmetry of DM systems is exploited to greatly reduce the number of DBP steps without compromising performance. In a fully dispersion-compensated link (zero residual CD after each span), the solution collapses to SSNC. In typical DM links with some residual CD, two or three steps may still be sufficient.

Figure 5(a) shows numerical simulation results of folded DBP for a 10×224-Gb/s 16-QAM system with an eight-span DM S-SMF link [22] on a standard International Telecommunication Union (ITU) grid of 50 GHz. The case labeled with folded DBP uses only three nonlinear sections separated by two dispersive sections; DBP-8s use one DBP step per each S-SMF



[FIG4] (a) The nonlinear threshold increase for 100-Gb/s CO-OFDM in a DM S-SMF link; (b) the SSNC for SPM and XPM compensation.

span; DBP-15s use a DBP step for each S-SMF and DCF span. The performance of folded DBP is clearly comparable to DBP-8s and DBP-15s. Different spans can be merged because a small amount of residual CD is beneficial for the approximation carried out by DBP, similar to that in the Wiener-Hammerstein model [23]. Figure 5(a) shows there is no penalty due to this coarse approximation and the computational cost is reduced by a factor of five.

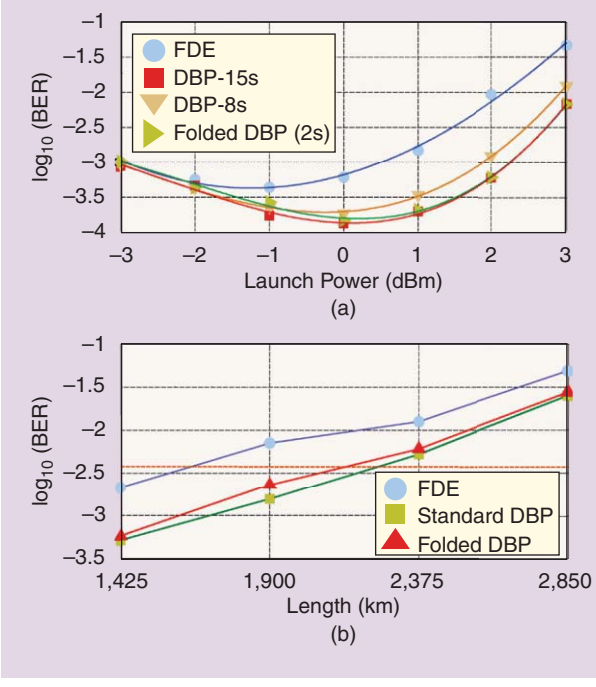
Similar experimental results for WDM transmission are displayed in Figure 5(b) on the setup described in [24], where 10×111 Gb/s POLMUX-RZ-DQPSK signals were propagated by employing standard ITU grid with 50-GHz channel spacing. Folded DBP has similar performance to DBP with one step per span and increases the reach by about 500 km, assuming a hard FEC limit of 3.8×10^{-3}, represented by the dashed orange line.

FILTERED DBP

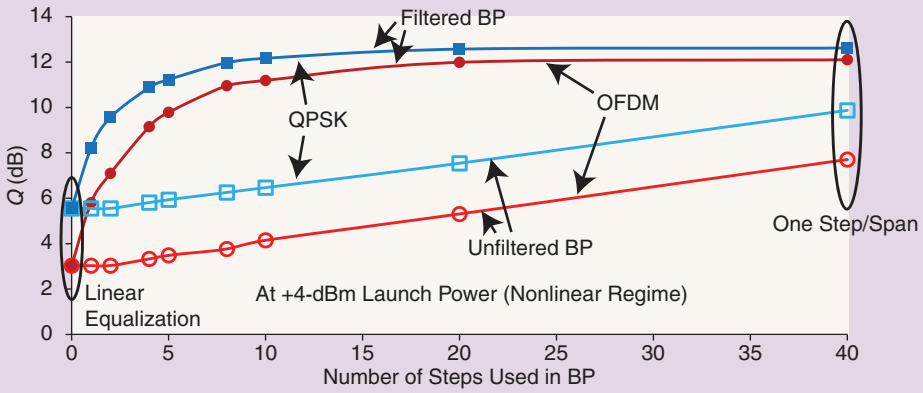
As with single-step methods, the effect of CD can partially be accounted for by suppressing the high-frequency intensity fluctuations with an LPF [6]. In the time domain, this can be understood as taking into account the effect of nonlinear phase shift on a particular symbol, arising from multiple symbols interfering (via dispersion) with that particular symbol. Such an effect can be modeled in DBP by taking a weighted average of the intensity waveform, which is simply an FIR filter [25]. This is commonly referred to as filtered or correlated DBP and enables compensation of multiple physical spans with each DBP step. This approach has been shown to be effective in both single- and multicarrier systems, as shown in Figure 6. For both CO-OFDM and CO-QPSK, four-step filtered DBP provides a similar improvement to 40-step unfiltered DBP for a 40-span NDM link. Therefore, filtered DBP reduces the step-size requirements by a factor of ten, for a similar improvement in performance.

NL COMPENSATION USING PHASE ESTIMATION

Pilot-tone-aided phase noise compensation was first proposed to compensate for intrasymbol laser phase noise in CO-OFDM systems [26]. This was then extended to compensate for fiber nonlinearity [7]. In this method, an unmodulated pilot tone is transmitted; this tone can be in the center of the information band, as shown in Figure 7 [8], or toward one side [27]. At the receiver, this tone is filtered out from the signal and used to capture the phase noise experienced by the signal. This method compensated for all forms of phase noise; laser phase noise is treated identically to fiber nonlinearity. There are two limitations: 1) there must be no signal on either side of the pilot tone to allow the pilot tone to be separated from the signal by filtering and 2) only phase distortions within the bandwidth of the filter can be compensated.



[FIG5] (a) Simulation results for a 10×224-Gb/s POLMUX 16-QAM over a 8×80-km DM link. (b) Experimental results for a 10×111 Gb/s POLMUX-RZ-DQPSK system at 0-dBm launch power into each span [24].



[FIG6] **The Q-factor of a 4-QAM 40×80-km NDM link with filtered and unfiltered DBP [6].**

The ASE close to the pilot tone will also be passed by the filter; this noise will be multiplied with the signal and cause distortions. In the ASE limited regime, the resulting SNR after pilot-based nonlinearity compensation (PB-NLC) is determined by both the SNR of the signal and of the pilot. Therefore, the amount of power allocated to the pilot should be optimized to maximize the SNR after compensation. The optimum power and the attainable SNR are also affected by the bandwidth of the LPF; a wider bandwidth will increase the amount of noise in the filtered pilot. The optimal bandwidth is around ~500 MHz and 15% of total power in a DM link [28]. This will be even lower for an NDM link, due to the phase array effect [4], [19]. PB-NLC is therefore only suitable for the compensation of narrow-bandwidth phase noise. Because SPM effects are typically wider band, PB-NLC is enhanced if SPM is first compensated using the inverse model technique [8].

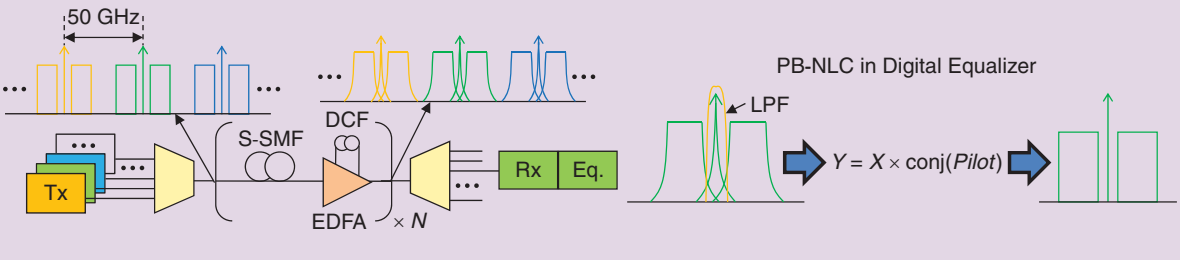
Figure 8 shows the benefit from PB-NLC and DBP [27] for an NDM link. The optimum power of the pilot was 20 dB below the signal with an optimized frequency gap of 24 GHz. The optimal filter bandwidth to extract information from the pilot was ~100 MHz. PB-NLC alone extended reach for both systems slightly; the benefit is enhanced if used after DBP. The broadband phase noise generated by SPM was compensated using DBP. The "cleaned" pilot gives a better estimation of narrowband phase noise generated by laser phase noise and XPM.

In general, all phase compensation methods work for all sources of phase noise, such as from fiber nonlinearity or the laser. Therefore, XPM can also be compensated using blind phase estimation techniques. However, the properties of the phase noises are different; phase noise generated from XPM is mostly between 10–100 MHz. Therefore, the parameters will most likely need to be adjusted to compensate for XPM effectively.

DIGITAL NONLINEARITY COMPENSATION IN NEXT-GENERATION SYSTEMS

FUTURE 1-Tb/s SYSTEMS

Future 1-Tb/s systems will be required to not only increase the capacity of the core optical network but also be more cost and energy efficient per bit. Transmitting more bits per symbol will increase the capacity utilizing similar hardware architecture, thereby lowering the cost per bit. Therefore, gross bits per symbol are likely to increase from four, in POLMUX CO-QPSK systems, to eight or more. However, reach is inversely proportional to the bits per symbol so this cannot be exploited indefinitely [1], [2]. Therefore, a 1-Tb/s signal will occupy around 130–180 GHz of optical bandwidth. This is certainly beyond the bandwidth of a single transmitter and coherent receiver so the signal will comprise multiple lower-rate channels. To maximize the spectral efficiency of the 1-Tb/s superchannel, the subchannels are likely to be squeezed together. Tighter channels will increase interchannel



[FIG7] **An optical system with PB-NLC.**

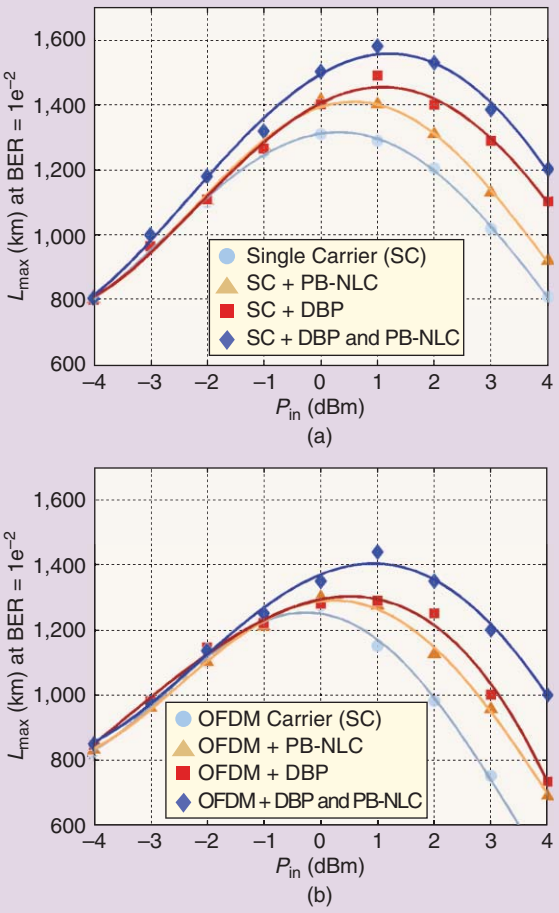
nonlinear distortions, which will also decrease the reach of the system. Therefore, improvements in the link and DSP, such as digital nonlinearity compensation, will be required to achieve a similar reach to current 100 Gb/s systems [2]. However, digital nonlinearity compensation is typically considered to be a complex approach, even at current commercial bit rates (100 Gb/s). Therefore, before discussing the application of nonlinearity mitigation in Tb/s, it is important to establish the required computational requirements for various digital nonlinearity mitigation techniques to give us a better understanding of what is possible.

A KEY CHALLENGE IN REALIZING NONLINEAR COMPENSATORS IS THE CHIP COMPLEXITY—THE NUMBER OF GATES REQUIRED TO IMPLEMENT A GIVEN ALGORITHM.

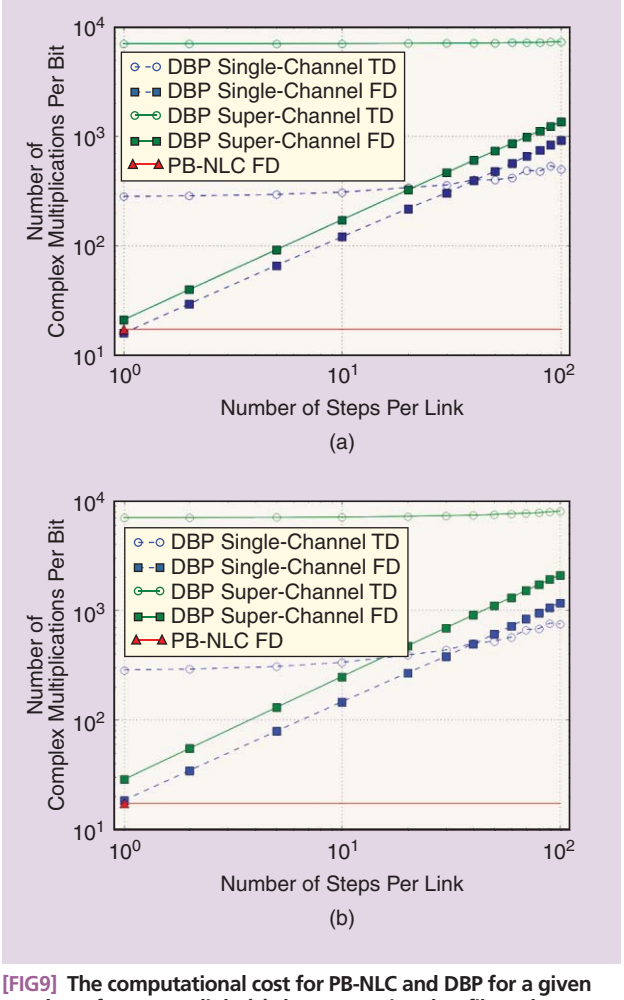
COMPLEXITY OF NONLINEARITY COMPENSATION

Advances in DSP usually mean more complex algorithms and higher computational cost. Hence, a key challenge in realizing nonlinear compensators is the chip complexity—the number of gates required to implement a given algorithm. As an example, we consider a 1-Tb/s superchannel with a 4% Ethernet overhead and 15% FEC overhead, giving a total bit rate of 1.204 Tb/s. This is split evenly between five subchannels, each modulated with DP-16-QAM at 30.1 Gbaud. A root-raised-cosine pulse shape with roll-off 0.2 was used to confine each subchannel to 36.1 GHz, making the superchannel 186.86 GHz. We assume a 10 × 80 km NDM link, using EDFA amplification and S-SMF.

Figure 9 shows the number of complex multiplications per transmitted bit for conventional single-channel DBP, superchannel DBP, and pilot-based nonlinearity compensation. The details on the number of multiplies needed are presented in [29]. For DBP, the linear operator is assumed to be implemented either in the time domain as an FIR filter or in the frequency domain using the overlap-save technique and FFT/IFFTs. Results for one step per link can be interpreted as the complexity of a conventional EDC. For PB-NLC, we assume that the pilot is filtered by a fifth-order Gaussian filter



[FIG8] The maximum reach distance versus launch power for 9 WDM 16-QAM at 224 Gb/s over an NDM link for (a) single carrier and (b) CO-OFDM [27]. The maximum distance was calculated by assuming a soft-decision FEC with a threshold of 1.0×10^{-2}.



[FIG9] The computational cost for PB-NLC and DBP for a given number of steps per link: (a) the conventional unfiltered DBP and (b) filtered digital DBP.

with a bandwidth of 100 MHz, which is implemented in the frequency domain.

Figure 9(a) shows the complexity of conventional unfiltered DBP (see the section "Nonlinear Digital Backward Propagation"). The complexity of the time-domain implementation is essentially independent of the number of steps used for DBP. This is because the length of the impulse response, N_{taps}, required in each section is inversely proportional to the number of steps and therefore the complexity for all steps together is constant. The small increase for high number of steps is caused by quantization effects, since the length of the filter is rounded up to integer values. As mentioned in the section "Nonlinear Digital Backward Propagation," the nonlinear sections are computationally trivial. In contrast, the complexity of the frequency-domain implementation is roughly proportional to the number of DBP steps and is much more efficient than the time-domain implementation for a small number of steps. The break-even between time-domain and frequency-domain solution is somewhere around 40 steps per link for single-channel DBP and far beyond 100 for superchannel DBP.

FOR HIGHER-ORDER CONSTELLATIONS, PB-NLC'S JOINT ABILITY TO COMPENSATE FOR LASER PHASE NOISE ALSO BECOMES MORE VALUABLE BECAUSE BLIND CARRIER PHASE RECOVERY IS MORE DIFFICULT.

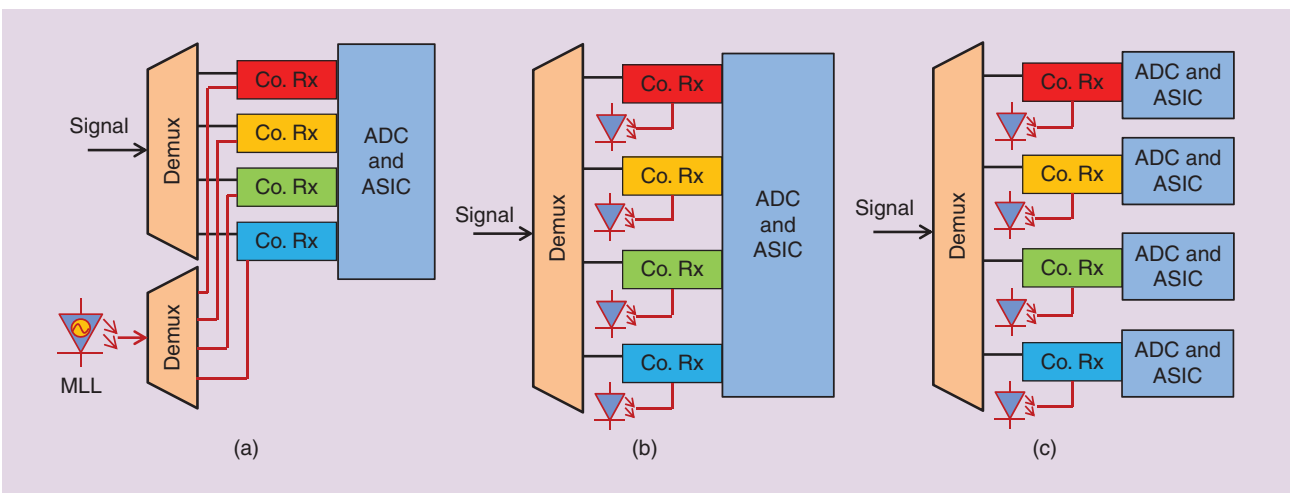
The complexity of PB-NLC is dominated by the filtering operation and therefore requires a similar number of multiplies to DBP with only one step per link. Since PB-NLC is independent of the number of DBP steps, the total complexity of combined DBP and PB-NLC can be determined by linearly adding the constant PB-NLC complexity on top of the DBP complexity for any number of steps.

Figure 9(b) shows the complexity of filtered DBP, where we assumed an additional real-valued FIR filter with ten and 30 taps for each DBP step for single-channel and superchannel DBP, respectively. For convenience, the complexity of PB-NLC is shown in this plot as well. The impact of additional filtering has only a very small impact for a small number of steps and becomes more noticeable as the number of steps increases. However, compared to the overall complexity, the additional increase due to filtering is still moderate.

NONLINEARITY COMPENSATION IN 1-Tb/s SYSTEMS

Even though the complexity of DBP seems excessive from today's perspective, the very fast development in DSP technology means that nonlinear compensation techniques, such as DBP, are likely to be a part of future commercial chip-integrated transponders. The novel system architecture of 1-Tb/s systems will produce new challenges for effective nonlinearity compensation. It has been observed that single-channel DBP reduces in effectiveness if the channel spacing is reduced [30]. This is because interchannel effects, such as XPM, will be stronger. Therefore, the benefit from subchannel DBP will be small in 1-Tb/s systems.

One way of overcoming XPM between the subchannels is to use DBP in the whole superchannel. This would require the optical local oscillators of the multiple receivers be phase locked to reproduce the received superchannel [11], as shown in Figure 10(a). Figure 9 shows superchannel DBP is only slightly higher in complexity than single-channel DBP if implemented in the frequency domain, which is optimum if <40 steps are used. However, the wider signal bandwidth of the superchannel will also mean that shorter steps will be required; multiple steps per span are likely to be required [21]. Filtered DBP may be useful in superchannel DBP to reduce the number of steps. Additionally, superchannel DBP must be performed on a single ASIC. This will pose new challenges for the ASIC design because of the required number of gates on a single chip.



[FIG10] Superchannel receiver architectures for a four subchannel system: (a) with a joint ASIC and phase-locked LOs for superchannel DBP, (b) with a joint ASIC and unlocked LOs for coupled-field DBP, and (c) with independent ASICs and coherent receivers for single-channel DBP. *Mode-locked laser* is represented by the acronym MLL.

Alternatively, the XPM of neighboring subchannels can be compensated using coupled-field DBP [11]. Coupled-field DBP does not require the local oscillators to be phase locked because a separate SSFM is used for each subchannel. However, a large amount of information transfer is required between the DBP hardware of the subchannels. This likely means that the DSP still must be contained in a single ASIC, as shown in Figure 10(b). This method does not compensate for FWM between the subchannels so will have inferior performance compared to superchannel DBP. The lower bandwidth signals will mean fewer steps are required.

Another possibility is to use the PB-NLC to compensate for the XPM between the subcarriers. The additional complexity of PB-NLC is similar to EDC of today's systems, as shown in Figure 9, making it relatively light in complexity. This method will allow individual ASICs to be used for each subchannel and phase-independent local oscillators to be used, as shown in Figure 10(c). This makes implementation of both the optics and the DSP easier. For higher-order constellations, PB-NLC's joint ability to compensate for laser phase noise also becomes more valuable because blind carrier phase recovery is more difficult. PB-NLC removes the possibility of cycle slips and therefore eliminates the need to use differential encoding. This partially offsets the loss in sensitivity. Despite the benefits, PB-NLC is an approximate method and will not deliver the performance of full superchannel DBP. We conjecture that this method will be always used together with others, such as subchannel DBP, since it is ineffective in compensating for SPM.

CONCLUSIONS

In this article, we have reviewed current commercial optical communication systems and discussed promising nonlinearity compensation methods that may help enabling Tb/s transmission over longer distances. The inverse fiber model, often known as DBP, is capable of doubling the data rate. However, the computational power required by DBP of a full WDM band might be prohibitive. We presented the recent advances to reduce the computational effort needed without compromising performance. Another promising method of nonlinearity compensation is to use pilot tones as probes to detect the nonlinear phase distortions experienced by the signal, thereby compensating interchannel effects. Such methods work well in conjunction with DBP. Finally, we provided an outlook for the application of fiber nonlinearity compensation techniques in 1-Tb/s transmission systems, and analyzed the computational effort of various digital nonlinearity mitigation methods, both for conventional single-channel and multichannel nonlinearity compensation.

ACKNOWLEDGMENTS

Liang Du and Arthur Lowery were supported by the Australian Research Council's (ARC's) Discovery Scheme (DP1096782) and the Centre of Ultrahigh Bandwidth Devices for Optical Systems (CUDOS) (CE110001018). Antonio Napoli was partially supported by the FP-7 IDEALIST project (317999) and would like to thank Alberto Diaz, Chien-Yu Lin, Zied Maalej, and Vincent Sleiffer for their help and fruitful discussion. Bernhard Spinnler was supported by the German Federal Ministry of Education and Research (BMBF) (01BP12300A, EUREKA-Project SASER).

AUTHORS

Liang B. Du (liangbydu@gmail.com) received his bachelor of engineering (honors) and commerce degree in 2008 and received his Ph.D. degree in 2012 from Monash University, Australia. His Ph.D. project was on fiber nonlinearity compensation for optical communication systems. He was awarded the Corning Outstanding Student Paper Award at the 2011 Optical Fiber Communication Conference (OFC) and the Best Student Paper Award in transmission systems at the 2012 Optoelectronics and Communications Conference. Since 2012, he has been a research fellow in the Department of Electrical and Computer Systems Engineering at Monash University. His research interests include optical OFDM, equalization techniques for coherent optical systems, and optical nonlinearity compensation techniques.

Danish Rafique (danish.rafique@coriant.com) holds B.Sc. (2007, Foundation for Advancement of Science and Technology–National University, Pakistan), M.Sc. (2009, KTH, Sweden), and Ph.D. (2012, Tyndall National Institute, University College Cork, Ireland) degrees, all in electrical engineering. During his Ph.D. studies, he published more than 34 articles and was a semifinalist in the OFC Best Student Paper Competition (in 2011 and 2012). Currently, he is pursuing a career with R&D Labs of Coriant GmbH, Germany, where he is working on next-generation transponder solutions involving flex-grid networks. His research interests include nonlinear effects in fiber, DSP, and flex-grid optical networks. He is a reviewer for *IEEE Photonics Technology Letters, IEEE Journal of Lightwave Technology*, and Optical Society of America's *Optics Express*.

Antonio Napoli (antonio.napoli@coriant.com) received the master's and Ph.D. degrees in electronics engineering from Politecnico di Torino in 2002 and 2006, respectively. He was a visiting student at the Technische Universität Wien (2001), and a visiting researcher at Universitat Politècnica de Catalunya in Barcelona (2004) and at University College London, United Kingdom (2005). In 2006, he joined Siemens COM, where he worked on EDFA transient suppression and DSP for long-haul optical networks. In 2007, he joined Nokia Siemens Networks, where he worked on optical transmission up to 400 Gb/s. He has been with Coriant since 2013, where is leading a work-package on data plane within the EU FP-7 IDEALIST project. He regularly supervises students and teaches at several universities. His interests are focused on DSP and advanced modulation formats for long-haul systems. He is author or coauthor of more than 60 papers and six patents.

Bernhard Spinnler (bernhard.spinnler@coriant.com) received the Dipl.-Ing. degree in communications engineering and the Dr.-Ing. degree from the University of Erlangen-Nürnberg, Germany, in 1994 and 1997, respectively. Since 1997, he has worked on low-complexity modem design of wireless radio relay systems at Siemens AG, Information and Communication Networks. In

2002, he joined the optical networks group of Siemens Corporate Technology, which later became Nokia Siemens Networks. He has been with Coriant since 2013, where he is working on robust and tolerant design of optical communications systems. His interests focus on advanced modulation and equalization for single and multicarrier systems.

Andrew D. Ellis (andrew.ellis@aston.ac.uk) received the B.Sc. degree in physics with a minor in mathematics from the University of Sussex, Brighton, United Kingdom, in 1987. He received the Ph.D. degree in electronic and electrical engineering from The University of Aston in Birmingham, United Kingdom, in 1997 for his study on all optical networking beyond 10 Gb/s. He previously worked for British Telecom Research Laboratories, Corning Research Centre, and Tyndall National Institute in Cork, Ireland, where he was also a member of the Department of Physics, University College Cork. He is currently a professor of optical communications at Aston University, an adjunct professor of electronic engineering at Dublin City University, and a founder of the Dublin-based start-up Pilot Photonics. He research interests include all optical OFDM and optical and electrical signal processing.

Maxim Kuschnerov (maxim.kuschnerov@coriant.com) graduated from the Technical University of Munich in 2006 and received his Ph.D. degree in electrical engineering from the University of the Federal Armed Forces, Munich, Germany, in 2011. His Ph.D. research was conducted in collaboration with Nokia Siemens Networks (NSN) in Munich focusing on the development of digital signal processing algorithms for optical coherent systems. In 2010, he joined NSN full time, focusing on the development of 100-Gb/s transponders and beyond. He has authored and coauthored more than 100 peer-reviewed papers and conference contributions. In 2009, he received the NSN Quality Award for Innovation. Currently, he leads the system trial work package of the Modegap project, focusing on space division multiplexing over multimode fibers.

Arthur J. Lowery (arthur.lowery@monash.edu) is an ARC Laureate Fellow and professor at Monash University, Australia, and leads Monash Vision Group's development of a bionic eye based on a cortical implant. He is also the science leader of the Tbit/s project of ARC CUDOS. His Laureate Fellowship is in developing novel combinations of electronics and photonics for optical networks. In 1996, he cofounded VPIsystems (now VPIphotonics) and led the development of VPItransmissionMaker and VPIcomponentMaker simulation tools. He is a Fellow of the IEEE and of the Australian Academy of Technological Sciences and Engineering.

REFERENCES

[1] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," *Nature*, vol. 411, pp. 1027–1030, 2001.

[2] D. Rafique, J. Zhao, and A. D. Ellis, "Digital back-propagation for spectrally efficient WDM 112 Gbit/s PM m-ary QAM transmission," *Opt. Express*, vol. 19, pp. 5219–5224, 2011.

[3] G. Goldfarb, M. G. Taylor, and G. Li, "Experimental demonstration of fiber impairment compensation using the split-step finite-impulse-response filtering method," *IEEE Photon. Technol. Lett.*, vol. 20, pp. 1887–1889, 2008.

[4] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightwave Technol.*, vol. 26, pp. 3416–3425, 2008.

[5] W. Shieh, H. Bao, and Y. Tang, "Coherent optical OFDM: Theory and design," *Opt. Express*, vol. 16, pp. 841–859, 2008.

[6] L. B. Du and A. J. Lowery, "Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems," *Opt. Express*, vol. 18, pp. 17075–17088, 2010.

[7] B. Inan, S. Randel, S. L. Jansen, A. Lobato, S. Adhikari, and N. Hanik, "Pilot-tone-based nonlinearity compensation for optical OFDM systems," in *Proc. European Conf. Optical Communication*, Torino, Italy, 2010, p. Tu.4.A.6.

[8] L. B. Du and A. J. Lowery, "Pilot-based cross-phase modulation compensation for coherent optical orthogonal frequency division multiplexing long-haul optical communications systems," *Opt. Lett.*, vol. 36, pp. 1647–1649, 2011.

[9] K. Roberts, M. O'Sullivan, K.-T. Wu, H. Sun, A. Awadalla, D. J. Krause, and C. Laperle, "Performance of dual-polarization QPSK for optical transport systems," *J. Lightwave Technol.*, vol. 27, pp. 3546–3559, 2009.

[10] M. Nazarathy, J. Khurgin, R. Weidenfeld, Y. Meiman, P. Cho, R. Noe, I. Shpantzer, and V. Karagodsky, "Phased-array cancellation of nonlinear FWM in coherent OFDM dispersive multi-span links," *Opt. Express*, vol. 16, pp. 15777–15810, 2008.

[11] E. Ip, "Nonlinear compensation using backpropagation for polarization-multiplexed transmission," *J. Lightwave Technol.*, vol. 28, pp. 939–951, 2010.

[12] D. Rafique and A. D. Ellis, "Various nonlinearity mitigation techniques employing optical and electronic approaches," *IEEE Photon. Technol. Lett.*, vol. 23, pp. 1838–1840, 2011.

[13] X. Chen and W. Shieh, "Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems," *Opt. Express*, vol. 18, pp. 19039–19054, 2010.

[14] C. Paré, A. Villeneuve, P. A. Bélanger, and N. J. Doran, "Compensating for dispersion and the nonlinear Kerr effect without phase conjugation," *Opt. Lett.*, vol. 21, pp. 459–461, 1996.

[15] R. Killey, "Dispersion and nonlinearity compensation using electronic predistortion techniques," in *The IEE Seminar on Optical Fibre Communications and Electronic Signal Processing*. London, U.K.: IEE, 2005, pp. 0_14–2/6.

[16] R. Danish and D. E. Andrew, "The impact of signal-ASE four-wave mixing in coherent transmission systems," in *Proc. Optical Fiber Communication Conf.*, Los Angeles, California, 2011, p. OThO2.

[17] C. Xu and X. Liu, "Postnonlinearity compensation with data-driven phase modulators in phase-shift keying transmission," *Opt. Lett.*, vol. 27, pp. 1619–1621, 2002.

[18] A. J. Lowery, "Fiber nonlinearity pre- and post-compensation for long-haul optical links using OFDM," *Opt. Express*, vol. 15, pp. 12965–12970, 2007.

[19] L. B. Du and A. J. Lowery, "Improved nonlinearity precompensation for long-haul high-data-rate transmission using coherent optical OFDM," *Opt. Express*, vol. 16, pp. 19920–19925, 2008.

[20] L. B. Du and A. J. Lowery, "Practical XPM compensation method for coherent optical OFDM systems," *IEEE Photon. Technol. Lett.*, vol. 22, pp. 320–322, 2010.

[21] L. Zhu and G. Li, "Folded digital backward propagation for dispersion-managed fiber-optic transmission," *Opt. Express*, vol. 19, pp. 5953–5959, 2011.

[22] Z. Maalej, A. Napoli, V. Sleiffer, M. Kuschnerov, B. Spinnler, and N. Hanik, "Reduced complexity back-propagation for optical communication systems," in *Proc. Optical Fiber Communication Conf. Expo./Nat. Fiber Optic Engineers Conf. (OFC/NFOEC)*, 2012, pp. 1–3.

[23] D. S. Millar, S. Makovejs, C. Behrens, S. Hellerbrand, R. I. Killey, P. Bayvel, and S. J. Savory, "Mitigation of fiber nonlinearity using a digital coherent receiver," *IEEE Select. Topics Quantum Electron.*, vol. 16, pp. 1217–1226, 2010.

[24] C. S. Fludger, T. Duthel, D. van den Borne, C. Schulien, E.-D. Schmidt, T. Wuth, J. Geyer, E. De Man, G.-D. Khoe, and H. de Waardt, "Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission," *J. Lightwave Technol.*, vol. 26, pp. 64–72, 2008.

[25] D. Rafique, M. Mussolin, M. Forzati, J. Mårtensson, M. N. Chugtai, and A. D. Ellis, "Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm," *Opt. Express*, vol. 19, pp. 9453–9460, 2011.

[26] S. L. Jansen, I. Morita, T. C. W. Schenk, N. Takeda, and H. Tanaka, "Coherent optical 25.8-Gb/s OFDM transmission over 4160-km SSMF," *J. Lightwave Technol.*, vol. 26, pp. 6–15, 2008.

[27] A. Diaz, A. Napoli, S. Adhikari, Z. Maalej, A. P. Lobato Polo, M. Kuschnerov, and J. Prat, "Analysis of back-propagation and RF pilot-tone based nonlinearity compensation for a 9x224Gb/s POLMUX- 16QAM system," in *Proc. Optical Fibre Communication Conf.*, Los Angeles, California, 2012, p. OTh3C.5.

[28] L. B. Du and A. J. Lowery, "Pilot-based XPM nonlinearity compensator for CO-OFDM systems," *Opt. Express*, vol. 19, pp. B862–B867, 2011.

[29] B. Spinnler, "Equalizer design and complexity for digital coherent receivers," *IEEE Select. Topics Quantum Electron.*, vol. 16, pp. 1180–1192, 2010.

[30] S. J. Savory, G. Gavioli, E. Torrengo, and P. Poggiolini, "Impact of interchannel nonlinearities on a split-step intrachannel nonlinear equalizer," *IEEE Photon. Technol. Lett.*, vol. 22, pp. 673–675, 2010.

[SP]

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

Tsuyoshi Yoshida, Takashi Sugihara, Kazuyuki Ishida, and Takashi Mizuochi

Cycle Slip Compensation with Polarization Block Coding for Coherent Optical Transmission

Two-dimensional phases constellation corresponds to a slip state

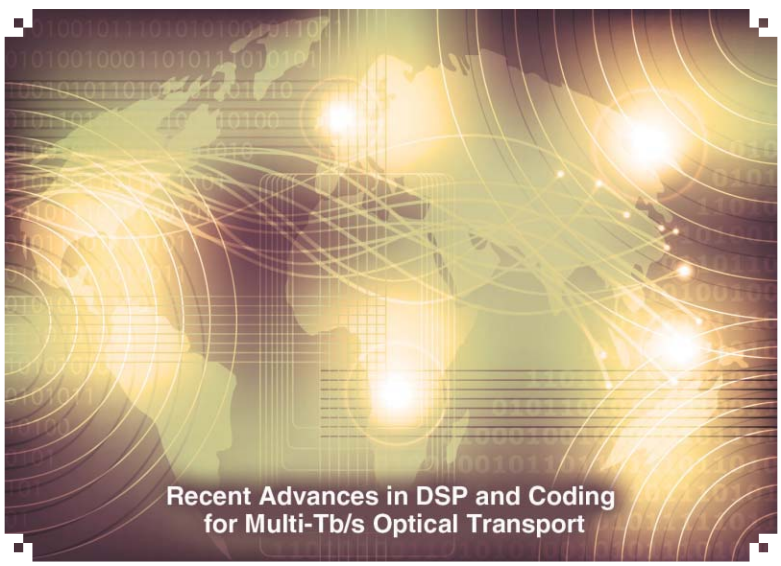


IMAGE LICENSED BY INGRAM PUBLISHING/JANMIKS

Cycle slip (CS) compensation is a critical technique for nondifferential coded coherent optical transmission. By monitoring the sparse, asymmetric polarization block-coded symbol mapped signal phases of two orthogonal polarizations, the CS can be estimated from a relatively short stretch of symbols (unit). Simulation shows that the polarization block coding-based method compensates CS and improves the Q-factor by 1 dB or more compared to differential coding.

Digital Object Identifier 10.1109/MSP.2013.2290326
Date of publication: 12 February 2014

INTRODUCTION

When seeking to raise the capacity of optical transport systems into the multi-Tb/s per channel realm, there are several critical issues: optical signal-to-noise ratio (OSNR), repeater bandwidth, fiber nonlinearity and electrical bandwidth of the devices used. With advances in coding and modulation technology, it is possible to improve the OSNR tolerance and spectral efficiency.

Forward error correction (FEC) is an effective way to improve OSNR tolerance. The latest FECs accept precorrection Q-factors down to around 5 dB and still obtain an output bit-error ratio (BER) of 10^{-15} with a net coding gain (NCG) of 12 dB with more than 25% redundancy at 100 Gb/s per channel or

more [1]. This is only 1 dB away from the soft-decision (SD) Shannon limit. SD low-density parity check (LDPC)-based coding is currently one such basic solution [2], [3].

Coherent detection using digital signal processing (DSP) has become practical in real-life optical communication systems [4]–[6]. The OSNR tolerance has been improved by using coherent detection instead of direct or differential detection. High-order modulation and polarization multiplexing are also able to be used to raise the spectral efficiency, thanks to linear optical/electrical conversion with coherent detection and DSP-based mapping, demapping, and equalization techniques. Spectral shaping in a transmitter DSP, the channel selectivity of coherent detection, and receiver side equalization together enable the channel spacing of wavelength division multiplexing (WDM) to be reduced to around the baud rate, referred to as Nyquist WDM.

BY MONITORING THE SPARSE, ASYMMETRIC POLARIZATION BLOCK-CODED SYMBOL MAPPED SIGNAL PHASES OF TWO ORTHOGONAL POLARIZATIONS, THE CYCLE SLIP CAN BE ESTIMATED FROM A RELATIVELY SHORT STRETCH OF SYMBOLS (UNIT).

Coherent detection has problems with burst errors due to CS [7]–[12]. The transmitted carrier phase is recovered by mixing the received signal light with a local oscillator (LO) light and recovering the carrier by DSP [13], [14]. The recovered carrier phase is ambiguous when the carrier is phase modulated. An M-ary phase-shift keyed (PSK) signal is modulated with M circularly symmetric phases, so its recovered carrier phase has $2\pi/M$ ambiguity. The recovered carrier phase can slip by the angle of ambiguity, which is known as CS. Once CS has occurred, the following data become incorrect, which can result in a very large burst of errors.

Although the most reliable way to reduce error bursts due to CS is to use differential encoding and decoding, this reduces the tolerance to random errors because of error propagation [15], especially at lower Q-factors. Therefore, investigators seek to avoid differential coding by applying various CS compensation (CSC) techniques. Analyzing a periodically inserted pilot symbol is one way to detect and compensate CS [10]–[12]. It is difficult to guarantee an improvement because of the outage rate due to miscompensation of CS and the required rate of pilot insertion, of which the latter increases the line-side bit rate and thus reduces the OSNR tolerance.

Multidimensional coded modulation, e.g., polarization block coding [16]–[19], with its time-block coding and the set-partitioning of higher-level quadrature amplitude modulation (QAM) [20], is being investigated to improve the OSNR tolerance. By combining polarization block coding, which maps the data blocks onto two orthogonally polarized signals, with simultaneous constellation monitoring in two polarizations at the receiver [18], it becomes possible to detect CS more reliably by using not just a pilot signal but also the modulated signal itself.

CYCLE SLIP IN FORWARD ERROR CORRECTION-BASED OPTICAL COMMUNICATIONS

The fundamental cause of CS is a combination of amplified spontaneous emission (ASE) noise, fiber nonlinearity effects, and laser phase noise. The widely used blind carrier phase recovery (CPR) techniques [21] not only find it hard to detect CS but can actually generate it. CS can lead to severe error bursts.

Most optical transport systems do not permit resending so the system has to guarantee a BER of less than 10^{-15} from the start to the end of life with a low rate of outages. A practical optical transport system can suffer from error bursts due to CS. Even if an adequate margin, greater than 3 dB say, in the average pre-FEC Q-factor is ensured, CS can still occur on rare occasions, and the resulting error bursts will be observed over the lifetimes of tens of thousands of wavelengths around the world.

An overview of the conventional methods for burst-error mitigation, differential coding and pilot symbol-aided CSC, is given in the section "Conventional Methods for Burst-Error Mitigation." The motivation for avoiding differential coding is discussed from the point of view of OSNR tolerance. To compare the CSC techniques, their burst-error correction capability is analyzed by introducing a simple model characterized by a number of key parameters (see the section "Analysis of Burst-Error Correction Capability"). The advantage of making the CSC unit small to mitigate frame errors is described.

CONVENTIONAL METHODS FOR BURST-ERROR MITIGATION

DIFFERENTIAL CODING

The first approach to avoiding the error-burst problem was to use differential coding. If recovered-carrier phase slip causes a symbol error, error propagation due to the CS is halted at the next symbol because differential decoding bases the recovery of the data only on the phase difference between pairs of symbols. Although this is robust against error bursts, it is weak in the face of random errors. A single symbol error will end up as two symbol errors due to error propagation. The penalty due to such error propagation depends on the modulation format [15].

For binary PSK (BPSK) and quaternary PSK (QPSK), the BER after differential decoding is almost double that before differential decoding. The relation between BER (P_e) and the corresponding Q-factor (Q_{BER}), given by the complementary error function (erfc)

$$Q_{\mathrm{BER}} = \sqrt{2}\,\mathrm{erfc}^{-1}[2P_e], \tag{1}$$

is highly nonlinear, so the penalty due to the increased OSNR required by the BER-doubling problem with differentially coded BPSK and QPSK is observed strongly at lower Q_{BER} in

the region of 7 dB or lower. The penalty is 1.4 dB at a Q_{BER} of 5 dB, while it is only 0.5 dB at a Q_{BER} of 10 dB.

FEC is now pushing the correctable limit to Q-factors of around 5 dB with an NCG of nearly 12 dB [1]. This issue of doubled BER incurs a significant loss of OSNR margin at such low Q-factors. An FEC with an NCG of 13.7 dB at 37.1% redundancy has been proposed [22], and such highly redundant code is likely to be put to practical use in the future but will suffer even more from the error propagation inherent to differential coding.

Error propagation also causes error bursts, which degrade the FEC's correction capability. Code interleaving can avoid the degradation of FEC correction capability due to such error bursts, but due to the differential coding it prevents reduction of the pre-FEC BER.

PILOT SYMBOL-AIDED CYCLE SLIP COMPENSATION

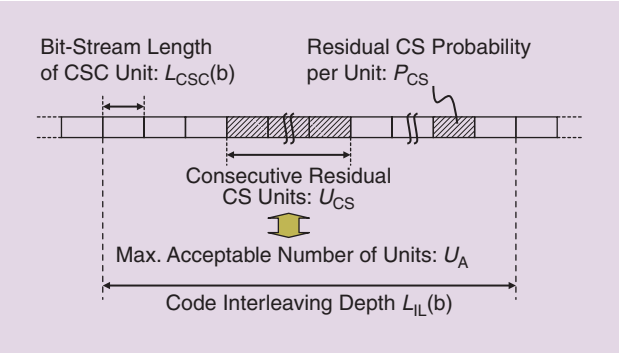
The second solution to the error-burst problem was to send a known pilot signal and to detect the carrier phase directly to avoid the error-propagation problem with differential coding described in the section "Conventional Methods for Burst-Error Mitigation." However, an out-of-band pilot tone requires additional transmission equipment, consumes extra signal power in a WDM system, and is seriously affected by frequency-dependent impairments. Therefore, most recent solutions are based on in-band pilot symbols [10]–[12]. Although this is easier to implement than polarization block coding, there is a tradeoff between accuracy of compensation and line-side bit rate. The pilot symbol insertion ratio is constrained by the need for its penalty to be less than that due to differential coding if it is to maintain an advantage in terms of OSNR tolerance.

ANALYSIS OF BURST-ERROR CORRECTION CAPABILITY

CSC aims to improve OSNR tolerance by avoiding the problems with differential coding (see the section "Conventional Methods for Burst-Error Mitigation"). To aid the discussion of desirable CSC techniques, including the pilot symbol-aided method described in the section "Conventional Methods for Burst-Error Mitigation," a simple analysis of the burst-error correction capability of FEC is given in this section, noting that residual post-FEC errors need to be eliminated.

The tolerance to random errors has mostly been discussed in terms of improving the tolerance to low OSNRs because that is currently the critical limiting factor. On the other hand, FEC performance is affected strongly by bursts of errors. Originally, error bursts had to be eliminated prior to FEC processing so that the FEC only had to correct random errors. Error bursts due to fiber nonlinearity can now be reduced by nonlinearity equalization, e.g., the technique described in [23], but there will still remain a finite probability of error bursts due to CS.

The ability of FEC to correct errors depends on the CS rate and the residual CS angle, e.g., $\pm\pi/2$ or π for QPSK, due to CS in the absence of differential coding. The data unit for CSC consists of a bit-block having a length of L_{CSC} bits long. The length of the unit has to be kept to sufficiently less than the code interleaving depth to protect the FEC frame. The code interleaving depth is itself restricted by the memory capacity and latency of practical large-scale integration (LSI) designs.



[FIG1] A model to estimate ability of FEC to correct burst errors with CSC.

To estimate how the signal parameters affect the ability of FEC to correct errors with CSC, we use the simple model shown in Figure 1. The FEC code is interleaved with a depth of L_{IL} bits. Code interleaving disperses the burst errors by changing their chronological order. The burst-error correction length (BECL) of the FEC is L_{BEC} bits. CS is compensated by an arbitrary technique over a unit of L_{CSC} bits. The probability of CS remaining after CSC is P_{CS}, which is defined as the number of units still with CS to the total number of CSC units. The number of consecutive units with CS after CSC is U_{CS}.

When $L_{CSC}U_{CS}$ is larger than L_{BEC}, the FEC cannot correct all the errors and an FEC frame error occurs. The FEC frame error rate (FER) due to error bursts is defined by the probability of $L_{CSC}U_{CS} > L_{BEC}$, which is the same as that given by $U_{CS} > U_A$, as in

$$P[L_{CSC}U_{CS} > L_{BEC}] = P[U_{CS} > U_A] \tag{2}$$

and has to be lower than the FER acceptable in the system. U_A is given by

$$U_A = \text{int}(L_{BEC}/L_{CSC}) \tag{3}$$

and is the maximum acceptable number of consecutive units with residual CS. When $L_{IL} >> L_{CSC}$, the probability of U_{CS} consecutive units with residual CS $P[U_{CS}]$ and the FER of the FEC are approximated by

$$P[U_{CS}] = P_{CS}^{U_{CS}}, \tag{4}$$

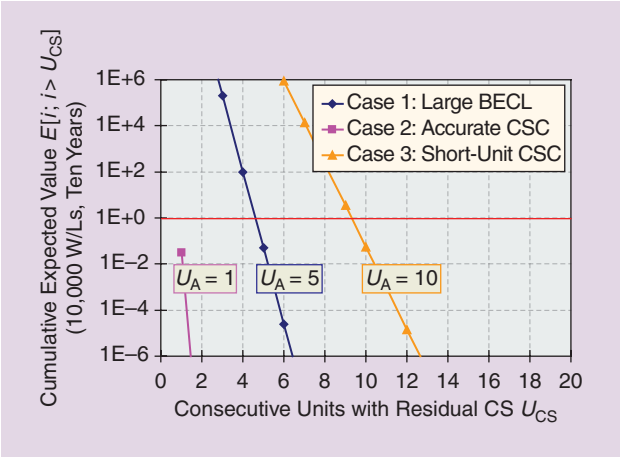
and

$$P[U_{CS} > U_A] = \sum_{i=U_A+1}^{U_{IL}} P_{CS}^{i} \approx P_{CS}^{U_A+1}, \tag{5}$$

where U_{IL} is defined by

$$U_{IL} = \text{int}(L_{IL}/L_{CSC} + 1). \tag{6}$$

From (5) and (6), L_{IL} is negligible because shorter L_{IL} makes L_{BEC} shorter. Therefore, to improve L_{BEC}, it is better to reduce the correlation between errors by making L_{IL} longer.



[FIG2] The cumulative expected value versus consecutive units with residual CS.

Some ways to reduce the residual post-FEC FER are 1) large BECL, 2) accurate CSC, and 3) short-unit CSC. We take, e.g., a code interleaving depth of 1 Mb and a bit rate of 1 Tb/s. We set the other parameters as in the sections "Large Burst-Error Correction Length," "Accurate Cycle Slip Compensation," and "Short-Unit Cycle Slip Compensation." Figure 2 shows the cumulative expected value $\mathrm{E}[i;\ i > U_{\mathrm{CS}}]$ versus the number of consecutive units with residual CS, U_{CS}, over 10,000 wavelengths for ten years for cases 1–3. The parameters are set to reduce the residual post-FEC errors to a negligible level. $\mathrm{E}[i;\ i > U_{\mathrm{A}}]$ is less than 0.1 for all cases.

LARGE BURST-ERROR CORRECTION LENGTH

The parameters set for this example in Figure 2 are $L_{\mathrm{CSC}} = 100$ bits, $L_{\mathrm{BEC}} = 500$ bits, and $P_{\mathrm{CS}} = 5 \times 10^{-4}$. U_{A} is five due to the high BECL. In this case, the requirement for P_{CS} can be relaxed to, e.g., 5×10^{-4}. Pilot symbol-aided CSC is relatively suitable for case 1, but increasing the BECL can impair the tolerance of the FEC to random errors.

ACCURATE CYCLE SLIP COMPENSATION

The parameters set for this example in Figure 2 are $L_{\mathrm{CSC}} = 100$ bits, $L_{\mathrm{BEC}} = 100$ bits, and $P_{\mathrm{CS}} = 1 \times 10^{-10}$. U_{A} is one due to the short BECL, so P_{CS} has to take the very low value of 1×10^{-10}. If U_{A} is zero, the requirement on P_{CS} becomes even more severe. Pilot symbol-aided CSC is relatively suitable for case 2 also, but extreme mitigation of P_{CS} is hard to verify and guarantee.

SHORT-UNIT CYCLE SLIP COMPENSATION

The parameters set for this example in Figure 2 are $L_{\mathrm{CSC}} = 10$ bits, $L_{\mathrm{BEC}} = 100$ bits, and $P_{\mathrm{CS}} = 1.6 \times 10^{-2}$, where U_{A} is ten. By making L_{CSC} small, U_{A} is increased and P_{CS} can take a higher value of 1.6×10^{-2}. One realistic approach to overcoming the error-burst problem is to shorten L_{CSC} with reasonable BECL and P_{CS}. It is difficult for case 3 to be realized with pilot symbol-aided CSC because the pilot insertion raises the line rate and thus reduces the OSNR tolerance.

POLARIZATION BLOCK-CODED MODULATION FOR CYCLE SLIP DETECTION

Short-unit CSC is one way to eliminate residual errors after FEC decoding, as shown in the section "Cycle Slip in Forward Error Correction-Based Optical Communications." Coded modulation [16]–[20], a technique used to improve OSNR tolerance, also enables the detection of CS for a modulated signal. By using both a pilot signal and inspection of the modulated signal, the CSC unit can be shortened. Figure 3 is a two-dimensional (2-D) depiction of the signal constellations using phase in the X and Y polarizations ($\phi_{\mathrm{X}}, \phi_{\mathrm{Y}}$) for several modulation formats, which aids recognition of CS and the CS state. The CS state is the difference between the true carrier phase and the recovered carrier phase before CSC. CSC using constellation monitoring does not detect CS events directly, but instead estimates the CS state symbol by symbol.

In the following sections, we review the principles of CS state estimation by monitoring the 2-D phase constellation of a polarization block-coded signal, and the difference from the case of nonpolarization block-coded signals. Polarization block coding as used in this article is defined as bit-block mapping onto the signal states, amplitude, phase etc., in two polarizations. This is not the same as multiplexing two independent single-polarization signals as there is a relationship of the symbol mapping of the polarization block-coded signal across the polarizations.

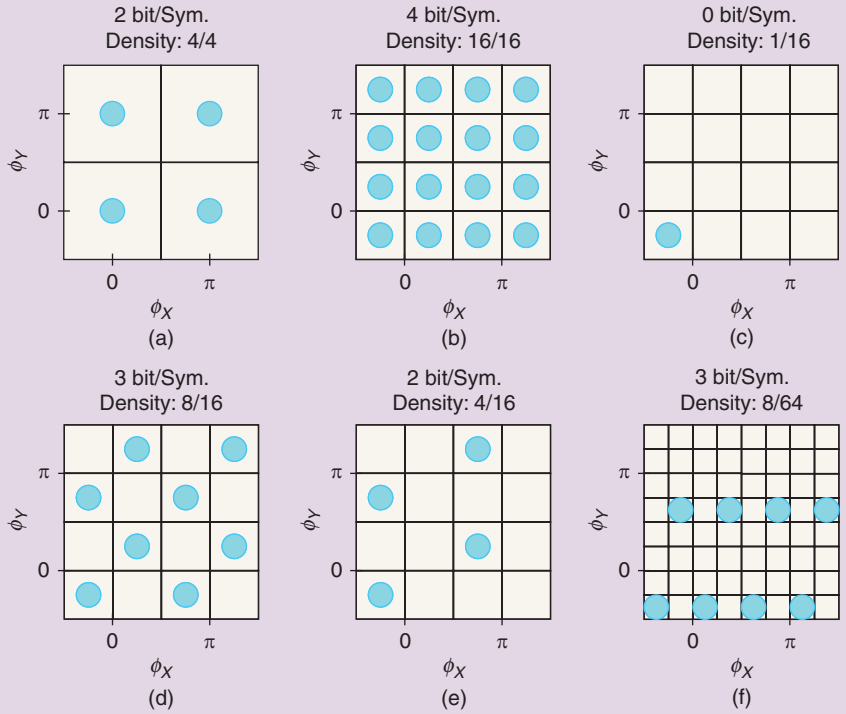
The CS state of a nonpolarization block-coded signal cannot be estimated by constellation monitoring. With a known signal such as a pilot symbol, the CS state can be estimated, but this is not suitable for use with a short-unit CSC. However, the constellation monitoring of a polarization block-coded signal does enable the use of a short-unit CSC.

NONPOLARIZATION BLOCK-CODED SIGNAL

For widely used dual-polarized (DP) signals such as DP-BPSK and DP-QPSK, the two polarizations' coding is independent, and they are just cases of orthogonal-polarization multiplexing. DP-BPSK has two independent spaces in each polarization, and the spaces are occupied by the signal points. DP-QPSK is similar to DP-BPSK, with 16 spaces occupied by 16 signal points. The number of spaces is restricted by the phase count and the power (order) of the CPR used. The carrier phase is recovered by a fourth power method for QPSK and a second-power method for BPSK [21] so the numbers of spaces in each polarization are four for QPSK and two for BPSK. The signal density defined by the number of signal points per space is 100% (4/4) for DP-BPSK and 100% (16/16) for DP-QPSK. Even when CS occurs in a DP-BPSK or DP-QPSK signal, its constellation in the 2-D phase domain does not change. Therefore, we cannot identify how the CS has occurred by looking only at the signal constellation without prior knowledge of the received data. Once CS occurs, the received data can no longer be decoded accurately.

PILOT SIGNAL

The most reliable method for detecting CS is to insert a known symbol stream as a pilot [10]–[12], provided the pilot symbol



[FIG3] **The constellations in the 2-D phase domain for several modulation formats (a) DP-BPSK, (b) DP-QPSK, (c) pilot for DP-QPSK, (d) PS-QPSK, (e) PC-DP-BPSK, and (f) PC-DP-B/QPSK.**

redundancy is acceptable. The pilot signal density is 1/16 for DP-QPSK. The number of possible CS events in each polarization is four, totaling $4 \times 4 = 16$ cases when CPR is conducted in a polarization-independent manner. Therefore all the CS events can be identified by using pilot symbols. This approach is, however, not suitable for short-unit CSC.

POLARIZATION BLOCK-CODED SIGNAL

We show in this section that the sparse and asymmetric coding typically proposed for improving OSNR tolerance is also useful for CS detection. The symbol mapping rule has to be chosen carefully. We cannot use a method that would cause serious degradation of the OSNR tolerance itself, even if it enables the detection of CS. When we restrict the symbol mapping rule, the bit/symbol ratio is reduced and the bandwidth has to be increased to maintain the bit rate. The increase in Euclidean distance should be such as to override the effect of the increase in bandwidth. As examples, polarization switched QPSK (PS-QPSK), polarization block-coded DP-BPSK (PC-DP-BPSK), and polarization block-coded BPSK and QPSK (PC-DP-B/QPSK) are described in the following sections.

PS-QPSK

A PS-QPSK signal [16], [17], [20] has a sparser constellation than DP-QPSK in the 2-D phase domain, and its signal density is 50% (8/16). While 50% of the spaces are filled by signal points, the remaining 50% of the spaces are empty. The sparseness of PS-QPSK has been understood as a key to improved tolerance to poor OSNR. The sparseness also helps with the detection of some cases of CS. If $\pm\pi/2$ CS occurs in the X-polarization but does not occur in the Y-polarization, the constellation is changed. As shown in Figure 3, all the PS-QPSK signal points are moved to empty spaces, which are not signal points allocated by the original symbol mapping rule. The change in the constellation depends on the CS angles in the X-polarization, $\phi_{CS,X}$, and in the Y-polarization, $\phi_{CS,Y}$. While the constellation is the same as the original when mod $[(\phi_{CS,X}, + \phi_{CS,Y}), \pi]$ is zero, the constellation is changed when mod $[(\phi_{CS,X}, + \phi_{CS,Y}), \pi]$ is $\pi/2$. Monitoring the constellation of a PS-QPSK signal enables the detection of some cases of CS, but it cannot enable the detection of cases where (sum of the CS angles of the two polarizations) modulo π is zero.

PC-DP-BPSK

PC-DP-BPSK [18] is suitable for the detection of CS, as explained in this section and again in the section "Example of Cycle Slip Detection and Compensation." The method described in [19] is similar from the point of view of bit-block mapping onto two polarizations with BPSK and QPSK. The signal density is 25% (4/16), so it can be used to identify four cases of CS. Figure 4 shows the recovered 2-D phase constellation for PC-DP-BPSK. CS occurs at angles of π in the X polarization when carrier frequency offset compensation and CPR are performed in a polarization-independent manner by splitting the signal into

orthogonally polarized components. The angles in both polarizations are taken to be integral multiples of $\pi/2$ to enable us to consider the case of joint-polarization CPR (JCPR) [9], [10], [24]. The carrier phase offset $\varphi_{O,P}$ is calculated from

$$\varphi_{O,P}[i] = \frac{1}{N_{PH}}\mathrm{Arg}\left[\sum_{j=i-(N_{CPR}-1)/2}^{i+(N_{CPR}-1)/2}\left(a_{XP}E_X^{N_{PH}}[i+j]+a_{YP}E_Y^{N_{PH}}[i+j]\right)\right]. \quad (7)$$

The suffix P indicates a polarization $P \in \{X, Y\}$ and i is the time index. E_X and E_Y are the electric fields in the X and Y polarizations, respectively. N_{PH} is the number of phases and N_{CPR} is the window size for averaging over time. The coefficients a_{XP} and a_{YP} determine the effects of the X and Y polarizations on polarization P.

The Y polarization has four phases, so CS occurs at multiples of $\pi/2$ in both polarizations for fully JCPR, achieved by setting N_{PH} to four and the coefficients a_{XP} and a_{YP} to 0.5, and applying these to the polarization block-coded signal. Although there are only two phases in the X polarization, fourth-power processing remains applicable because the two phases are subsets of the four phases of QPSK.

Fully joint-polarization fourth-power CPR is essential for PC-DP-BPSK and is a disadvantage compared to DP-BPSK, which can employ second-power CPR after splitting the signal into orthogonal polarizations. For accurate estimation of the phase error, the acceptable phase variation over the averaging window for fourth-power CPR is half that for second-power CPR. To mitigate the effects of phase noise, the averaging window size for CPR needs to be increased for PC-DP-BPSK at lower OSNRs [18].

The CS in the X and Y polarizations takes four states. Labeling the slip states in the X and Y polarizations m and n, the CS angles in the X and Y polarizations are $m\pi/2$ and $n\pi/2$, respectively.

With JCPR, the instances of slip states m and n are reduced. There are four groups:

- $G_0 \in \{00, 11, 22, 33\}$
- $G_1 \in \{10, 21, 32, 03\}$
- $G_2 \in \{20, 31, 02, 13\}$
- $G_3 \in \{30, 01, 12, 23\}$.

The group seen depends on the initial phase difference between the X and Y polarizations. Once the group is fixed, the CS takes only four states in the case of PC-DP-BPSK.

The 2-D phase constellations of all the states of Group G_0 are highlighted with bold borders in Figure 4. We define the CS states $(0, 0), (\pi/2, \pi/2), (\pi, \pi)$, and $(-\pi/2, -\pi/2)$ as States 0, 1, 2, and 3. The slip state can be identified by monitoring the recovered 2-D phase constellation because all four constellations are unique and all the 16 signal points are exclusive in the 2-D phase domain, e.g., if the received signal takes the signal points $(\pi/4, \pi/4)$, $(-3\pi/4, 3\pi/4)$, $(\pi/4, -3\pi/4)$, $(-3\pi/4, -\pi/4)$, the slip state is State 0. This is due to the sparse, asymmetric symbol mapping so it is not easy to extend this approach to a signal having a crowded constellation such as a QPSK-based signal.

PC-DP-B/QPSK

As an example of a 3 bits/symbol format, Figure 3 also shows a hybrid of PC-DP-B/QPSK. The method described in [19] is similar in having 3 bits/symbol and bit-block mapping onto two polarizations, but the apparent numbers of phases are different. The larger the constellation, the more careful the design of the symbol mapping needs to be to maintain both asymmetry for CS detection and a good Euclidean distance for OSNR tolerance. By setting the power in the Y polarization to half the power in the X polarization, the Euclidean distance is balanced across the respective electric fields.

The X and Y polarization have eight and two phases, respectively. By fully applying eighth-power JCPR, CS can be detected as well as for PC-DP-BPSK, except for slips of π because of the π rotational symmetric mapping. PC-DP-B/QPSK does not suffice to enable detection of all CSs by itself. However, it is possible to destroy the π rotational symmetry and detect π slip by switching it with another asymmetric symbol mapping rule.

EXAMPLE OF CYCLE SLIP DETECTION AND COMPENSATION

We describe an example of a CS detection and compensation technique [18] in this section. The key features are

- sparse, asymmetric polarization block-coded symbol mapping
- joint-polarization carrier frequency and phase recovery
- predecision and cost-function generation
- CS state estimation (SSE)
- symbol selection based on the estimated slip state.

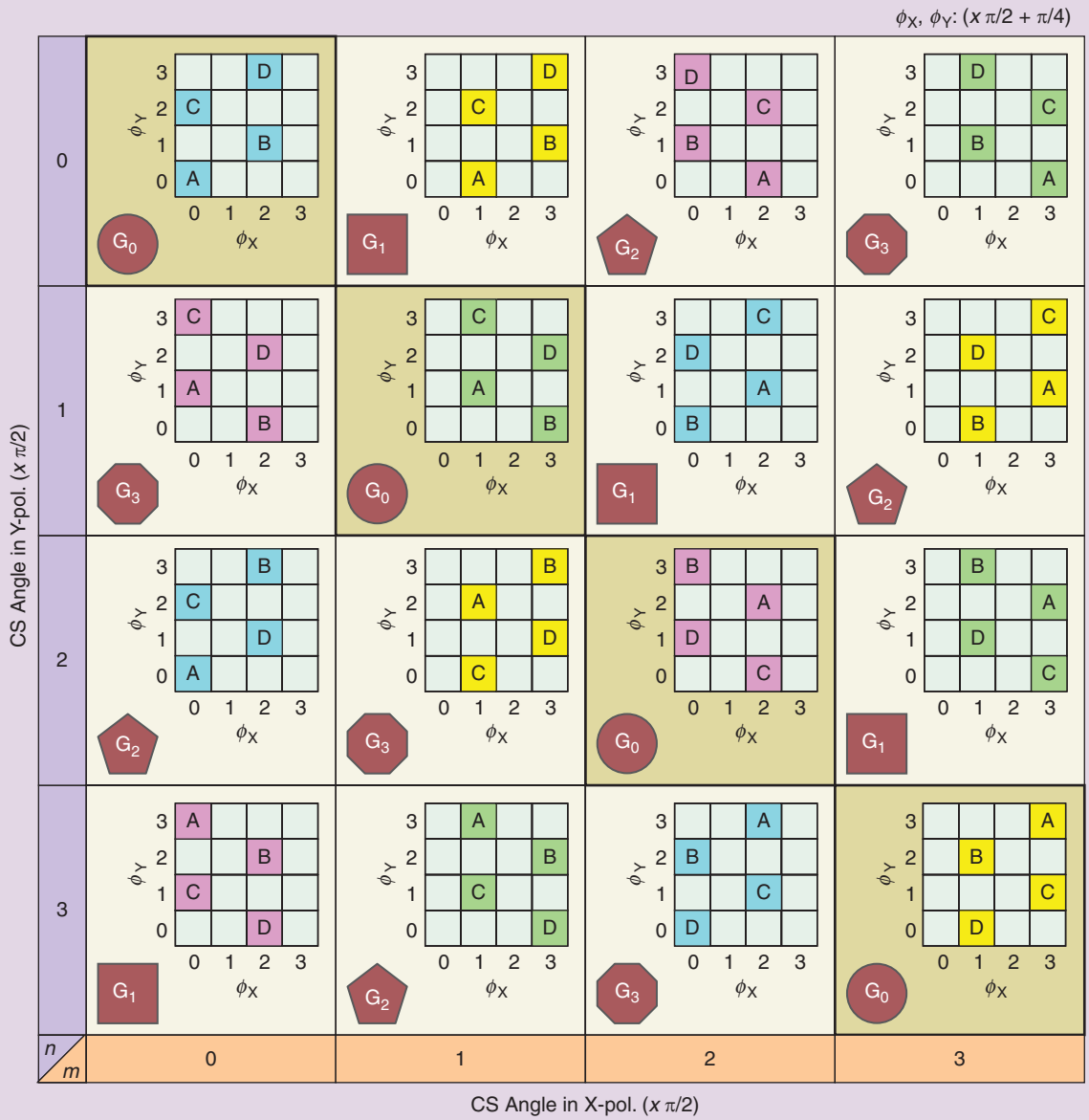
Details of these functions are given in the following sections. The numbers of elements required for the look-up tables (LUTs) N_{LUT}, real multipliers N_{MUL}, real adders and subtracters N_{ADD}, and real comparators N_{CMP} are summed to estimate the ability of this approach to be implemented.

SPARSE ASYMMETRIC POLARIZATION BLOCK-CODED SYMBOL MAPPING

Table 1 shows the symbol mapping rule for PC-DP-BPSK. To realize the sparse, asymmetric constellation of PC-DP-BPSK in the 2-D phase domain shown in Figure 3, four-lane (XI, XQ, YI, and YQ) data is generated from the two-lane input data d_0 and d_1. The optical phases of the X and Y polarizations ϕ_X and ϕ_Y are modulated by the four-lane data d_{XI}, d_{XQ}, d_{YI}, and d_{YQ} using, e.g., a polarization multiplexed I/Q modulator [25]. To implement this function, $N_{LUT} = 1$ is necessary.

JOINT-POLARIZATION CARRIER FREQUENCY AND PHASE RECOVERY

Carrier frequency offset compensation and CPR are performed in a polarization-independent manner by splitting the signal into orthogonally polarized components. The frequency offset between the carrier and local oscillator Δf_{avg} and the phase offset φ_{avg} are calculated in a fully joint-polarization manner. Both the X and Y polarization signals are compensated for frequency offset and phase offset based on common values of Δf_{avg} and φ_{avg}.



[FIG4] The recovered 2-D phase constellations for PC-DP-BPSK with sparse, asymmetric polarization block-coded symbol mapping.

These are averaged over the X and Y polarizations and time. The frequency offsets in the X and Y polarizations are estimated by a suitable algorithm, e.g., that given in [26], and averaged to obtain Δf_{avg}. From (7), the carrier phase φ_{avg} is given by

$$\varphi_{\text{avg}}[i] = \frac{1}{4}\text{Arg}\left[\sum_{j=i-(N_{\text{CPR}}-1)/2}^{i+(N_{\text{CPR}}-1)/2}\left(E_X^4[i+j]+E_Y^4[i+j]\right)/2\right]. \quad (8)$$

The JCPR function outputs the signal in polar coordinates to simplify the subsequent processing (see the section "Predecision and Cost-Function Generation").

To implement JCPR for PC-DP-BPSK, $N_{\text{MUL}} = 6$ and $N_{\text{ADD}} = 4$ for a fourth-power input signal $E_P[i]$, $N_{\text{ADD}} = 2$ $(N_{\text{CPR}} - 1)$ to sum ${E_P}^4[i]$, and $N_{\text{LUT}} = 1$ to calculate $\text{Arg}\{E_P[i]\}$ are required in each polarization. $N_{\text{ADD}} = 2$ to sum ${E_X}^4[i]$ and ${E_Y}^4[i]$ and $N_{\text{LUT}} = 1$ to calculate $\varphi_{\text{avg}}[i]$ are required. $N_{\text{ADD}} = 1$ to subtract $\text{Arg}\{E_P[i]\}$ and $\varphi_{\text{avg}}[i]$ are required in each polarization. In total, $N_{\text{LUT}} = 3$, $N_{\text{MUL}} = 12$, and $N_{\text{ADD}} = 4\ N_{\text{CPR}} + 8$. To

[TABLE 1] SYMBOL MAPPING RULE FOR PC-DP-BPSK.

SYMBOL	INPUT DATA		OPTICAL PHASE		OUTPUT DATA IN EACH LANE			
	d_0	d_1	ϕ_X	ϕ_Y	d_{XI}	d_{XQ}	d_{YI}	d_{YQ}
A	0	0	$\pi/4$	$\pi/4$	0	0	0	0
B	1	0	$-3\pi/4$	$3\pi/4$	1	1	1	0
C	0	1	$\pi/4$	$-3\pi/4$	0	0	1	1
D	1	1	$-3\pi/4$	$-\pi/4$	1	1	0	1

implement CPR for DP-DEBPSK, $N_{\mathrm{MUL}} = 3$ and $N_{\mathrm{ADD}} = 2$ for a second-power input signal $E_P[i]$, $N_{\mathrm{ADD}} = 2(N_{\mathrm{CPR}} - 1)$ to sum $E_P{}^2[i]$, $N_{\mathrm{LUT}} = 2$ to calculate $\mathrm{Arg}\{E_P[i]\}$, and argument of sum of $E_P{}^2[i]$, and $N_{\mathrm{ADD}} = 1$ to subtract $\mathrm{Arg}\{E_P[i]\}$ and the estimated carrier phase are required in each polarization. In total, $N_{\mathrm{LUT}} = 4$, $N_{\mathrm{MUL}} = 6$, $N_{\mathrm{ADD}} = 4$ $N_{\mathrm{CPR}} + 2$ are required. The calculation complexity for PC-DP-BPSK is higher than for DP-DEBPSK because PC-DP-BPSK uses fourth-power processing and DP-DEBPSK uses second-power processing.

This joint-polarization processing mitigates any random noise of the relative phase across the polarizations. To employ joint-polarization frequency offset compensation and CPR, the input signal phases in the X and Y polarizations should be synchronized. The requirement for adaptive equalization by butterfly impulse response filters, which means signal processing before frequency offset compensation and CPR, is described in [24]. When a least-mean square (LMS) algorithm is used [27], the phases after adaptive equalization are synchronized. When a constant modulus algorithm (CMA) is used, feedback of the estimated carrier phase is required because the phases in the two polarizations are not synchronized.

PREDECISION AND COST-FUNCTION GENERATION

Figure 5 shows an example of the circuitry for CSC. First, predecision and cost-function generation are performed for all the possible slip states. By applying fully JCPR, CS by the same angle is generated simultaneously over both the X and Y polarizations. The phase distances are calculated in each slip state from the symbols, where Symbol = {A, B, C, D}, as $D_{\mathrm{PH,X}}$ [State][Symbol] in the X polarization, and as $D_{\mathrm{PH,Y}}$ [State][Symbol] in the Y polarization. The signal points corresponding to symbols A, B, C, and D depend on the slip state. The phase distance in each polarization is summed as the cost function

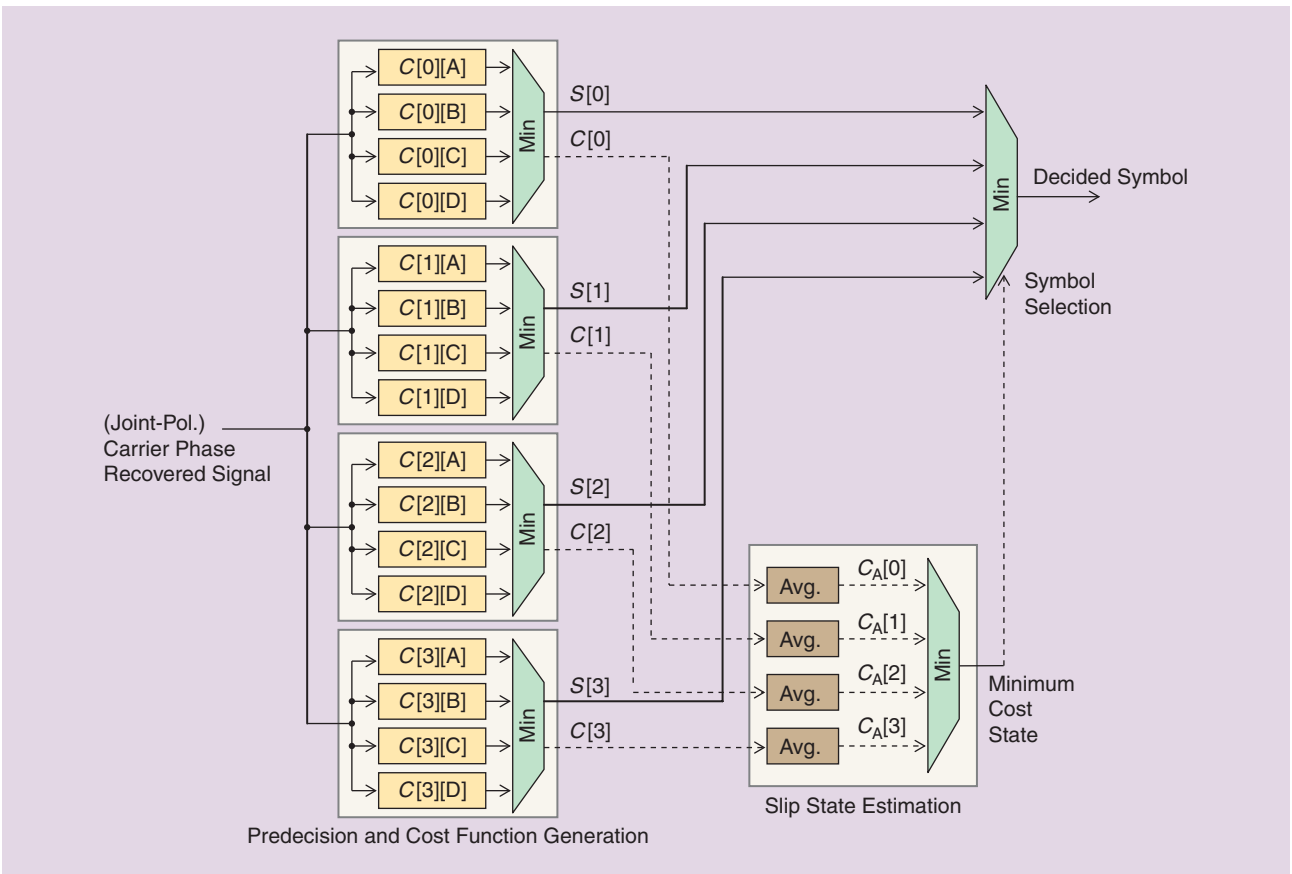
$$C[\mathrm{State}][\mathrm{Symbol}] = D^k_{\mathrm{PH,X}}[\mathrm{State}][\mathrm{Symbol}] + D^k_{\mathrm{PH,Y}}[\mathrm{State}][\mathrm{Symbol}], \tag{9}$$

where k is a power factor to weight the phase distance. The minimum-cost symbol for each state, S[State], is selected and sent to the symbol selection circuit for the operation described in the section "Symbol Selection Based on the Estimated Slip State." The minimum cost in each state

$$C[\mathrm{State}] = \underset{S=\{A,B,C,D\}}{\mathrm{Min}}\{C[\mathrm{State}][\mathrm{Symbol}]\}, \tag{10}$$

is sent to the SSE circuit in Figure 5 whose processing is described in the section "Slip State Estimation" as a single-shot cost function.

To implement this function, $N_{\mathrm{ADD}} = 4$ and $N_{\mathrm{CMP}} = 4$ to calculate $D_{\mathrm{PH},P}$, the phase distance from 4 phases $(i\pi/2 + \pi/4)$, $N_{\mathrm{MUL}} = 4(k-1)$ to calculate $D_{\mathrm{PH},P}{}^k$ in each polarization P,



[FIG5] **The circuitry for CSC including predecision and cost-function generation, slip state estimation, and symbol selection.**

$N_{ADD} = 4$ to calculate the cost function C[State][Symbol], and $N_{CMP} = 3$ to select the minimum cost function in each state are required. In total, $N_{MUL} = 8(k-1)$, $N_{ADD} = 24$, and $N_{CMP} = 20$ are required.

SLIP STATE ESTIMATION

In the SSE circuit in Figure 5, the single-shot cost function in each state C[State] comes from a predecision and cost-function generation circuit as described in the section "Predecision and Cost-Function Generation." The single-shot cost functions are not reliable due to noise, so they are averaged as C_A[State] in the time domain, e.g., by a moving average with a window size of N_{SSE}. Either a finite or an infinite impulse response filter is suitable for windowing. The minimum-cost state is selected as the estimated slip state and the state information is sent to the decision circuit described in the next section.

RAPIDLY VARYING COMPONENTS OF THE POLARIZATION CROSSTALK OR OPTICAL PHASE DUE TO FIBER NONLINEARITY CAN REMAIN AFTER POLARIZATION DEMULTIPLEXING USING BUTTERFLY FINITE IMPULSE RESPONSE ADAPTIVE FILTERS AND FREQUENCY OFFSET COMPENSATION.

To implement this function, $N_{ADD} = N_{SSE} - 1$ to calculate C_A [State] in each state, $N_{CMP} = 3$ to select the minimum-cost state, and in total $N_{ADD} = 4\,N_{SSE} - 4$ and $N_{CMP} = 3$ are required.

SYMBOL SELECTION BASED ON THE ESTIMATED SLIP STATE

In the symbol selection circuit in Figure 5, a symbol is selected and output as decided data from among the possible symbols corresponding to all the incoming slip states S[State] from the predecision and cost-function generation circuit described in the section "Predecision and Cost Function Generation," based on the incoming State from the SSE circuit described in the section "Slip State Estimation." The slip state can in principle be updated for every symbol. However, if implementation is difficult due to the complexity of the calculation, the slip state update period can be lengthened. Only a selector is required for this function.

SIMULATION

A numerical simulation was conducted to verify the CSC capability provided by the PC-DP-BPSK discussed in the section "Example of Cycle Slip Detection and Compensation" when used as a polarization block-coded signal. We simulated for tolerance to OSNR, polarization crosstalk, and phase variation, and then compared the results with those for differential coding.

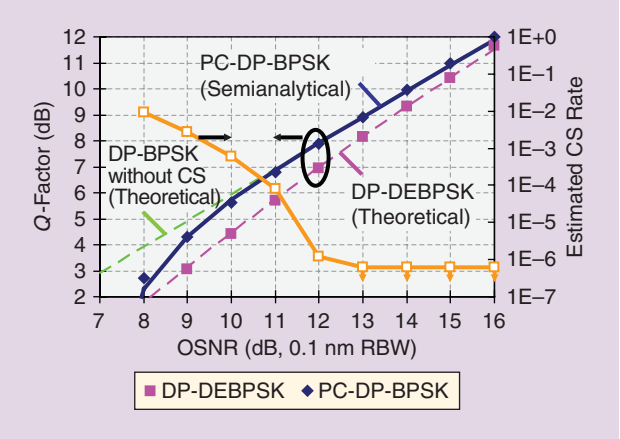
The bit rate was 128 Gb/s. The total number of symbols used was 1.64×10^6 and the signal pattern was a repeated $2^{11}-1$ pseudorandom binary sequence. The BER was calculated from the error count by Monte Carlo simulation and converted to Q-factor in (1). Three channel models were used: additive white Gaussian noise (AWGN) from ASE noise, ASE noise and polarization crosstalk, and ASE noise and phase variation.

The ASE noise resolution bandwidth defining the OSNR was set to 0.1 nm. The ASE noise was assumed to be AWGN and was generated by the Box–Muller method from random numbers [28]. The laser line width and the frequency offset between the transmitter and the LO were set to zero. Optical-electrical conversion was assumed to be linear, and quantization noise was neglected. Polarization multiplexing and demultiplexing were considered ideal.

Rapidly varying components of the polarization crosstalk or optical phase due to fiber nonlinearity, such as cross phase modulation and cross polarization modulation [29], can remain after polarization demultiplexing using butterfly finite impulse response adaptive filters [27] and frequency offset compensation [26]. Therefore, the tolerances to residual polarization crosstalk and phase variation are simulated numerically in the sections "Polarization Crosstalk Tolerance" and "Phase Variation Tolerance."

As a reference, the characteristics of DP-BPSK without CS, shown in Figure 6 as "DP-BPSK without CS (Theor.)" and DP-differential encoded (DE) BPSK, shown as "DP-DEBPSK (Theor.)," which are respectively the calculated theoretical limit and numerically simulated DP-DEBPSK, were also calculated. CPR was performed jointly for the orthogonal polarizations as shown in the section "Joint-Polarization Carrier Frequency and Phase Recovery." The power factor k in (6) was set to two. CPR for DP-DEBPSK is calculated by second-power CPR independently of polarization after splitting the signal into orthogonal polarizations. The window size N_{CPR} for averaging CPR is set to 33 symbols. The averaging for SSE used a moving average, whose window size N_{SSE} was also set to 33 symbols.

CS was not generated directly, instead coming from the simulated ASE noise, fast polarization crosstalk, fast phase variation, and JCPR used in this simulation. The CS events were counted by monitoring the changes in the estimated CS state for PC-DP-BPSK. The CS count was divided by the total number of symbols



[FIG6] **The simulated Q-factor and estimated CS rate versus OSNR. Dashed lines: theoretical; solid blue line: semianalytical; and plotted points: numerically simulated characteristics.**

to calculate the CS rate, because the CS was being compensated symbol by symbol. The minimum observable CS rate is 6×10^{-7}, which is the value due to a single CS in the simulation.

The PC-DP-BPSK performance with CSC, calculated by a simple semianalytical approach, is shown as "PC-DP-BPSK (Semianalytical)" in Figure 6. The BER after CSC was calculated from

$$P_e = P_{e,\text{th}} + eP_{\text{CS}}. \tag{11}$$

$P_{e,\text{th}}$ is the BER of DP-BPSK without CS, P_{CS} is the CS rate, and e is the expected number of bit-errors in a single CS event. The parameter e was set to four by the AWGN fitting process described in the section "OSNR Dependence."

OSNR DEPENDENCE

The OSNR dependence was simulated numerically. Figure 6 shows the Q-factor and estimated CS rate for PC-DP-BPSK as a function of OSNR. DP-DEBPSK is the lower limit to be exceeded and DP-BPSK without CS is the ideal and upper limit. PC-DP-BPSK degrades more below 6-dB Q-factor than does DP-BPSK without CS due to the residual CS after CSC. On the other hand, compared to DP-DEBPSK, PC-DP-BPSK has an advantage of more than 1 dB over DP-DEBPSK at Q-factors lower than 7 dB. The performance of PC-DP-BPSK depends on the OSNR and on the N_{CPR} and N_{SSE} window size settings. At lower OSNRs, N_{CPR} and N_{SSE} should be lengthened to reduce the difference from DP-BPSK without CS [18]. The estimated CS rate is reduced by increasing the OSNR. By keeping the estimated CS rate below 10^{-4}, the degradation from the theoretical limit is lower than 0.2 dB in the case of an AWGN channel. By using (11), a fit to the semianalytical result was obtained by setting the expected number e of bit-errors in a single CS event to four in the simulation. Once the CS rate is characterized, the performance of PC-DP-BPSK with CSC can be predicted by this approach.

THE MANAGEMENT OF CYCLE SLIP IS ESSENTIAL FOR COHERENT OPTICAL COMMUNICATION SYSTEMS TO ACHIEVE MULTI-Tb/s PER CHANNEL OPTICAL TRANSPORT.

We have compared the CS rates of the DP-QPSK experiment reported in [8] with PC-DP-BPSK. In [8], the CS rate was between 10^{-4} and 10^{-5} for 100 Gb/s DP-QPSK with 16 dB OSNR. JCPR used a fourth-power method, and the number of symbols averaged was six, which means $N_{\text{CPR}} = 3$ for both polarizations. A CS rate for PC-DP-BPSK of 5×10^{-5} was obtained by numerical simulation using fourth-power JCPR with $N_{\text{CPR}} = 3$ and 16-dB OSNR. The experimental CS rate for fourth-power JCPR was matched precisely by the numerical simulation.

POLARIZATION CROSSTALK TOLERANCE

The polarization crosstalk was simulated numerically. ASE noise was also taken into account. The OSNR was set to 12 dB in 0.1-nm resolution bandwidth. N_{CPR} and N_{SSE} were set to the same value of 33 symbols. The polarization crosstalk was simulated using the Jones matrix $J[\theta]$

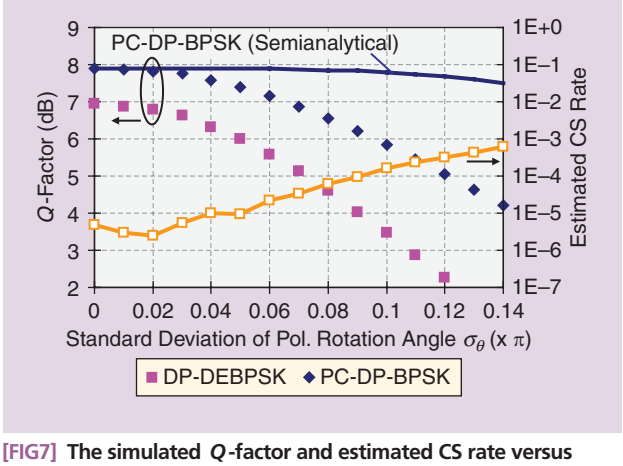
$$J[\theta] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{12}$$

The polarization rotation angle θ was represented as random white Gaussian noise. Its mean was zero and its standard deviation σ_θ was variable. When the optical field is described as $\mathbf{E} = [E_{\text{X}} \; E_{\text{Y}}]^{\text{T}}$, where E_{X} and E_{Y} are the signal magnitudes in the X and Y polarizations, the polarization crosstalk input/output function is

$$\mathbf{E}_{\text{OUT}} = J[\theta]\mathbf{E}_{\text{IN}}. \tag{13}$$

This polarization crosstalk was assumed to remain after DSP-based polarization demultiplexing. DP-DEBPSK was also simulated to provide a reference. We calculated the polarization crosstalk by changing the difference in carrier phase (ψ) between the X and Y polarizations at which crosstalk occurs. The performance of DP-DEBPSK is worse near $\psi = 0$ because the crosstalk causes in-phase interference, and it is better near $\psi = \pi/2$ because the crosstalk causes phase-quadrature interference. On the other hand, for PC-DP-BPSK, half of the crosstalk is in-phase and half is in-phase quadrature in both the $\psi = 0$ and $\psi = \pi/2$ cases due to the asymmetric symbol mapping. Therefore, the degradation due to polarization crosstalk is less for PC-DP-BPSK than for DP-DEBPSK.

Figure 7 shows the Q-factor and estimated CS rate for PC-DP-BPSK as a function of the standard deviation of the polarization rotation angle σ_θ for $\psi = 0$, which was the worst case for both DP-DEBPSK and PC-DP-BPSK. With no polarization crosstalk ($\sigma_\theta = 0$), PC-DP-BPSK has a 0.9 dB advantage over DP-DEBPSK. When $\sigma_\theta = 0.1\pi$, the advantage increases to 2.5 dB regardless of the higher estimated CS rate for PC-DP-BPSK. Although PC-DP-BPSK depends on joint-polarization processing of the symbol mapping, carrier recovery, and CSC, its tolerance to polarization



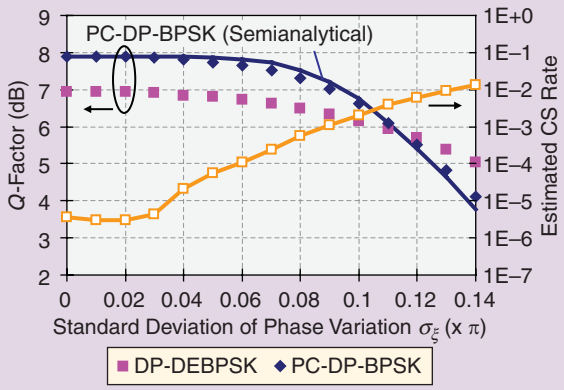
[FIG7] The simulated Q-factor and estimated CS rate versus polarization crosstalk. Solid line: semianalytical, plotted points: numerically simulated characteristics.

crosstalk is superior to that of DP-DEBPSK. CS is not the main cause of the Q-factor degradation in this case because the results from the semianalytical calculation were far higher than the numerically simulated ones.

PHASE VARIATION TOLERANCE

The phase variation was simulated numerically. ASE noise was also taken into account. The OSNR was set to 12 dB in 0.1-nm resolution bandwidth. N_{CPR} and N_{SSE} were set to the same value of 33 symbols. The phase rotation angles in the X and Y polarizations, ξ_X and ξ_Y, were represented as independent and identically distributed random white Gaussian noise. Their means were zero and their standard deviations were a common variable σ_ξ. The phase variation input/output function is described as

$$\mathbf{E}_{OUT} = \begin{bmatrix} \exp[j\xi_X] & 0 \\ 0 & \exp[j\xi_Y] \end{bmatrix} \mathbf{E}_{IN}. \tag{14}$$



[FIG8] The simulated Q-factor and estimated CS rate versus phase variation. Solid line: semianalytical; plotted points: numerically simulated characteristics.

Figure 8 shows the Q-factor and estimated CS rate for PC-DP-BPSK as a function of standard deviation of phase rotation angle σ_ξ. With no phase variation ($\sigma_\xi = 0$), PC-DP-BPSK shows 0.9 dB advantage over DP-DEBPSK. The advantage decreases with increasing phase variation and reaches zero at $\sigma_\xi = 0.11\pi$ and an estimated CS rate of 4×10^{-3}. This is because PC-DP-BPSK requires coherence of the relative carrier phase between the X and Y polarizations, which is degraded by differential phase rotation of the two polarizations.

Severe phase variation is usually caused by fiber nonlinearity so it can be mitigated with a nonlinearity equalization technique, e.g., that described in [23], which should be applied before using constellation monitoring-based CSC. By holding σ_ξ to $< 0.11\pi$ and the estimated CS rate to less than 4×10^{-3}, we can enjoy the Q-factor advantage of PC-DP-BPSK. CS is the main cause of the Q-factor degradation here because the results from the semianalytical calculation, which consider only AWGN and CS, agree with the numerically simulated ones.

Inadequate CPR window size, fiber nonlinearity, especially in a chromatic dispersion managed link, and its imperfect equalization can cause phase noise. PC-DP-BPSK performs well with AWGN, but it is degraded by any increase in phase noise. Therefore, balancing the AWGN and phase noise is key to taking advantage of PC-DP-BPSK.

In Figure 8, the maximum advantage of PC-DP-BPSK at 12 dB OSNR is $\sigma_\xi < 0.11\pi$. The in-phase and quadrature-phase components of AWGN are $\sigma_I^2 = \sigma_Q^2 = 0.162$ when the average single-polarization power is normalized to one. The phase noise power in the I/Q field can be approximated to $\sigma_{PN}^2 = (0.11\pi)^2 = 0.119$, by applying a first order Taylor expansion $(\sin(x) \sim x)$. The ratio of phase noise to total noise (PNR) is calculated as $\sigma_{PN}^2/(\sigma_I^2 + \sigma_Q^2 + \sigma_{PN}^2) = 0.43$. In this case, a noise balanced PNR $= 0.43$ is the limit of applicability of PC-DP-BPSK.

DSP CALCULATION COMPLEXITY

The complexity of the calculations undertaken in the DSP is discussed in this section to evaluate the possibility of implementing PC-DP-BPSK in a practical DSP-LSI. Table 2 compares its calculation complexity with DP-DEBPSK for each DSP block. The numbers of elements for the LUTs N_{LUT}, the real multipliers N_{MUL}, the real adders and subtracters N_{ADD}, and the real comparators N_{CMP} are counted to enumerate the DSP blocks.

[TABLE 2] A COMPARISON OF PC-DP-BPSK AND DP-DEBPSK DSP CALCULATION COMPLEXITIES.

DSP BLOCK	DP-DEBPSK	PC-DP-BPSK	COMPARISON
SYMBOL MAPPING	TWO DIFFERENTIAL ENCODERS	$N_{LUT} = 1$ (FOUR ADDRESSES)	LESS COMPLEX FOR PC-DP-BPSK
ADAPTIVE EQUALIZATION	LMS OR CMA	LMS OR CMA	NOT SIGNIFICANTLY DIFFERENT
FREQUENCY OFFSET COMPENSATION	CALCULATION IN EACH POL. INDEPENDENTLY	ONE ADDITIONAL ADDER IS REQUIRED FOR AVERAGING OVER POLARIZATION	NOT SIGNIFICANTLY DIFFERENT
CPR	$N_{LUT} = 4$ $N_{MUL} = 6$ $N_{ADD} = 4N_{CPR} + 2$	$N_{LUT} = 3$ $N_{MUL} = 12$ $N_{ADD} = 4N_{CPR} + 8$	NOT SIGNIFICANTLY DIFFERENT
DECISION	TWO DIFFERENTIAL DECODERS	$N_{LUT} = 2$ $N_{MUL} = 8(k-1)$ $N_{ADD} = 4N_{SSE} + 20$ $N_{CMP} = 23$	MORE COMPLEX FOR PC-DP-BPSK. THE INCREMENTAL COMPLEXITY IS HALF THAT OF THE SECOND-POWER CPR.

By comparing the element counts for DP-DEBPSK and PC-DP-BPSK, the incremental calculation complexity for PC-DP-BPSK is found to be the same as for one DP-DEBPSK CPR block. This means that it is possible to implement the signal processing for PC-DP-BPSK in a practical LSI. Each functional block is detailed in the following sections: symbol mapping, adaptive equalization, frequency offset compensation, CPR, and decision.

When PC-DP-BPSK is extended to higher-level modulation formats such as QPSK or multilevel QAM, the calculation complexities should be estimated separately because the algorithms are significantly different, as discussed in the sections "PC-DP-BPSK" and "PC-DP-B/QPSK."

SYMBOL MAPPING

While DP-DEBPSK needs two sets of differential encoders, PC-DP-BPSK requires one LUT whose number of addresses is four corresponding to the four signal points in the section "Sparse Asymmetric Polarization Block-Coded Symbol Mapping." A differential encoder needs a feedback path to calculate the exclusive logical sum (XOR) of the input data and the one-symbol delayed output data. Therefore, implementing symbol mapping for PC-DP-BPSK is easier than for DP-DEBPSK.

ADAPTIVE EQUALIZATION

While DP-DEBPSK employs both LMS and CMA algorithms, PC-DP-BPSK prefers an LMS algorithm to synchronize the phases in the X and Y polarizations for JCPR, as shown in the section "Joint-Polarization Carrier Frequency and Phase Recovery." Both LMS and CMA algorithms can be implemented, so this does not restrict the ability to implement adaptive equalization. When CMA is used, the miscapture problem needs to be solved for both DP-DEBPSK and PC-DP-BPSK [30].

SEVERAL SIMULATIONS CONDUCTED FOR PC-DP-BPSK SHOW THE POSSIBILITY OF 1 DB OR MORE IMPROVEMENT IN *Q*-FACTOR COMPARED TO DIFFERENTIAL CODING.

FREQUENCY OFFSET COMPENSATION

There is no difference between DP-DEBPSK and PC-DP-BPSK in terms of frequency offset compensation, except for averaging the frequency offsets over the X and Y polarizations for PC-DP-BPSK as described in the section "Joint-Polarization Carrier Frequency and Phase Recovery." The averaging process requires only a real adder, so the incremental calculation complexity is negligible.

CARRIER PHASE RECOVERY

As described in the section "Joint-Polarization Carrier Frequency and Phase Recovery," whereas $N_{\mathrm{LUT}} = 4$, $N_{\mathrm{MUL}} = 6$, and $N_{\mathrm{ADD}} = 4\,N_{\mathrm{CPR}} + 2$ are required for DP-DEBPSK, $N_{\mathrm{LUT}} = 3$, $N_{\mathrm{MUL}} = 12$, and $N_{\mathrm{ADD}} = 4\,N_{\mathrm{CPR}} + 8$ are required for PC-DP-BPSK. The difference comes from needing second-power CPR for DP-DEBPSK and fourth-power for PC-DP-BPSK.

DECISION

DP-DEBPSK requires a differential decoder, which consists of two comparators for deciding the two consecutive symbols, and an XOR for them in each polarization. The added calculation complexity is negligible.

PC-DP-BPSK requires CSC as described in the sections "Pre-Decision and Cost-Function Generation," "Slip State Estimation," and "Symbol Selection Based on the Estimated Slip State." $N_{\mathrm{MUL}} = 8(k-1)$, $N_{\mathrm{ADD}} = 24$ and $N_{\mathrm{CMP}} = 20$ for predecision and cost-function generation, $N_{\mathrm{ADD}} = 4\,N_{\mathrm{SSE}} - 4$ and $N_{\mathrm{CMP}} = 3$ for SSE, totaling $N_{\mathrm{MUL}} = 8(k-1)$, $N_{\mathrm{ADD}} = 4\,N_{\mathrm{SSE}} + 20$, and $N_{\mathrm{CMP}} = 23$ are required. This function is practical to implement because the incremental calculation complexity for PC-DP-BPSK is half that of second-power CPR or a quarter that of fourth-power CPR.

CONCLUSIONS

The management of CS is essential for coherent optical communication systems to achieve multi-Tb/s per channel optical transport. We reviewed the parameter design for CSC from the viewpoint of the error-burst correction capability of FEC. One way to eliminate residual burst errors is to shorten the CSC unit. This relaxes the requirements for miscompensation probability and BECL. To realize this by estimating the slip state of a modulated signal, we described a technique for constellation monitoring of a polarization block-coded signal. The recovered 2-D phase constellation becomes unique to a slip state with this technique even if the signal is modulated. Based on the resulting unique constellation, the slip state can be identified and compensated. As an example of a polarization block-coded signal, we described PC-DP-BPSK and the configuration of its DSP. Several simulations conducted for PC-DP-BPSK show the possibility of 1 dB or more improvement in Q-factor compared to differential coding.

While we have described a method to detect CS by examining a 2-D phase constellation, it is expected to be possible to detect CS in systems with higher spectral efficiency by examining higher-dimensional constellations by, e.g., signal point selection from QAM and time-block coding, which maps the bit-block data into signal states with multiple time indices.

ACKNOWLEDGMENT

This research is in part supported by the National Institute of Information and Communication Technology of Japan under its "λ Reach Project."

AUTHORS

Tsuyoshi Yoshida (Yoshida.Tsuyoshi@ah.MisubishiElectric.co.jp) received the B.E. and M.Info. degrees in electrical engineering from Kyoto University, Japan, in 2005 and 2007, respectively. Since 2007, he has been with Mitsubishi Electric Corporation, Kamakura, Japan, where he has been engaged in the research and development of high-speed optical transceiver and digital signal processing for high-capacity optical transmission. He is a

member of the IEEE Communications Society, IEEE Photonics Society, Optical Society of America, Institute of Electronics, Information and Communication Engineers of Japan, and Laser Society of Japan.

Takashi Sugihara received the B.E., M.E., and Ph.D. degrees in electrical engineering from Osaka University, Japan, in 1992, 1994, and 1998, respectively. Since 1998, he has been with Mitsubishi Electric Corporation, Kamakura, Japan, where he has been engaged in the research and development of high-speed optical transceiver and electrical technologies for transmission impairment mitigation. He is a member of the IEEE Communications Society, IEEE Photonics Society, Optical Society of America, Institute of Electronics, Information and Communication Engineers of Japan, Japan Society of Applied Physics, and Laser Society of Japan.

Kazuyuki Ishida received the B.E. and M.E. degrees in electrical engineering from Waseda University, Tokyo, Japan, in 1993 and 1995, respectively. Since 1995, he has been with Mitsubishi Electric Corporation, Kamakura, Japan, where he has been engaged in the research and development on optical wavelength division multiplexing transmission systems. He is a member of the Institute of Electronics, Information, and Communication Engineers of Japan.

Takashi Mizuochi received the B.E., M.E., and Ph.D. degrees in electrical engineering from Osaka University, Japan, in 1986, 1988, and 2004, respectively. Since 1988, at Mitsubishi Electric Corporation, he has been engaged in research on coherent optical fiber communications, long-haul transmission systems, undersea WDM communication systems, FEC for transport systems, and electronic signal processing. His current interests include the application of FEC and digital signal processing to optical networks. He is currently the general manager at Electro-Optics, Microwave and Communication Technologies Labs, Mitsubishi Electric Corporation. He is the associate editor of *IEEE Journal of Quantum Electronics*.

REFERENCES

[1] K. Sugihara, Y. Miyata, T. Sugihara, K. Kubo, H. Yoshida, W. Matsumoto, and T. Mizuochi, "A spatially-coupled type LDPC code with an NCG of 12 dB for optical transmission beyond 100 Gb/s," in *Proc. Optical Fiber Communication Conf.*, 2013, paper OM2B.4, pp. 1–3.

[2] D. Chang, F. Yu, Z. Xiao, Y. Li, N. Stojanovic, C. Xie, X. Shi, X. Xu, and Q. Xiong, "FPGA verification of a single QC-LDPC code for 100 Gb/s optical systems without error floor down to BER of 10-15," in *Proc. Optical Fiber Communication Conf.*, 2011, paper OTuN2, pp. 1–3.

[3] D. A. Morero, M. A. Castrillon, F. A. Ramos, T. A. Goette, O. E. Agazzi, and M. R. Hueda, "Non-concatenated FEC codes for ultra-high speed optical transport networks," in *Proc. IEEE Global Telecommunications Conf.*, 2011, pp. 1–5.

[4] C. R. S. Fludger, J. C. Geyer, T. Duthel, and C. Schulien, "Digital signal processing—From simulation to silicon," in *Proc. European Conf. Optical Communication*, 2010, paper Tu.5.A.1, pp. 1–3.

[5] E. Yamazaki, S. Yamanaka, Y. Kisaka, T. Nakagawa, K. Murata, E. Yoshida, T. Sakano, M. Tomizawa, Y. Miyamoto, S. Matsuoka, J. Matsui, A. Shibayama, J. Abe, Y. Nakamura, H. Noguchi, K. Fukuchi, H. Onaka, K. Fukumitsu, K. Komaki, O. Takeuchi, Y. Sakamoto, H. Nakashima, T. Mizuochi, K. Kubo, Y. Miyata, H. Nishimoto, S. Hirano, and K. Onohara, "Fast optical channel recovery in field demonstration of 100-Gbit/s Ethernet over OTN using real-time DSP," *Opt. Express*, vol. 19, no. 14, pp. 13179–13184, 2011.

[6] C. Rasmussen, Y. Pan, M. Aydinlik, M. Crowley, J. C. Geyer, P. Humblet, F. Liu, B. Mikkelsen, P. Monsen, N. Nadarajah, G. Pendock, and B. Shah, "Real-time DSP for 100+ Gb/s," in *Proc. Optical Fiber Communication Conf.*, 2013, paper OW1E.1, pp. 1–3.

[7] E. Ibragimov, B. Zhang, T. J. Schmidt, C. Malouin, N. Fediakine, and H. Jiang, "Cycle slip probability in 100 G PM-QPSK systems," in *Proc. Optical Fiber Communication Conf.*, 2010, paper OWE2, pp. 1–3.

[8] C. R. S. Fludger, D. Nuss, and T. Kupfer, "Cycle-slips in 100G DP-QPSK transmission systems," in *Proc. Optical Fiber Communication Conf.*, 2012, paper OTu2G, pp. 1–3.

[9] H. Zhang, Y. Cai, D. G. Foursa, and A. N. Pilipetskii, "Cycle slip mitigation in POLMUX-QPSK modulation," in *Proc. Optical Fiber Communication Conf.*, 2011, paper OMJ7, pp. 1–3.

[10] Y. Gao, A. P. T. Lau, and C. Lu, "Cycle-slip resilient carrier phase estimation for polarization multiplexed 16-QAM systems," in *Proc. Opto-Electronics and Communications Conf.*, 2012, paper 4B2-4, pp. 1–2.

[11] M. A. Castrillon, D. A. Morero, and M. R. Hued, "A new cycle slip compensation technique for ultra high speed coherent optical communications," in *Proc. IEEE Photonics Conf.*, 2012, paper MU2, pp. 175–176.

[12] M. Magarini, L. Barletta, A. Spalvieri, F. Vacondio, T. Pfau, M. Pepe, M. Bertolini, and G. Gavioli, "Pilot-symbols-aided carrier-phase recovery for 100-G PM-QPSK digital coherent receivers," *Photon. Technol. Lett.*, vol. 24, no. 9, pp. 739–741, 2012.

[13] M. G. Taylor, "Phase estimation methods for optical coherent detection using digital signal processing," *J. Lightwave Technol.*, vol. 27, no. 7, pp. 901–914, 2009.

[14] K.-T. Wu and H. Sun, "Techniques in carrier recovery for optical coherent systems," in *Proc. Optical Fiber Communication Conf.*, 2012, paper OTh4C.3, pp. 1–3.

[15] W. J. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Trans. Commun.*, vol. 26, no. 3, pp. 385–391, 1978.

[16] E. Agrell and M. Karlsson, "Power-efficient modulation formats in coherent transmission systems," *J. Lightwave Technol.*, vol. 27, no. 22, pp. 5115–5126, 2009.

[17] J. K. Fischer, L. Molle, M. Nölle, D.-D. Groß, and C. Schubert, "Experimental investigation of 28-GBd polarization-switched quadrature phase-shift keying signals," in *Proc. European Conf. Optical Communication*, 2011, paper Mo.2.B.1, pp. 1–3.

[18] T. Yoshida, T. Sugihara, K. Ishida, and T. Mizuochi, "Slip-state estimation by polarization-phase constellation monitoring in phase-shift keyed transmission," in *Proc. OptoElectronics and Communication Conf./Int. Conf. Photonics in Switching*, 2013, paper ThR3.2, pp. 1–2.

[19] A. Dochhan, H. Griesser, and M. Eiselt, "First experimental demonstration of a 3-dimensional simplex modulation format showing improved OSNR performance compared to DP-BPSK," in *Proc. Optical Fiber Communication Conf.*, 2013 paper JTh2A.40, pp. 1–3.

[20] L. D. Coelho and N. Hanik, "Global optimization of fiber-optic communication systems using four-dimensional modulation formats," in *Proc. European Conf. Optical Communication*, 2011, paper Mo.2.B.4.

[21] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inform. Theory*, vol. 29, no. 4, pp. 543–551, 1983.

[22] T. H. Lotz, X. Liu, S. Chandrasekhar, P. J. Winzer, W. Sauer-Greff, and R. Urbansky, "Experimental PDM-OFDM signal transmission with shaped signal constellations achieving a net spectral efficiency of 11.15 b/s/Hz over a distance of 800 km," in *Proc. Int. OFDM Workshop*, 2012, pp. 1–6.

[23] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightwave Technol.*, vol. 26, no. 20, pp. 3416–3425, 2008.

[24] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "DSP for coherent single-carrier receivers," *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3614–3622, 2009.

[25] Optical Internetworking Forum. (2013, Aug. 27). Implementation agreement for integrated polarization multiplexed quadrature modulated transmitters, IA # OIF-PMQ-TX-01.0. [Online]. Available: http://www.oiforum.com/public/documents/OIF-PMQ-TX-01.1.pdf

[26] L. Li, Z. Tao, S. Oda, T. Hoshida, and J. C. Rasmussen, "Wide-range, accurate and simple digital frequency offset compensator for optical coherent receivers," in *Proc. Optical Fiber Communication Conf.*, 2008, paper OWT4, pp. 1–3.

[27] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Express*, vol. 16, no. 2, pp. 804–817, 2008.

[28] G. E. P. Box and M. E. Muller, "A note on the generation of random normal deviates," *Ann. Math. Stat.*, vol. 29, no. 2, pp. 610–611, 1958.

[29] M. Winter, D. Setti, and K. Petermann, "Cross-polarization modulation in polarization-division multiplex transmission," *Photon. Technol. Lett.*, vol. 22, no. 8, pp. 538–540, 2010.

[30] M. Yan, Z. Tao, H. Zhang, W. Yan, T. Hoshida, and J. C. Rasmussen, "Adaptive blind equalization for coherent optical BPSK system," in *Proc. European Conf. Optical Communication*, 2010, paper Th.9.A.4, pp. 1–3.

[SP]

Moshe Nazarathy and Alex Tolmachev

Subbanded DSP Architectures Based on Underdecimated Filter Banks for Coherent OFDM Receivers

Overview and recent advances

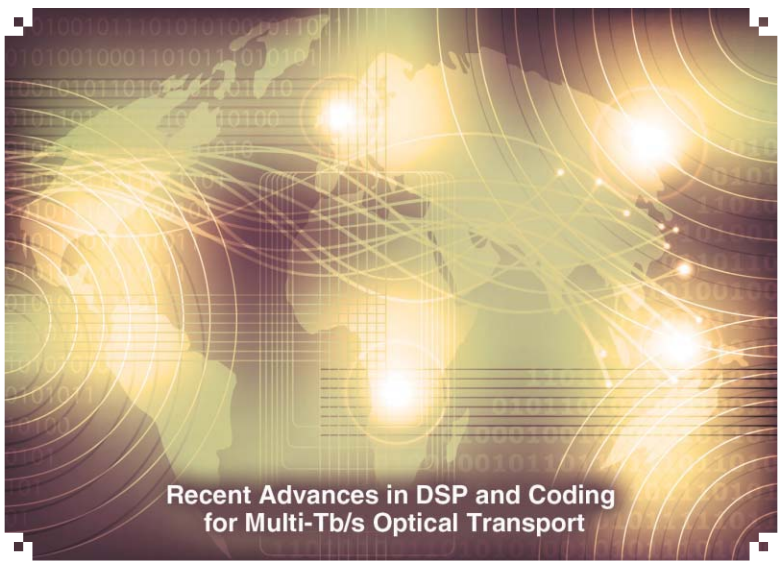


IMAGE LICENSED BY INGRAM PUBLISHING /JANMIKS

Subbanded digital signal processing (DSP) with underdecimated (Udeci) filter banks (FBs), is a recent DSP technique whereby the optical channel bandwidth is digitally sliced into multiple spectrally disjoint subbands (SBs) to be processed in parallel. In terms of DSP hardware (HW) architecture, digital subbanding amounts to an alternative mode of parallelizing the receiver signal processing task to multiple slower processors, whereby the parallelization performed in the frequency domain (FD) rather than in the time domain (TD). We show that FD parallelization is especially suited to the long-haul optical fiber channel and present novel receiver DSP structures based on Udeci FBs, providing substantial complexity savings for ultra-high-speed optically coherent transmission.

Digital Object Identifier 10.1109/MSP.2013.2291015
Date of publication: 12 February 2014

INTRODUCTION

FBs have been known in DSP for decades and applied to electrical wireless and wireline communication [1]. Oversampled FBs [2], more recently referred to as *Udeci FBs* [3], form an FB subclass enjoying several advantages over critically decimated (Cdeci) or

1053-5888/14/$31.00©2014IEEE

critically sampled FBs. We show that digital subbanding based on Udeci FBs may provide sizable savings of complexity in the DSP section of coherent optical receivers (Rx) over ultra-high-speed fiber-optic links, as well as improve spectral efficiency and performance.

Udeci FBs are likely to become a key enabling technology for multi-Tb/s optical receivers, as they help improve the energy efficiency of the multi-Tb/s networks as the data throughput continues to grow.

OVERVIEW OF RECENT OPTICAL DETECTION TECHNIQUES AND THEIR USE OF DSP TECHNIQUES

We briefly review recent trends in photonic transmission, exploring commonalities and differences versus electrical communication techniques, motivating the role of FB DSP techniques in the upcoming generation of optical communication systems for multi-Tb/s transmission.

OPTICALLY COHERENT DETECTION

In recent years, we have witnessed the advent and commercial deployment of optically coherent detection (OCD) [4], whereby a laser is used in the Rx as an optical local oscillator (OLO), in addition to having a laser in the transmitter (Tx) as an optical source. The coherent optical Rx demodulates the received optical field, mixing it with the OLO field via the quadratic detection effect provided by the photodiodes (PDs). If the LO frequency does not coincide with that of the signal, then the net result is the downconversion of the optical spectrum to an IF electrical frequency, providing the front end of a coherent optical heterodyne receiver. Having the two frequencies (nearly) coincide and using two OLOs in quadrature results in homodyne (or intradyne) coherent detection, generating the in-phases and quadrature (IQ) components of the optical field complex envelope (CE) (or equivalently magnitude and phase) Subsequently, analog to digital conversions followed by DSP may be applied in the optical receiver. OCD is a recent trend, rapidly making its way into large-scale commercial deployment. In contrast, earlier optical communication techniques had been traditionally based on optical direct detection (ODD), essentially using the Rx PD as a power detector. The ODD quadratic detection generates the squared magnitude of the optical field, erasing field phase information. Gaining access to the full CE by means of OCD is a "game changer," motivating advanced DSP in the Rx.

THE AGE OF DSP IN OPTICALLY COHERENT TRANSMISSION

Once optical field CE becomes observable by means of OCD, there should be, in principle, little difference between modern electrical wireless or wireline Rx-s versus coherent optical Rx-s. Photonic OCD is just as an extremely wideband case of detection of ultra-high rate data modulated onto an extremely high-frequency carrier. This point of view has spurred within the optical communication community a trend of porting to OCD a host of detection structures and DSP techniques originally developed in the context of electrical transmission. Optical communication researchers have been competing in identifying various modulation formats and DSP techniques to be adopted for advantageous OCD use.

DISTINGUISHING TRAITS OF PHOTONIC TRANSMISSION

Nevertheless, the long-haul fiber-optic channel bears unique traits, differing from electrical wireless/wireline media in five major respects.

1) The fiber-optic channel is ultra-broadband, enabling long-haul fiber links for data and voice communications backbones of national or global coverage. Over the broadband optical spectrum, wavelength division multiplexing (WDM) techniques are used to densely frequency-multiplex multiple optical channels, each of which typically occupies several tens of GHz and carries data rates the order of 100 GB/s. An optical channel is typically two to three orders of magnitude faster than what is customary in electrical transmission, making DSP design and application-specific integrated circuit (ASIC) realization of the optically coherent Rx extremely challenging. Here we address for definiteness the Rx DSP for a single optical channel, with typical bandwidth (BW) of 25 GHz, transmitting quadrature phase shift keying (QPSK) or 16-QAM (BW refers in this article to the full spectral support of a signal, not 3 dB BW). Ultra-dense WDM of multiple such channels to form Tb/s optical superchannels, is treated elsewhere in this issue of *IEEE Signal Processing Magazine*.

2) Third-order nonlinear effects unique to the fiber-optic propagation (in contrast, wireless/wireline electrical channels are typically linear). This article focuses on linear processing—nonlinear impairment mitigation is addressed elsewhere in this issue.

3) Extremely long delay spread [duration of the channel impulse response (IR)] of long-haul fiber links due to the accumulation of chromatic dispersion (CD) for the wideband signals. As CD is a linear effect, it may be mitigated by linear equalizer (EQZ) filters. Note that dispersive delay spread is also present in wireline transmission, e.g., over telephony DSL and electrical power lines. However, the channel memory of long-haul fiber-optic links is on the order of 100 symbols, which is one to two orders of magnitude longer than in electrical transmission. This indicates that the CD EQZ requires an excessive number of taps, even if optimized by using an FD EQZ such as overlap save (OLS). Notice that for second-order dispersion multiple CD EQZ filter taps are characterized by a single degree of freedom, specifically the CD differential group delay parameter (higher-order CD would entail additional degrees of freedom.

4) The single-mode fiber (SMF)-optic channel is inherently 2×2 multiple input, multiple output (MIMO), as it supports two orthogonal polarizations (POLs), referred to here as X- and Y- POL. The polarization mode dispersion (PMD) effect implies that the fiber dispersive delay spread is POL-dependent. PMD is impacted by environmental fiber disturbances on a sub-ms scale, hence it must be adaptively tracked by the 2×2 MIMO EQZ (moreover, the EQZ taps must be tunable to absorb the uncertainty in the estimation of the amount of CD). Granted, MIMO techniques of even higher dimension have proliferated in wireless transmission. Nevertheless, at

ultra-high symbol rate, the 2×2 MIMO EQZ with $4\times O(10)$ taps memory takes its toll onto the overall complexity of coherent Rx DSP.

5) The acceptable outage probabilities in the high-capacity long-haul optical trunks are several orders of magnitude lower than in typical wireless links, impacting the selection of DSP algorithms, e.g., to mitigate outage effects due to PMD.

COHERENT Rx COMPLEXITY

It turns out that in a typical modern coherent Rx operating at tens of GS/s, the two linear EQZ functions—CD EQZ and POL/PMD 2×2 MIMO EQZ—dominate the Rx DSP computational complexity (reflected mainly in the number of multipliers), and place excessive requirements on area and power consumption of ASIC implementations of the DSP section of the Rx. As transmission rates are expected to follow a Moore-like growth law, the Rx DSP complexity is identified as a bottleneck about to stifle the expansion of the Internet backbone infrastructure. Minimizing DSP computational load without sacrificing performance is recognized as a critical goal of the modern photonic receiver algorithmic system design. We aim to show that known FB DSP techniques may be applied along with novel extensions to advance the complexity reduction/energy-efficiency objective, contributing to the continued growth of the global communications.

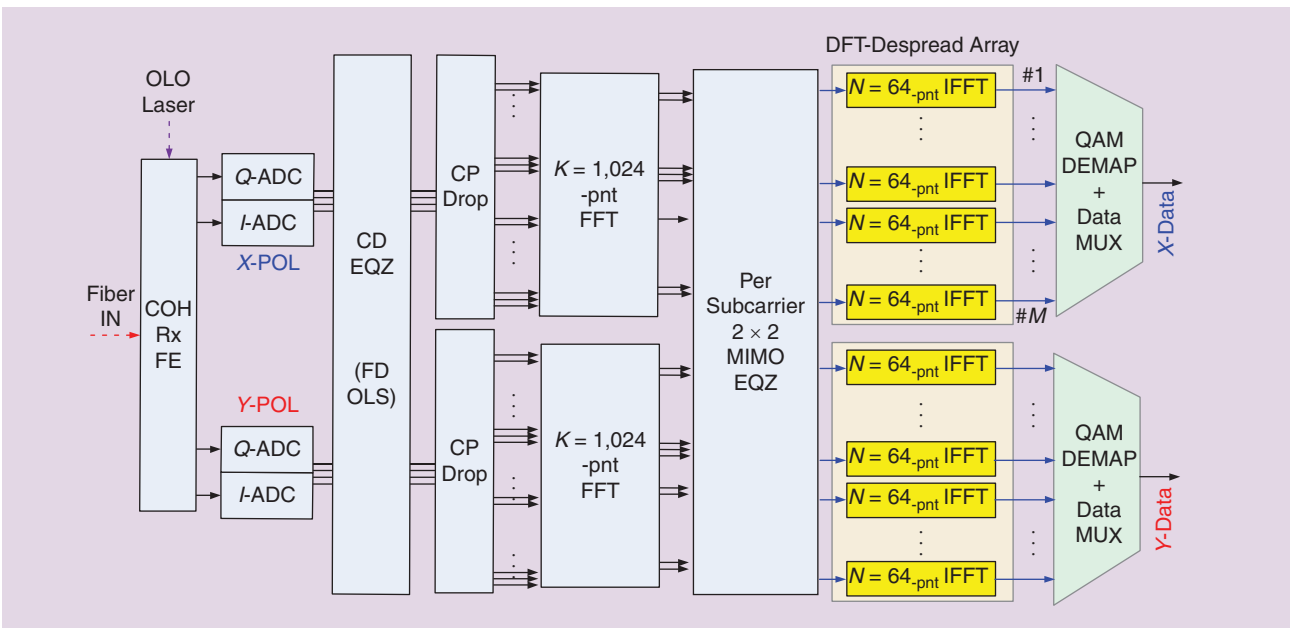
MULTICARRIER TECHNIQUES IN OPTICAL COMMUNICATIONS

Multicarrier techniques such as OFDM are widespread in electrical transmission and have made their way into optical communication as well [5]–[9], though not yet commercially deployed. Intense research and debate continues on the relative merits of single-carrier (Nyquist) versus multicarrier (OFDM) formats for optical transmission [10], [11]. Figure 1 presents the block diagram of a state-of-the-art conventional coherent optical OFDM Rx for a single optical channel. In fact, the shown system is augmented to incorporate the so-called discrete Fourier transform (DFT)-spread (DFT-S) variant (also adopted in fourth-generation wireless transmission under the revised name single-carrier FDMA, but still dubbed DFT-S in optical communication [7]). Note that in wireless transmission, this technique is used in a multiple access context, whereas in photonic long-haul transmission it is proposed to be used to transmit multiple tributaries point-to-point. In DFT-S [7], in addition to the main MN-pnt OFDM FFTs in the Rx (one for each POL), an array of M (I)FFTs each of size N are used in the Tx and the Rx to effectively generate and detect M parallel tributaries, which may be viewed as frequency multiplexed narrower band single carrier signals (SCs). Advantages of DFT-spread OFDM include reduced peak-to-average power ratio and improved nonlinear [12] and phase noise tolerances.

BRIEF HISTORICAL ACCOUNT OF DIGITAL FBs FOR OCD

Actually, FBs are routinely used in optical communication without naming them as such. In particular, the WDM optical channelization technique, essentially assembling/separating out multiple transmissions at different "colors" (wavelengths/optical frequencies), may be viewed as an analog form of FB, albeit in the optical domain. However, our interest is in digital FBs.

In 2009, Ho [13] was the first to propose a coherent optical link based on Udeci (oversampled) FBs, pointing out a substantial complexity advantage upon performing the CD EQZ on an SB basis. He observed that Udeci FBs may be potentially realized with lower complexity than a Cdeci FB, for a given channel BW and number of SBs and for comparable distortion, but did not elaborate on specific DSP structures realizing such advantage.



[FIG1] The conventional high-speed DFT-S OFDM Rx, where FE stands for front end and CP for cyclic prefix. The OFDM Tx is not shown but it is essentially the DSP dual of the Tx, comprising two IFFTs, one for each POL. A plain OFDM Rx (non-DFT-S) may be obtained by simply removing the DFT-despread array.

CD EQUALIZATION ADVANTAGE OF SUBBANDED Rx

Here we reinterpret Ho's argument regarding the CD EQZ complexity reduction, expanding and quantifying it as follows: Let $C[\mu]$ measure the complexity cost (e.g., the multipliers count) of the CD EQZ, with $\mu = \lceil \Delta\tau T_s^{-1} \rceil$ the CD delay spread gauged in units of discrete-time sampling intervals, T_s. For a finite IR (FIR) TD implementation of the EQZ, $C = O(\mu)$, whereas for an FD implementation, $C = O(\log\mu)$ (the log stemming from the Cooley–Tukey DFT fast algorithm). In turn, the analog delay spread is $\Delta\tau = \beta_2 L \cdot B$, where L is the fiber length and β_2 is the CD parameter. The sampling rate also varies proportionally to the BW, $T_s^{-1} = \eta B$ (with η a factor related to the spectral efficiency and oversampling ratio). The CD delay spread is then $\mu = \lceil \beta_2 \eta L \cdot B^2 \rceil$, quadratic in the BW. The impact of slicing the total channel BW, B, into M SBs (B/M per SB) and using a separate CD EQZ in each SB is dramatic—the delay spread in each SB is reduced by a factor of M^2. For an FIR TD EQZ, the complexity per SB is reduced by a factor of $O(M^2)$, whereas for an overlap-save FDE EQZ the complexity per SB is reduced by a factor of $O(M \log M)$. Since there are M SBs operating in parallel, the total CD EQZ complexity is M times larger and the savings and the final CD EQZ savings are $O(M)$ for a TD EQZ and $O(\log M)$ for an FD EQZ. For example, for a $B = 25$ GHz channel transmitted over a 2,000 km link of standard SMF, $\mu = 150$ taps, slashed down to $\mu/M^2 = 150/15^2 = 2/3$ for $M = 15$ SBs, i.e., a fraction of a single tap per SB. Thus, just a single tap EQZ suffices to fix the CD in each SB, as further elaborated in the section "Subband Processors–2×2 MIMO DFT-S OFDM Receivers." Once the channel spectrum has been sliced into multiple SBs, the SB processors become extremely simple. This is an instance of a divide and conquer algorithm, with the divide performed in the FB and the conquer occurring in the multiple SB Rx-s, the complexity of all of which together is much lower than the complexity of a full-band Rx—provided the FB realization overhead is kept sufficiently low (in the section "Novel Efficient Realization of a 2 $\times$Udeci FB Core by Combining a Pair of Cdeci FBs," we show how to efficiently implement the Udeci FB).

Fast-forwarding to 2013, Liu et al. [14] experimentally demonstrated this concept by long-haul optical transmission over a subbanded system constructed on the basis of a DFT-spread OFDM structure acting as an equivalent FB, realized by inserting appropriate spectral guardbands within each SB. Their offline SW verified that the CD EQZ becomes more computationally efficient when the number of SBs is increased. Further improvements are desired to reduce the spectral overhead used for the guardbands that incurred about 10% spectral efficiency loss, and to increase in the number of SBs to enjoy even higher subbanding complexity advantage. This spectral efficiency challenge is addressed in our own Udeci FB approach, but prior to that, let us mention that the subbanding concept surfaced in the last few years in the coherent optical OFDM context. Du and Lowery [15] proposed an OFDM-based subbanding technique at the Tx-side, without addressing its complexity of realization. An optical OFDM review by Jansen [5] accounted for the subbanding advantage in OFDM transmission in terms of a reduction in the cyclic prefix (CP) duration, which may be taken as the CD delay spread per SB, hence should vary as $(1/M)\Delta\tau = (1/M)\beta_2 L \cdot B$, substantially reduced for large SB count, M. However, Jansen assumed that the slicing into SBs is optically performed, by means of numerous narrow band optical channels, each with its DAC, optical modulator at the Tx, optical channel filter, coherent optical front-end, analog-to-digital converter (ADC), DSP ASIC at the Rx, which would be highly inefficient in HW resources as well as inefficient in spectral efficiency due to the required optical guardbands. Evidently, the digital subbanding approach based on FBs would be preferable, provided the digital complexity could be reduced in both the FB core and in the SB processors, which act as mini-Rx-s in their own right albeit over narrowband SBs. We made our contribution to this objective [16]–[23] by introducing efficient implementations of Udeci FBs and dechannelizers for (DFT-S) OFDM Rx-s. We have been inspired in our own approach mostly by harris [3], [24], [25] (note that this author insists on spelling his name in lowercase) but our approach has evolved beyond it in several respects, as we focus on OFDM. For the Udeci FB core itself, we present an alternative different realization equivalent to harris'. Their DSP Rx structures, constructed around their Udeci FB core realization, are not equivalent to ours but typically have a different focus—intertwining data modulation with up/down frequency conversion and multiple sampling rate conversions. This is a distinct perspective than ours—we rather devote the Udeci FB to the specific DFT-S OFDM context, utilizing it for the "divide" part of "divide and conquer." The harris Udeci FB structure differs from our novel structure yet incurs the same multipliers count, however, it might be less suitable for high-speed optical communication, as it is less amenable to time-parallelization due to its state-machine structure of the cyclic buffers at the I/O of the polyphase filter array. A harris dechannelizer would be more costly to implement in an ultra-high-speed ASIC as required for optical communication. In contrast, our Udeci FB realization is much simpler, "recycling" two well-known Cdeci FB modules, interconnecting them to form a Udeci FB (see the section "Novel Efficient Realization of a 2$\times$ Udeci FB Core by Combining a Pair of Cdeci FBs"). HW designers would appreciate this realization simplicity.

In the HW efficiency area, we make a contribution to both the Udeci FB core and the rest of the Rx, by combining the generic Udeci FB core with additional DSP (interpolators/decimators) to generate a complete multichannel (de)channelizer. The resulting scheme may be interpreted as a DFT-Spread OFDM receiver (DFT-S features PAPR and phase noise advantages versus plain OFDM). To our knowledge, we are the first to disclose OFDM Rx structures based on Udeci FBs. In contrast, in [26], multichannel techniques based on Cdeci FBs for electrical communication were considered as an alternative to OFDM. In our approach we do not pit the FB versus OFDM techniques against each other as mutually exclusive options, but we combine them to advantage in efficient OFDM receiver structures based on Udeci FBs rather than Cdeci FBs.

Our most recent contribution is to "go experimental" [23], but not in an offline optical transmission experiment (optically propagated signals acquired by ultra-fast samplings scopes and processed offline, which gains the title "experimental" in optical

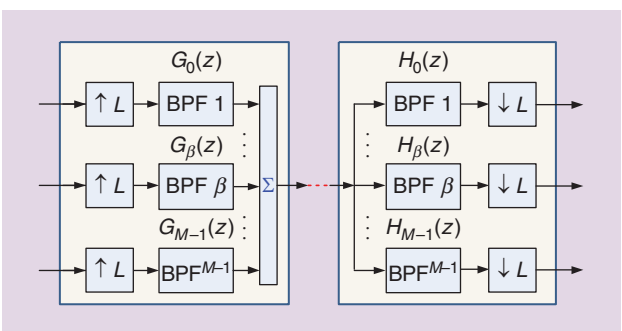
[TABLE 1] SOME OPERATORS OF MULTIRATE THEORY—OUR NOTATION.

${}^{v}O_{u}^{(\text{param})}$	GENERIC OPERATOR	${}^{v}O_{u}\{x[u]\} = y[v]$	LABELED BY I/O VARIABLES AND PARAMETERS
IDFT^{M} DFT^{M}	(INVERSE) DFT	$X[\beta] = {}^{\beta}\text{IDFT}_{p}^{M-\text{pnt}}\{x[p]\} = \sum_{p=0}^{M} x[p]\bar{W}^{p\beta}$ $x[p] = {}^{p}\text{DFT}_{\beta}^{M-\text{pnt}}\{X[\beta]\} = (1/\text{M})\sum_{p=0}^{M} X[\beta]\bar{W}^{-\beta p}$	$\bar{W}_M = e^{j2\pi/M}$; $W_M = e^{-j2\pi/M}$
$[x]_R$	MODULO-R REDUCTION	$[x]_R \equiv x \bmod R \Leftrightarrow x = \lfloor x/R \rfloor R + [x]_R$ $[\bullet]_R: \mathbb{R} \mapsto [\nu_0, \nu_0 + R),\ R, \nu_0 \in \mathbb{R}$ or $[\bullet]_R: \mathbb{Z} \mapsto \{k_0, k_0+1, \ldots k_0+R-1\},\ R, k_0 \in \mathbb{Z}$	DEFINED FOR EITHER REALS OR INTEGERS FOR REALS, TYPICALLY $\nu_0 = 0$ OR $\nu_0 = -R/2$
$C^{[s]_R}$	CIRCULAR SHIFT MODULO-R	$C_{\nu}^{[s]_R}\{X(\nu)\} \equiv X([\nu - s]_R)$	s IS THE DELAY PARAMETER
$\downarrow L$	DOWNSAMPLING	$\{x[k]\}_{\downarrow L} \equiv x[kL],\ \ L \in \mathbb{Z}$	
$\uparrow L$	UPSAMPLING	$\{y[k]\}_{\uparrow L} \equiv \{\ldots, \underbrace{y[0], 0, 0, \ldots, 0}_{L}, \underbrace{y[1], 0, 0, \ldots, 0}_{L}, \ldots\}$	
$x^{[p]_P^{type}}[k]$	pTH-POLYPHASE COMPONENT MODULO-P (TYPE $\in$ {1, 2, 3}) OF A SEQUENCE $\{x[k]\}$	$x^{[p]_P^{\prime}}[k] \equiv x[kP + p] = \{x[k+p]\}_{\downarrow P}$ $x^{[p]_P^{\natural}}[k] \equiv x[kP + (P-1-p)]$ $x^{[p]_P^{\prime\prime\prime}}[k] \equiv x[kP - p] = \{x[k-p]\}_{\downarrow P}$	FOR TYPE-1 BY DEFAULT, WE DISCARD THE TYPE SUPERSCRIPT: $x^{[p]_P}[k] \equiv x^{[p]_P^{\prime}}[k]$
$S/P^{(1:L)}$	1:L SERIAL-TO-PARALLEL MODULE	$S/P^{(1:L)}$: $x[k] \rightarrow \left\{x^{[0]_L^{\prime\prime\prime}}[k], x^{[1]_L^{\prime\prime\prime}}[k], \ldots, x^{[L-1]_L^{\prime\prime\prime}}[k]\right\}$	GENERATES THE TYPE-3 POLYPHASE COMPONENTS

communication) but rather we prepare for the fastest ever "online" experiment. We have recently constructed and tested a real-time FPGA at the unprecedented 25 Gbaud (Gsamp/s) symbol rate for DFT-spread OFDM with 16-QAM modulation. This is work in progress, with back-to-back electrical fiber transmission already demonstrated over the real-time field-programmable gate array (FPGA) incorporating all Rx processing functions, but optical transmission is not shown yet. Electrically, our Rx was shown to process 160 Gb/s over the 25-GHz optical channel BW over dual POLs and was able to inspect any SB.

In the last two years, other contributions to digital subbanding by FBs have been made by Randel et al. [27], [28], focusing on spectrally efficient real-time implementation at the Tx side, using multicarrier offset-QAM. This complements our own FB-based work on real-time Rx-s [23].

A recent work on digital subbanding from Huawei and related academics [29] has reinvented elements of our previously published approach, bringing CD equalization down to a single tap along with integer delay per SB, duplicating a concept from our previous proposal and simulations [16]–[19], [21], [22], which we recently demonstrated in FPGA HW [23].



[FIG2] Back-to-back synthesis and A-FBs, Cdeci ($L = M$), and Udeci ($L < M$).

Cdeci VERSUS Udeci UNIFORM FBs

As the Udeci FBs are DSP structures, we must be precise in our development of these concepts, so we are now "going mathematical" aided by the operator notation listed in Table 1.

An FB is defined as a collection of filters with common input or output. In a communication link based on FBs (Figure 2), at the Tx we place a synthesis FB (S-FB), forming the Tx output as the sum of the outputs of S-FB filters, in turn fed by M tributary signals to be "multiplexed" for transmission. The tributary signals may undergo some preprocessing at the tributary level. At the Rx, a corresponding analysis FB (A-FB) acts as a demultiplexer. The received signal is 1:M split to feed a collection of filters. After some optional postprocessing, the filter outputs (M outputs of the A-FB) should ideally reconstruct the tributary signals launched at the Tx-side.

We are interested in uniform FBs, meaning that filters forming the FB are all band-pass filters (BPFs) with spectral passbands appearing identical around their center frequencies, which are on a regular grid along the frequency axis, ν, at S [Hz] separation. The total frequency span for the whole FB is $R = MS$. In our case, R coincides with the optical channel BW (and with the inter-channel frequency spacing in a WDM multichannel system). The FB slices the channel spectral support into M SBs with $S = R/M$ spectral separation). The uniform FB requirement is equivalent to having all BPFs transfer functions (TFs) consist of uniformly shifted replicas of a prototype filter (PF) TF: $H_\beta(e^{j2\pi\nu/R}) = H_0(e^{j2\pi(\nu-\beta S)/R})$. Equivalently the IRs $h_\beta[k] = Z^{-1}\{H_\beta(z)\}$ (with Z the z-transform) are harmonic modulations of the IR, $h_0[k]$, of the PF: $h_\beta[k] = e^{j2\pi\beta k/M} h_0[k] = \bar{W}_M^{\beta} h_\beta[k]$.

As BW is reduced by a factor of M upon going from the channel down to the SB level, it is useful to downsample the BPF outputs by some integer factor, L. In particular, for $L = M$, we have a classic FB structure, referred to as *critically sampled* or *critically Udeci*. This Cdeci FB has been known for

half a century. In this case, if the PF is designed with one-sided BW S, then each SB is sampled at its Nyquist rate, $S = R/M$. If $L < M$, the FB is referred to as oversampled or Udeci (e.g., [3], [24], and [25]). In such Udeci FB, the SBs are oversampled by a factor of $V = M/L > 1$, referred to here as the *Udeci-factor*; the corresponding Udeci FB is referred to as $V\times$ Udeci. This article treats twice-Udeci ($2\times$ Udeci) FBs ($V = 2, L = M/2$), but for generality, some derivations will be carried out for arbitrary V, setting $V = 2$ at the end.

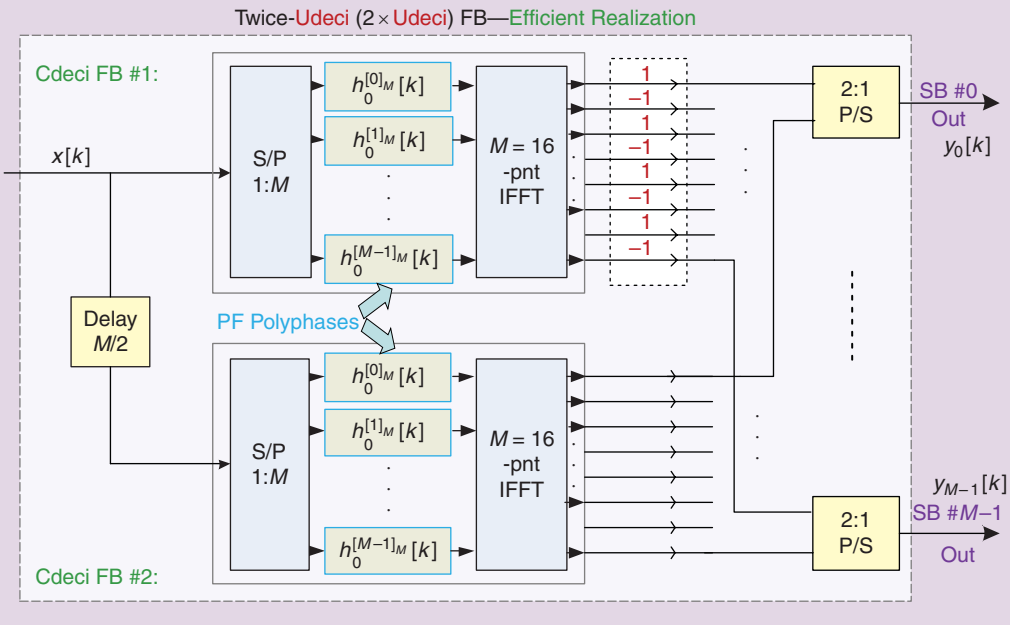
We commence with the Cdeci class case ($V = 1$), specifically Figure 2 with $L = M$, reviewing its well-known polyphase realization, referred to as *uniform DFT* (*U-DFT*) *FB*. This efficient Cdeci FB realization is not allocated a separate figure here, as it is taught in most DSP textbooks, e.g., [25], but its specification is as follows: a 1:M serial-to-parallel (S/P) module feeding an array of M polyphase filters, the outputs of which in turn feed an M-pnt DFT. The polyphase filters have IRs $\{h_0^{[\gamma]_M}\}_{\gamma=0}^{M-1}$, given by the M polyphase components of the PF filter $h_0[k]$ (see Table 1 for polyphase components definition). The classic U-DFT realization of the Cdeci FB (shown as the submodule marked Cdeci #1 or #2 in Figure 3) is far more efficient than a brute-force approach of M BPFs in parallel (Figure 2), nevertheless it not efficient enough, since the requisite number of taps of the PF (equal to the sum of the number of taps of the polyphase filters) is still prohibitive. The reason for having too many taps is that the BPFs forming the A-FB, used to carve the channel into spectrally disjoint SBs, must have nearly rectangular TFs, to prevent either crosstalk between SBs or loss of spectral efficiency. Such nearly brickwall PF typically require $O(10^3)$ taps. This hampers the applicability of classic Cdeci FBs to energy-efficient communication applications such as optical transmission, motivating the adoption of Udeci FBs. However, the U-DFT Cdeci FB will emerge as a building block in our novel efficient implementation of the Udeci FB.

NOVEL EFFICIENT REALIZATION OF A $2\times$ Udeci FB CORE BY COMBINING A PAIR OF Cdeci FBs

Udeci FB structures are obtained by setting $L = M/V < M$, or equivalently $V > 1$ in Figure 2 (in particular, $V = 2$ for $2\times$ Udeci FBs). These FBs are inefficient in the form of Figure 1, as the realization of multiple sharp BPFs at high speed is very complex. Here we disclose for the first time an efficient polyphase realization of a uniform $2\times$ Udeci FB, conveniently structured as interconnection of a pair of U-DFT Cdeci FB modules (Figure 3), making it easy to implement the FB in HW (or in software (SW) at much lower speeds) as the Cdeci FB, to replicated twice to make a Udeci, is a well-known module for which HW cores already exist (one major FPGA maker supplies a Cdeci FB core).

The "glue" connecting the two Cdeci modules to form a $2\times$ Udeci A-FB is very simple, as shown in Figure 3. One Cdeci FB input is delayed by $L = M/2$, whereas the outputs of the other Cdeci FB are alternated in sign (modulated by ±1). Equivalently, in case we have access to the internals of the Cdeci FB, we may replace the ±1 alternation by a half-wave circular shift, swapping of the upper and lower halves of the input vector into the DFT, statically crossing the wires.

The two Cdeci FB modules provide the even and odd components of the outputs $y_\beta[k]_{\beta=1}^{M}$ of the $2\times$ Udeci A-FB, thus if we require the FB output in serial form, we just interleave the two M-pnt output vectors generated by the two Cdeci FBs, by means of $M/2$ 1:2 Parallel/Serial (P/S) modules. Our proof that Figure 3 is equivalent to the generic A-FB structure at the Rx side of Figure 2



[FIG3] A novel efficient structured $2\times$ Udeci FB realization based on interleaving a pair of Cdeci FBs. The *M* outputs of Cdeci FB#1 are alternated in sign and time-interleaved with the corresponding *M* outputs of Cdeci FB#2 yielding for each of the *M* output SB signals at rate $2R/M$ where R is the rate of $x[k]$.

(with L there taken as $M/2$) is based on first deriving the polyphase modulo L IRs, $\{h_\beta^{[\gamma]_L}[k]\}_{\gamma=1}^{L}$, for each of the $M = 2L$ BPFs of the $2\times$ Udeci FB, in terms of the L polyphases $\{h_0^{[\gamma]_L}[k]\}$ of the PF

$$h_\beta^{[\gamma]_L}[k] = h_0^{[\gamma]_L}[k](-1)^{\beta k}\bar{W}_{2L}^{\beta\gamma}, \qquad \beta = 0,1,\ldots,2L-1; \quad \gamma = 0,1,\ldots,2L-1. \tag{1}$$

A polyphase representation of the output of any filter followed by $\downarrow L$ downsampling is derived in most DSP textbooks based on Noble's identities (e.g., [25]), and is applied here to express the βth output of the Udeci FB, in response to the wideband input $x[k]$ into the FB

$$y_\beta[k] = \sum_{\gamma=0}^{L-1} x^{[\gamma]_L^{III}}[k] \otimes h_\beta^{[\gamma]_L}[k]. \tag{2}$$

Interestingly, the type-3 polyphases of the input signal interact with the type-1 polyphases of same order of the filter IR. This equation fully specifies the I/O mapping of the 2xUdeci FB (which is a 1:2L SIMO system), but its direct realization would not be most efficient. It remains to manipulate this equation [with (1) set in] to bring it to a form equivalent to the highly efficient form of Figure 3. The derivation is omitted but it is based on Table 1 and multirate properties.

Let us compare our resulting efficient Udeci FB realization of Figure 3 with prior ones by harris et al. [3], [24], [25], and also by our group (our current novel realization differs from our prior realization introduced in [18], which was based on $M/2$ single-input, dual output (SIDO) polyphase filters preceding the M-pnt DFT). It turns out that all three Udeci FB realizations incur the same multipliers count, however our two alternative realizations (in [18] and here) feature a high-speed operational advantage over that of harris', due to easier time-parallelization, as explained in the section "Brief Historical Account of Digital FBs for OCD."

(DE)CHANNELIZER BASED ON Udeci FBs PLUS INTERPOLATORS/DECIMATORS

In our recent optical communication-oriented research [16]–[23], we presented DSP structures augmenting the Udeci FBs at the Tx and Rx by additional elements, converting them to multichannel (de)muxes, or (de)channelizers with nearly perfect reconstruction (Figures 4 and 5). The added elements are interpolators/decimators, combined with the efficient Udeci core of the last section to form a complete distributed (de)channelizer system. At the Tx-side, by attaching V-fold interpolators on the input ports of the S-FB at the Tx, we obtain an FD-mux of multiple input tributaries; at the Rx-side, by attaching V-fold interpolators on the input ports of the A-FB, we obtain a dechannelizer or FD-demux of the transmitted tributaries. Circular shifts are further inserted between the interpolators/decimators and the FB ports to perform frequency up/down conversions.

The resulting (de)channelizers feature four advantages: 1) low HW complexity, 2) the SB are extracted flat and sharp (brickwall-like filtering per SB, nearly perfect reconstruction), 3) no spectral guardbands required between SBs, and 4) the SBs retain orthogonality.

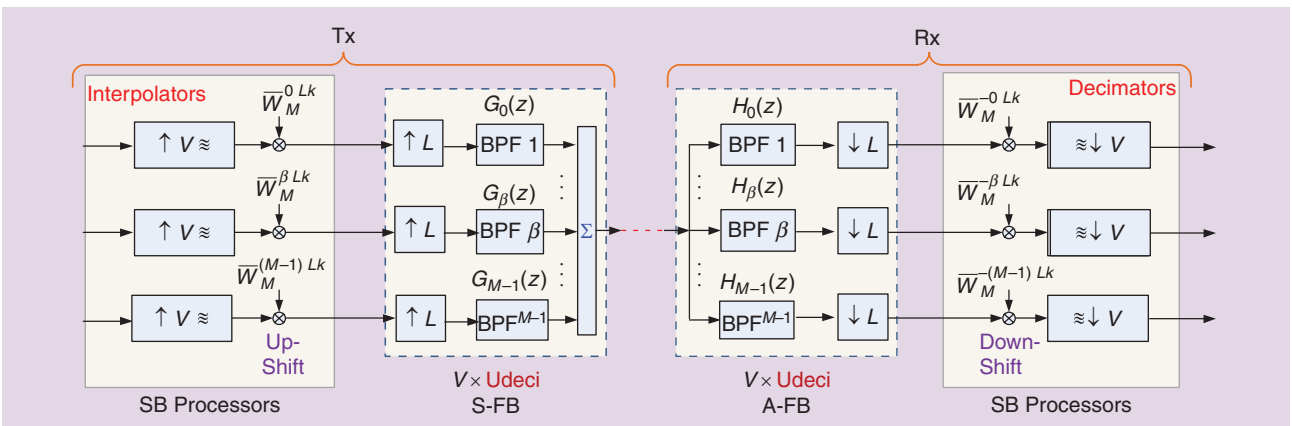
Using our dechannelizers, channel impairments may be equalized on an SB basis. Unlike in FB structures with perfect reconstruction [2], our nearly perfect reconstruction is not based on mutual cancellation of spectral crosstalk among neighboring SBs (which would be impaired by the distortion generated by the transmission channel).

As opposed to [3], our approach is to keep the (de)channelization and per-SB processing functions well separated. Any subsequent processing per SB is to be performed in separate array of SB processors attached to each of the SBs (not shown in Figure 4).

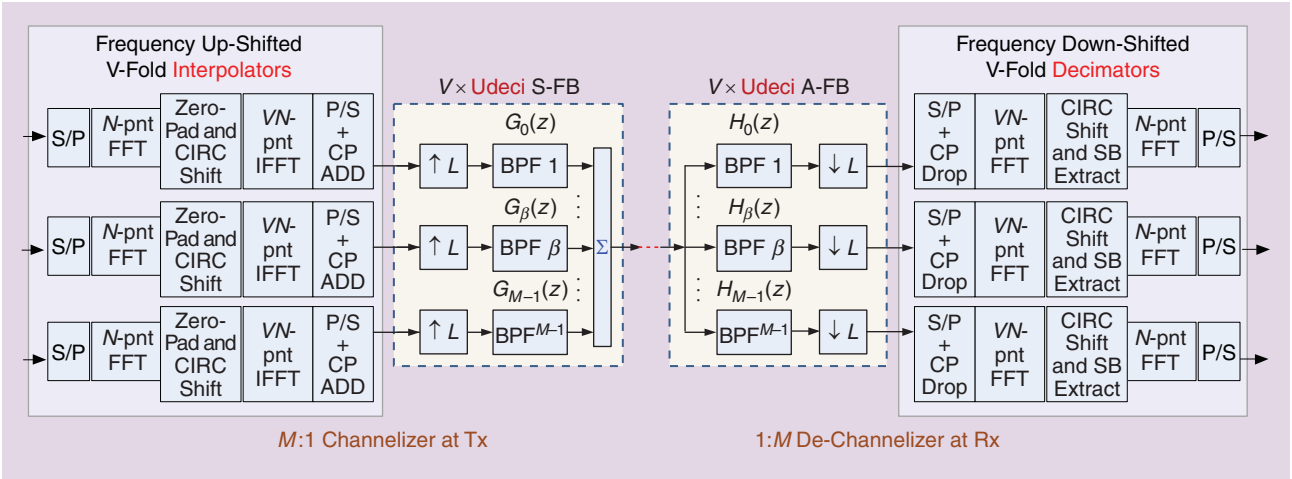
Our principle of operation is that the filtering be partitioned into two tiers:

- coarse prefiltering at the FB level within each of the constituent BPFs is no longer required to be brickwall, but is now relaxed to have mild frequency responses (Figure 6), shaped akin to trapezoids rather than rectangles, flat over the SB extent, S, with mild transitions spanning neighboring S intervals
- sharp pre- and postfiltering in the interpolators/decimators, efficiently realized in the FD by back-to-back (I)FFTs of different sizes, as further explained.

This structure provides (de)channelization at reduced overall computational load relative to a Cdeci FB. Indeed, the extra complexity of the interpolators/decimators (which operate at slower rate) turns out to be just a fraction the complexity



[FIG4] The (de)channelizer with no spectral overhead and with nearly perfect reconstruction. The interpolators and decimators comprise up/downsampling followed/preceded by brickwall low-pass filtering performing image rejection/antialiasing, respectively.



[FIG 5] The Tx-side channelizer and Rx-side dechannelizer with nearly perfect reconstruction based on efficient (I)FFT-based realization of the interpolators/decimators in the (de)channelizer of Figure 4.

reduction gained by replacing the sharp brickwall filters of a Cdeci FB, requiring lots of taps, by milder-sloping filters of the Udeci FB, requiring far fewer taps.

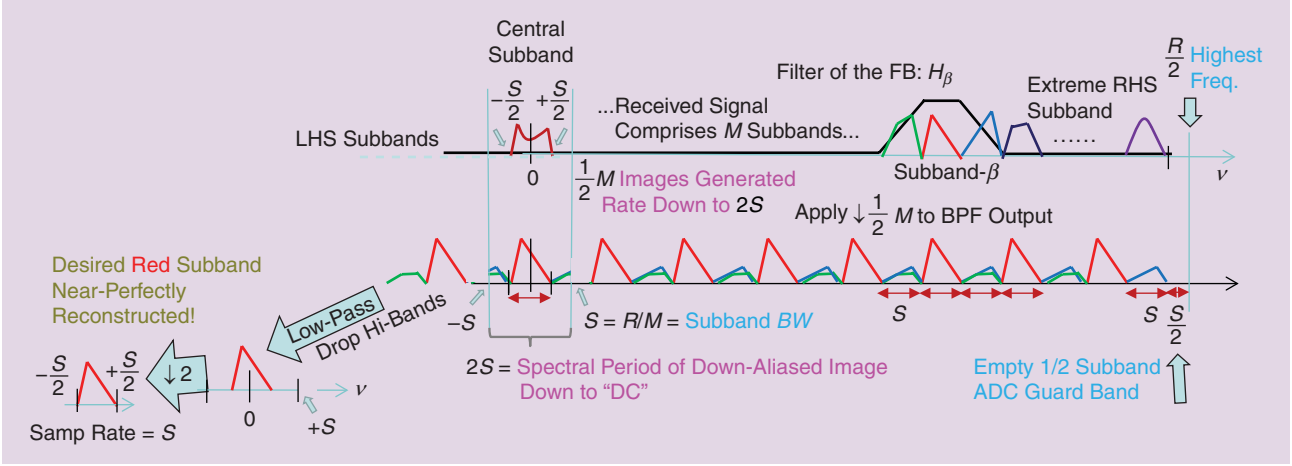
At this point, let us invoke DSP-duality to heretofore exclusively treat Rx-side A-FB structures, as the Tx-side S-FB structure is the DSP-dual of the A-FB, with all signal flows running in reverse. Moreover, as shown in the next section, the S-FB at the Tx is not really necessary but may be replaced by a conventional DFT-S OFDM Tx. Thus, from now on our focus is A-FBs at the Rx.

The structure of Figure 4 is next transformed into the equivalent one of Figure 5, essentially realizing the V-fold decimators in the FD as cascades of an N-pnt DFT a circular shift and a VN-pnt IDFT (and its dual at the Tx). The equivalence is due to decimation amounting to band-limiting low-pass filter (LPF) followed by downsampling. The LPF is performed in the cyclic FD, discarding DFT frequency samples corresponding to the blocked band at the IDFT input. The modulations (frequency up/down shifters) appearing in Figure 4 are realized by the FD circular shifts in Figure 5.

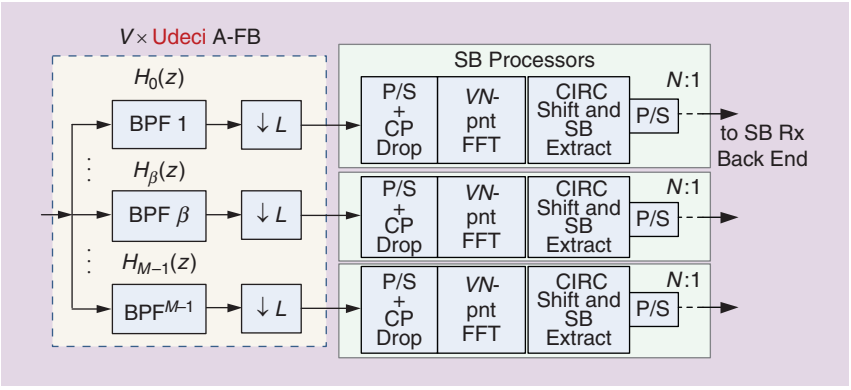
It is useful to analyze the (de)channelizer scheme of Figure 5 (specialized to $V = 2$) in the spectral domain, as depicted in Figure 6. The mild trapezoidal filter, H_β, nearly perfectly passes through over its flat-band the βth SB, along with residual interference from the two neighboring SBs. The figure illustrates how the $L = M/2$-fold decimation generates L spectral images, aliasing the filtered signal all the way down to baseband, where a nearly perfect image of the desired SB is carved out by the twofold decimator. This figure applies to the case of an even SB index, β. In the case of an odd SB index, a similar analysis (not illustrated) would indicate that the twofold decimator should be preceded by a half-band circular shift, essentially picking up the upper (high-pass) spectral half of the downsampled output of the odd-indexed BPF.

THE 2 × UDECI FB BASED DECHANNELIZER APPLIED TO (DFT-S) OFDM TRANSMISSION

In Figure 5 we presented an FB-based (de)channelizer, at the Tx and Rx respectively. To transmit an OFDM channel, multiple



[FIG 6] A spectral analysis of the Rx-side dechannelizer in Figure 5 (for an even SB index, β). The mild trapesoidal-like filter of the FB selects the desired SB along with remnants of the adjacent SBs. The downconversion effect generated by (M/2)-fold downsampling brings down to baseband the desired SB along with interference from the two neighbors. The final low-pass filtering action of the decimator singles out the desired SB. LHS stands for left-hand side.



[FIG 7] **A novel FB-based OFDM Rx obtained by placing onto the output ports of Figure 5 *N*-pnt IFFTs (associated with the OFDM Rx tributaries) and canceling them out versus the *N*-pnt FFTs of the decimators. The V-fold Udeci A-FB is efficiently realized as in Figure 3.**

mini-OFDM tributaries may be launched within each SB. In this case, the Rx side may be simplified (Figure 7). Once multiple SB OFDM Rx-s are added at A-FB outputs in Figure 5 to detect the OFDM tributaries, the N-pnt IFFTs within these SB OFDM Rx-s end up stacked back-to-back with the N-pnt FFTs of the decimators of Figure 5, and mutually cancel out. Thus, for OFDM tributaries, the overall subbanded Rx may be realized as in Figure 7.

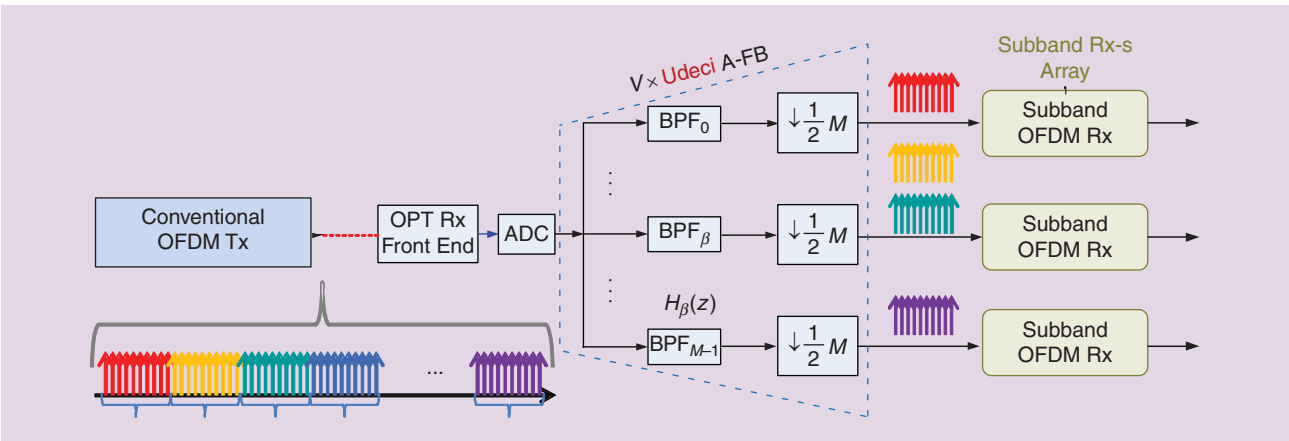
Moreover, it turns out that the Tx-side channelizer is not necessary for OFDM– it may be replaced by a conventional OFDM Tx as shown in Figure 8, wherein each SB carries a subset of OFDM tones—a "mini-OFDM" signal in itself. The overall OFDM spectrum, as generated by a conventional OFDM Tx, essentially an IFFT, may be conceptually viewed as composed of a juxtaposition of contiguous SBs filled up with "mini-OFDM" signals. Nevertheless, the same logic is not extensible to the Rx-side—the OFDM FFT is not equivalent to an A-FB. Indeed, the dispersive delay (CD) over the fiber channel causes OFDM symbols in different SBs to arrive at Rx at different times. The Rx [right-hand side (RHS) in Figure 8] must slice the SBs by means of "true" linear filtering (rather than cyclic filtering as realizable in the DFT domain), by means of an FB followed by an array of SB Rx-s, essentially realizing the structure of Figure 7. The OFDM Rx structure of Figure 8, with an efficient Udeci FB, implements a divide and conquer strategy, performing the "divide" (slicing into SBs) by means of the 2 × Udeci A-FB while the "conquer" is performed in the multiple SB Rx-s, each processing a SB, with reduced BW S, a factor of M lower than the BW R of the full optical channel. The SB Rx-s become extremely simple, even M of them together are more than two times simpler than a full-band Rx. For example, if R = 25 GHz is the channel BW, then for M = 16 SBs we require 16 SB Rx-s—actually 15, as the 16th SB, the outer one, split into two halves at the two extreme ends of the channel, is used as transition region for the ADC—without incurring any spectral efficiency penalty provided the technique of orthogonal band multiplexing (OBM) [8] is used.

(DFT-S) OFDM RECEIVER BASED IN 2 × UDECI FBS—TOP VIEW

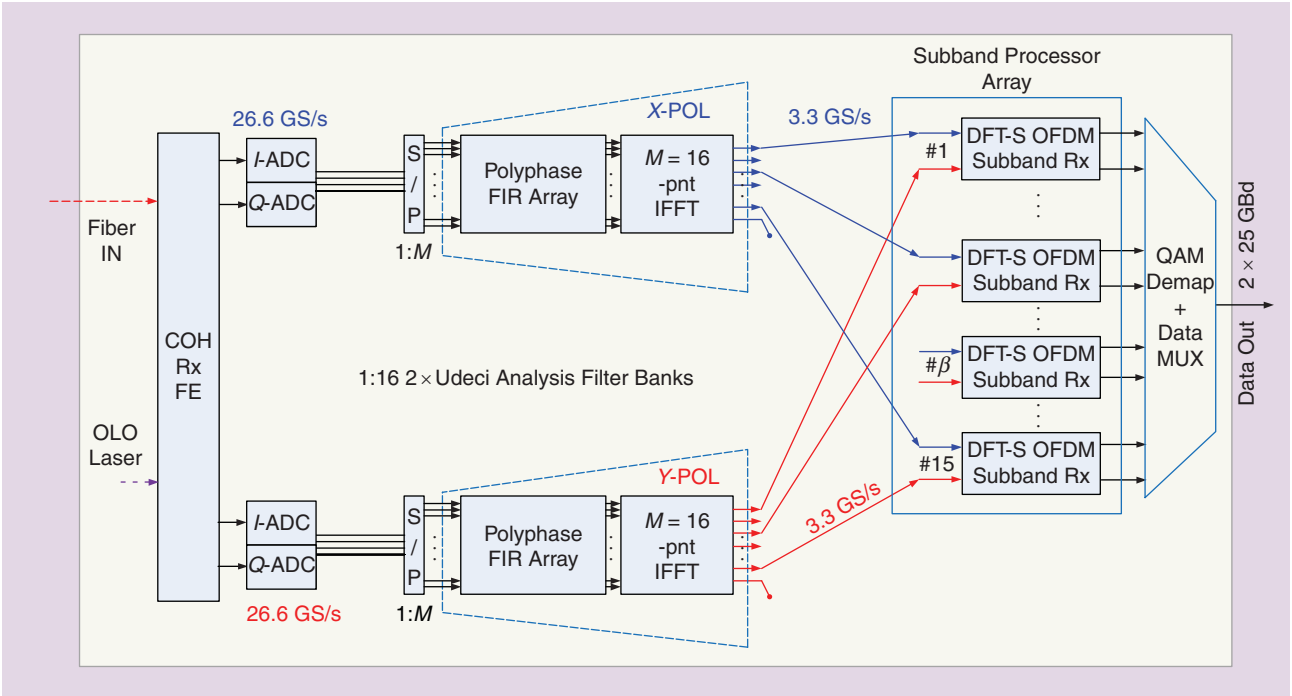
The description above pertained to OFDM over a scalar channel. For coherent optical OFDM, each digitally subbanded Rx must contend with two orthogonal POLs, therefore the top level of the overall Rx (Figure 9) is equipped with a pair of 2xUdeci FBs, slicing out the the X- and Y-POL signals into SBs, sent in pairs to the SB processors, which are actually 2 × 2 MIMO OFDM Rx-s (the βth SB outputs of X- and Y-FBs are routed to the βth SB Rx, with β indexing the M output ports of the two FBs).

SUBBAND PROCESSORS–2 × 2 MIMO DFT-S OFDM RECEIVERS

Figure 10 "zooms" into one of the SB Rx-s, presenting its internals. This is a MIMO OFDM Rx, relaxed to detect an M-times narrower spectral band (S=R/M=25 GHz/15=1.66 GHz rather than 25 GHz in our system, which is substantial relief, despite the multiplicity of 15 such Rx-s). Due to the 2xUdeci nature of



[FIG8] **The OFDM link using a conventional OFDM Tx. The Rx essentially comprises a 2 × Udeci FB followed by an array of SB OFDM Rx-s. The *MN*-tones OFDM symbol generated by the Tx may be conceptually viewed as the spectral juxtaposition of *M* OFDM symbols, each containing *N* tones, each associated with an SB. The analysis FB in the Rx separates out the individual "mini-OFDM" signals and presents them to *M* SB OFDM Rx-s, each one of which processes one SB.**



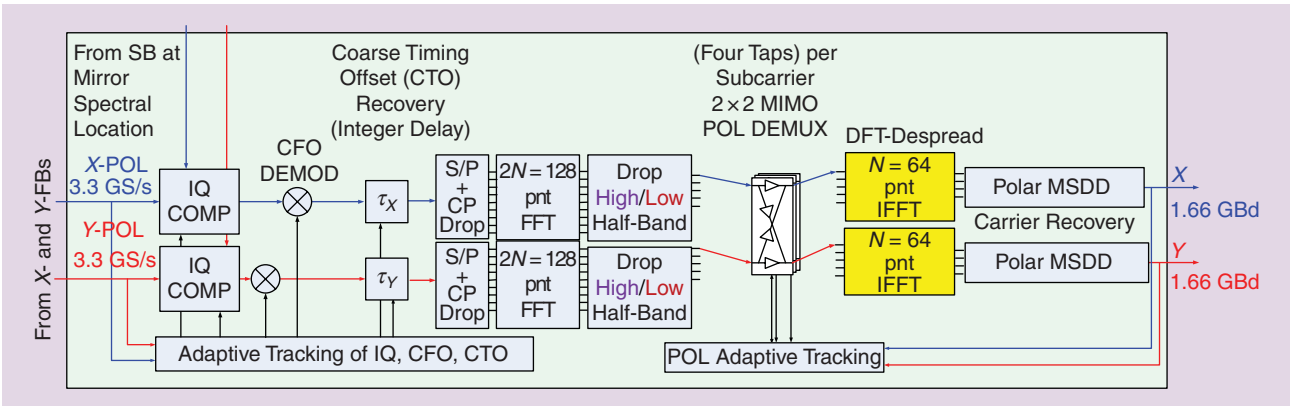
[FIG 9] **A subbanded (DFT-S) OFDM Rx. The top level shows the two FBs for the X- and Y-POLs and the array of SB Rx-s. Notice that each SB Rx is fed by the corresponding SBs from the two 2 × Udeci FBs associated with the two POLs.**

the FB, each SB Rx is twice-oversampled at sampling rate $2S$=3.33 GHz.

The backbone of the SB Rx is the $2N$-FFT and the N-IFFT, forming (for V=2) the V-fold decimator of Figure 5, completing dechannelization with nearly perfect reconstruction, in conjunction with the preceding A-FB cores shown in Figure 9. The extra DSP functions incorporated in the SB Rx-s are IQ imbalance and carrier frequency offset (CFO) mitigation, coarse (integer) and fine (fractional) timing recovery, joint CD+ POL EQZ (2 × 2 MIMO) and carrier phase recovery (CPR). The SB Rx features multiple innovations in these individual DSP modules. In particular 1) a recently introduced polar variant [30] of the multisymbol delay detection (MSDD) CPR with low complexity eliminating all multipliers, 2) in each SB Rx, the CD and POL-demux EQZ is jointly performed with just four taps, whereas in a conventional receiver there would be a full OLS FD EQZ required to mitigate CD, as well as 4 x O(10) taps to mitigate POL rotation and PMD. This underscores the substantial complexity advantage of our FB-based Rx versus a conventional full-band one, and 3) robust IQ imbalance correction scheme acting on SBs in pairs.

TOTAL Rx COMPLEXITY

The total complexity savings for 4,000-km transmission over standard SMF is 57%, as itemized in Figure 11 (*complexity* is defined here as multipliers count in the FPGA). At 2,000 km, also for a MIMO memory duration of 12 samples for a conventional full-band system, the subbanded savings would be 54.5%. Lower savings (16.6% at 2,000 km and 26.5% at 4,000 km)



[FIG 10] **The detail of each of the SB Rx-s in Figure 9. This 2 × 2 MIMO DFT-S OFDM Rx contains a full suite of DSP algorithms for signal conditioning, albeit at M/2 = 8 times reduced rate, relative to the full channel BW+ADC oversampling guardband.**

would be obtained if the MIMO equalization were to be incorporated within the OLS FD EQZ, however this option is still under investigation in the literature, pending a satisfactory solution to training the coefficients of such an EQZ.

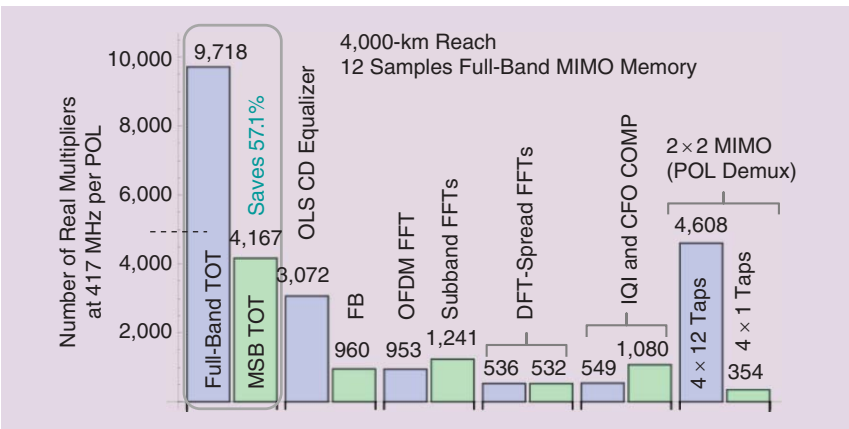
The key rationale for per-SB timing recovery and channel (CD and POL) EQZ is that each SB is practically frequency-flat—the quadratic phase profile of the CD over frequency appears like a sloped straight segment over the narrowband S=1.66 GHz BW of each SB—a linear phase slope means a constant delay per SB (regularly increasing from each SB to the adjacent one). The integer part of the SB delay (measured in SB samples units) may be readily corrected by a digital buffer, whereas the fractional part of the delay and any other residual distortion of the almost frequency-flat SB may be corrected by a single complex tap EQZ. This indicates a substantial advantage in CD equalization complexity for an SB receiver. For an OFDM Rx, the timing recovery function is much simplified relative to a single-carrier Rx, as delay and correlate (D&C) algorithms such as Schmidl–Cox are degraded by CD. Here, the frequency-flat SBs yield improved D&C performance per SB. Moreover, the quite complex parallelization of D&C algorithms [31] is now simplified. The net result is more robust and simpler timing recovery. More generally, in a full-band Rx, control path functions such as channel estimation and adaptive tracking generally take a major toll on complexity and performance. In contrast, the digitally subbanded Rx does not require separate CD estimation and enjoys substantial advantage in all its adaptive algorithms, which converge much faster and more accurately over each frequency-flat SB. The underlying reasons for adaptive advantage are 1) the taps count per SB is very small (typically two per POL) and 2) the eigenvalue spread of the autocorrelation matrix of the received signal is very small here. It is well known in adaptive filtering theory that the eigenvalue spread of matrix controls the speed of convergence of an adaptive filter and that increased frequency ripple or variation of the signal power spectral density is indicative of increased eigenvalue spread. Thus, convergence of adaptive filters in each SB is generally much faster than convergence of an adaptive filter operating in each SB. Rapid and accurate adaptive algorithms convergence means low data-aided overhead.

REAL-TIME HW IMPLEMENTATION OF A 16-QAM DFT-S OFDM RX

The complexity savings and treelike nature of the FB → SB array enabled our setting a record in the real-time FPGA realization of a DFT-S OFDM Rx, albeit fed at this point not by actual ADCs, but from synthetic 16-QAM simulated data played out from a FIFO at full-speed (25 Gbaud / POL × 2 POL, carrying an aggregate 160 Gb/s going through the FB). We used two ML628 boards and layouts of the two XC6HX380T Virtex6 FPGAs, one for the FB and the other for the SB Rx. For now we just implemented a single FB but a full 2×2 MIMO SB Rx out of the $M = 15$ ones (feeding one POL input of the SB Rx from data passing in real time through the FB, whereas the data on the other path is simulated offline, establishing the real-time functionality of both the FB and the MIMO OFDM SB Rx. For demo purposes, a single SB Rx may be used to access and detect any of the 15 SBs of the full 25-GHz optical channel, one at a time (each SB carries (1/15) 160 Gb/s = 10.66 Gb/s). Note that the limitation in accessing all SBs at once is neither of algorithmic DSP nature, nor due to computational inefficiency, but it is rather extraneous—inter-FPGA interconnectivity bottlenecks, which would be irrelevant in actual ASIC implementation. This validates real-time feasibility of our FB-based 2×25 Gbaud 16-QAM Rx.

DISCUSSION AND CONCLUSIONS

We aimed to establish that FB-based digital subbanding is a preferred architecture for high-speed ASICs for optical communication. The subbanding scheme is not analog, but it is rather purely digital, performed after A/D conversion for each individual channel in the WDM multiplex, amounting to a second tier of fine frequency division demultiplexing. Digital subbanding provides benefits similar to those obtained by multiband OFDM techniques [7]–[9]. In those techniques, relatively narrowband bands (3–6 GHz) are photonically generated within a superchannel structure, multiplexed at the Tx and demultiplexed at the Rx, an approach recently increasingly adopted in superhero OFDM superchannel experiments, e.g., [6]. In contrast, our digital subchannel demux requires much simpler, lower-cost, and energy-efficient HW. As our subbanding realization is digital rather than analog, we do away with the cumbersome finely spaced multitone generator, and we eliminate a



[FIG 11] The complexity comparison of a full-band conventional Rx versus a multisubband (MSB) DFT-S OFDM Rx for 4,000-km transmission over SMF and for 12 taps of memory for the conventional POL-demux 2×2 MIMO EQZ. Most MSB savings stem from more efficient CD and 2×2 MIMO (PolDemux) EQZ. The 2xUdeci FB "overhead," which enables these savings, is seen to be just several percent of total Rx complexity. Both systems designed to target the same high spectral efficiency over 2,000-km SSMF: Very low CP spectral overhead of 1.56% = 2/128 = 8/1,024. Full-band DFT-S OFDM Tx (used with both Rx-s) uses 1,024-pnt OFDM symbols and inserts eight samples of CP in each of the MSB subband Rx-s, simply dropping one CP sample every 128 samples. In contrast, the full-band Rx needs heavy CD and adaptive 2×2 MIMO EQZs in the TD, before the OFDM.

large number of DACs, ADCs, modulators, optical filters, and analog optical receivers, which would be customarily used in transmission of a finely spaced (3–6 GHz) analog-generated superchannel. Nevertheless, we enjoy the full benefits of having narrowband frequency-flat SBs, which are now digitally (de) muxed. Furthermore, despite efficiently crowding the multiple 1.6 GHz SBs with zero spectral guardbands, we are nevertheless able to maintain a nearly perfect degree of orthogonality between the individual SBs (i.e., eliminate inter-SB crosstalk), which would have been impossible in a fine-muxed analog/optical-generated superchannel. Finally, by OBM [8] at the OFDM WDM channel level (in our exemplary system between neighboring 25-GHz channels) we are able to further eliminate the interchannel guardbands among the WDM bands.

Another aspect pertains to the effectiveness of the current Udeci subbanded approach to detection of SC modulation. For SC coherent receivers, CD, timing, POL, frequency offset, and other impairments may also be best compensated per SB basis. The A-FB and the array of SB receivers must be complemented by an S-FB or equivalent means assembling the final SC signal out of the individually processed SBs. Such SC subbanded architecture turns out to be more effective than conventional single-carrier structures, however, this topic is deferred to a future publication.

The FB-based receivers presented here were developed in the photonic context but may also be found useful for wireline and wireless electrical transmission.

AUTHORS

Moshe Nazarathy (nazarat@ee.technion.ac.il) received the B.S. (cum laude) and doctor of science degrees in electrical engineering from the Technion–Israel Institute of Technology, Haifa. He is a tenured associate professor with the Electrical Engineering Department of the Technion–Israel Institute of Technology. He cofounded Harmonic, Inc., served as senior vice president of R&D and corporate chief technology officer, and was a member of Harmonic's Board of Directors from 1988 to 2001 and a general manager of the company's Israeli subsidiary. He is also a venture partner with Giza Ventures, Ltd.

Alex Tolmachev (talex@techunix.technion.ac.il) received the B.S. degree in electrical engineering and the B.A. degree in physics (both summa cum laude) from the Technion–Israel Institute of Technology, Haifa, in 2004. He is currently working toward the Ph.D. degree in the Technion Electrical Engineering Department in the research area of coherent optical communication focusing on broadband transmission methods and mitigation of impairments by advanced DSP techniques.

REFERENCES

[1] P. Vaidyanathan, "Filter banks in digital communications," *IEE Circuits Syst. Mag.*, vol. 1, no. 2, pp. 4–25, 2001.

[2] Z. Cvetkovic and M. Vetterli, "Oversampled filter banks," *IEEE Trans. Signal Processing*, vol. 46, no. 5, pp. 1245–1255, May 1998.

[3] M. Awan, Y. Le Moullec, P. Koch, and F. Harris, "Polyphase filter banks for embedded sample rate changes in digital radio front-ends," *ZTE Commun.*, vol. 8, no. 4, pp. 3–9, 2011.

[4] X. Liu and M. Nazarathy, "Coherent, self-coherent and differential detection," in *Impact of Nonlinearities on Fiber Optic Communications*. New York: Springer, 2011, Chap. 1.

[5] S. L. Jansen, "OFDM for optical communications," in *Proc. Optical Fiber Communication Conf.* (*OFC'10*), 2010, p. SC341.

[6] W. Shieh and I. Djordjevic, *OFDM for Optical Communications*. Amsterdam, The Netherlands: Elsevier, 2010.

[7] Y. Tang, W. Shieh, and B. S. Krongold, "DFT-spread OFDM for fiber nonlinearity mitigation," *IEEE Photon. Technol. Lett.*, vol. 22, no. 16, pp. 1250–1252, Aug. 2010.

[8] W. Shieh, Q. Yang, and Y. Ma, "107 Gb/s coherent optical OFDM transmission over 1000-km SSMF fiber using orthogonal band multiplexing," *Opt. Express*, vol. 16, no. 9, pp. 6378–6386, Apr. 2008.

[9] X. Liu and S. Chandrasekhar, "High spectral-efficiency transmission techniques for beyond 100-Gb/s systems, SPMA1," in *SPPCom—Signal Processing in Photonic Communications—OSA Techn. Dig.*, 2011, pp. 1–36.

[10] S. Kilmurray, T. Fehenberger, P. Bayvel, and R. I. Killey, "Comparison of the nonlinear transmission performance of quasi-Nyquist WDM and reduced guard interval OFDM," *Opt. Express*, vol. 20, no. 4, pp. 4198–4208, 2012.

[11] R. Schmogrow, M. Winter, M. Meyer, D. Hillerkuss, S. Wolf, B. Baeuerle, B. Nebendahl, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, W. Freude, and J. Leuthold, "Real-time Nyquist pulse generation beyond 100 Gbit/s and its relation to OFDM," *Opt. Express*, vol. 20, no. 1, pp. 317–337, 2012.

[12] G. Shulkind and M. Nazarathy, "An analytical study of the improved nonlinear tolerance of DFT-spread OFDM and its unitary- spread OFDM generalization," *Opt. Express*, vol. 20, no. 23, pp. 25884–25901, 2012.

[13] K.-P. Ho, "Subband equaliser for chromatic dispersion of optical fibre," *Electron. Lett.*, vol. 45, no. 24, p. 1224, 2009.

[14] X. Liu, P. Winzer, C. Sethumadhavan, S. Randel, and S. Corteselli, "Multiband DFT-spread-OFDM equalizer with overlap-and-add dispersion compensation for low-overhead and low-complexity channel equalization," in *Proc. Optical Fiber Communication Conf.* (*OFC'13*), 2013, vol. 2, no. 2, p. OW3B.2.

[15] L. B. Du and A. J. Lowery, "Mitigation of dispersion penalty for short-cyclic-prefix coherent optical OFDM systems," in *Proc. European Conf. and Exhibition on Optical Communication* (*ECOC'10*), 2010, p. Tu.4.A.5.

[16] A. Tolmachev and M. Nazarathy, "Real-time-realizable filtered-multi-tone (FMT) modulation for layered-FFT Nyquist WDM spectral shaping," in *Proc. European Conf. and Exhibition on Optical Communication* (*ECOC'11*), 2011, paper We.10.P1.65.

[17] A. Tolmachev and M. Nazarathy, "Filter-bank based efficient transmission of reduced-guard-interval OFDM," *Opt. Express*, vol. 19, no. 26, pp. 370–384, 2011.

[18] A. Tolmachev and M. Nazarathy, "Low-complexity multi-band polyphase filter bank for reduced-guard-interval coherent optical OFDM," in *Proc. Signal Processing in Photonic Communications (SPPCom), Advanced Photonics OSA Conf.*, 2011, p. SPMB3.

[19] A. Tolmachev, M. Nazarathy, I. Tselniker, and I. Sigron, "Oversampled digital filter banks simplify and improve signal processing in RGI OFDM receivers," in *Proc. Optical Fiber Communication Conf.* (*OFC'12*), 2012, p. OM2H.3.

[20] M. Nazarathy and A. Tolmachev, "Digital subbanding—A signal processing architecture radically improving OFDM coherent optical receivers," in *Proc. SPPCom—Signal Processing in Photonic Communications—OSA Tech. Dig.*, June 2012.

[21] M. Nazarathy, A. Tolmachev, and S. Ben-Ezra, "Subbanding DSP for flexible optical transceivers," in *Proc. Int. Conf. Transparent Optical Networks (ICTON)*, 2012, no. 1.

[22] M. Nazarathy and A. Tolmachev, "Filter-bank based digital subbanding ASIC architecture for coherent optical receivers," in *Proc. Photonics West 2013*, 2013, pp. 8647–17.

[23] A. Tolmachev, M. Orbach, M. Meltsin, R. Hilgendorf, Y. Birk, and M. Nazarathy, "Real-time FPGA implementation of efficient filter-banks for digitally subbanded coherent DFT-S OFDM receiver," in *Proc. Optical Fiber Communication Conf.* (*OFC'13*), 2013, p. OW3B.1.

[24] F. J. Harris, C. Dick, and M. Rice, "Digital receivers and transmitters using polyphase filter banks for wireless communications," *IEEE Trans. Microwave Theory Tech.*, vol. 51, no. 4, pp. 1395–1412, 2003.

[25] F. J. Harris, *Multirate Signal Processing for Communication Systems*. Englewood Cliffs, NJ: Prentice Hall, 2004.

[26] B. Farhang-Boroujeny, "OFDM versus filter-bank multicarrier," *IEEE Signal Processing Mag.*, vol. 92, pp. 92–112, May 2011.

[27] S. Randel, A. Sierra, X. Liu, S. Chandrasekhar, and P. Winzer, "Study of multicarrier offset-QAM for spectrally efficient coherent optical communications," in *Proc. European Conf. and Exhibition on Optical Communication* (*ECOC 2011*), p. Th.11.A.1.

[28] S. Randel, S. Corteselli, S. Chandrasekhar, A. Sierra, X. Liu, and P. J. Winzer, "Generation of 224-Gb/s multicarrier offset-QAM using a real-time transmitter," in *Proc. Optical Fiber Communication Conf.* (*OFC 2012*), vol. 30, p. OM2H.2.

[29] I. Slim, A. Mezghani, L. G. Baltar, J. Qi, F. N. Hauske, and J. A. Nossek, "Delayed single-tap frequency-domain chromatic-dispersion compensation," *Photon. Technol. Lett.*, vol. 25, no. 2, pp. 167–170, 2013.

[30] A. Tolmachev, I. Tselniker, M. Meltsin, I. Sigron, and M. Nazarathy, "Efficient multiplier-free FPGA demonstration of for high performance phase recovery of 16-QAM," in *Proc. Optical Fiber Communication Conf.* (*OFC 2013*).

[31] Q. Yang, N. Kaneda, X. Liu, S. Chen, W. Shieh, and Y. K. Chen, "Digital signal processing for multi-gigabit real-time OFDM," in *Proc. Signal Processing in Photonic Communications (SPPCom'11), Advanced Photonics OSA Conf.*, 2011, vol. 1, p. SPMB2.

[SP]

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

Alan Pak Tao Lau, Yuliang Gao, Qi Sui, Dawei Wang, Qunbi Zhuge, Mohamed H. Morsy-Osman, Mathieu Chagnon, Xian Xu, Chao Lu, and David V. Plant

Advanced DSP Techniques Enabling High Spectral Efficiency and Flexible Transmissions

Toward elastic optical networks

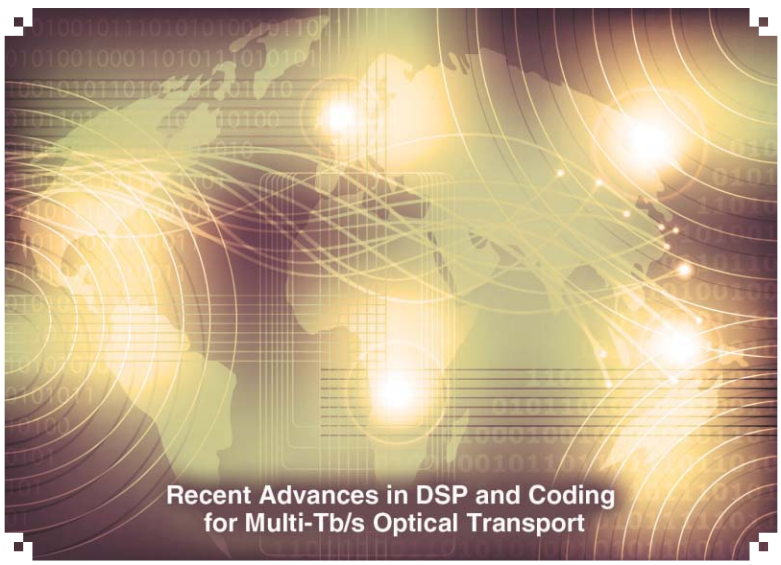


IMAGE LICENSED BY INGRAM PUBLISHING /JANMIKS

In this article, we describe advances in digital signal processing (DSP) techniques that enable Tb/s transmission, and software-defined flexible transponders that support adaptive modulation formats and elastic optical networks (EONs).

INTRODUCTION

Over the past decade, advances in optical hardware such as external modulators, narrow-linewidth lasers and 90° hybrids have enabled DSP algorithms from wireless/copper-wire communications [1] to be applied in coherent optical communication systems. Coherent transceivers supporting 100 Gb/s per channel using polarization-multiplexed quadrature phase-shift keying (QPSK) were announced by system vendors in 2009 and subsequently deployed in transmission systems several years later. DSP units such as chromatic dispersion (CD) compensation, polarization-mode dispersion (PMD) compensation, timing phase recovery (TPR), frequency offset estimation (FOE), as well as carrier phase estimation (CPE) are standard in commercial coherent receivers. Moving forward, 400 Gb/s and 1 terabit per second (Tb/s) per-channel transmission using 16-QAM and above were demonstrated and commercialized in 2012. The next generation of hardware [e.g., digital-to-analog converters (DACs)] and software (e.g., programmability) advances will enable EONs, networks that are capable of adapting to network conditions to maximize overall efficiencies. One of the key

Digital Object Identifier 10.1109/MSP.2013.2287021
Date of publication: 12 February 2014

1053-5888/14/$31.00©2014IEEE

components in EONs is the adaptive/flexible receiver DSP units supporting multiple baud rates, bandwidths, and modulation formats. In addition, imperfections in both the transmit and the local oscillator lasers combined with the requirement to perform signal processing at tens of GHz impose unprecedented constraints and challenges on DSP units for 16-QAM formats and above. As a result, DSP research for coherent optical communications is becoming more divergent from traditional wireless communications. In this article, we review recent DSP developments for high spectral efficiency coherent optical transmission and discuss ongoing efforts to realize adaptive DSP platforms for future EONs.

ONE OF THE KEY COMPONENTS IN EONS IS ADAPTIVE/FLEXIBLE RECEIVER DSP UNITS SUPPORTING MULTIPLE BAUD RATES, BANDWIDTHS, AND MODULATION FORMATS.

DSP DEVELOPMENTS FOR OPTICAL COMMUNICATIONS BEYOND 100 Gb/s

Polarization diverse coherent detection, along with DSP, started the era of electronic signal processing in coherent optical communications and significantly increased transmission capacities over thousands of kilometers at 100 Gb/s. Fundamental signal processing philosophies and building blocks used in wireless/copper-wire communications are utilized in first generation 100 Gb/s PM-QPSK transponders. As a brief introduction, let $E_{t,x(y)}(t)$ denote the transmitted signal in the x- (y-)polarization, which is given by

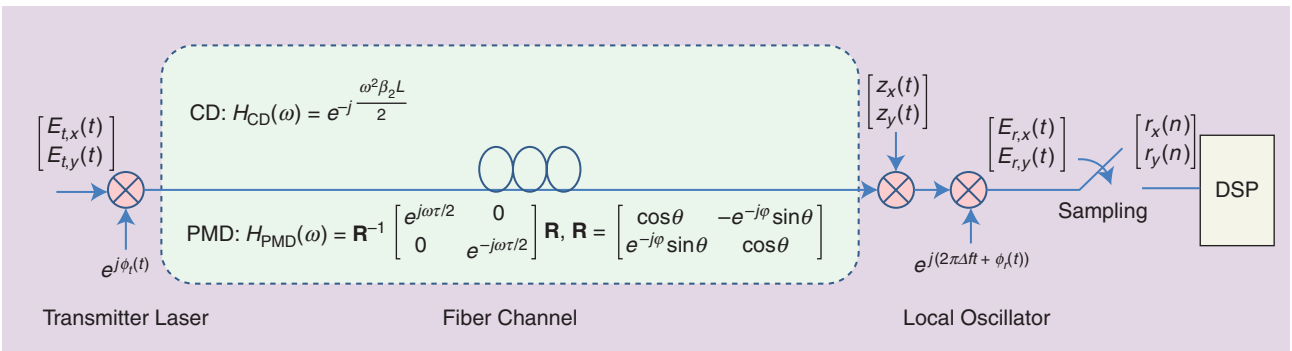
$$E_{t,x(y)}(t) = \sum_k s_{k,x(y)} p(t - kT), \tag{1}$$

where $s_{k,x(y)}$ are the kth information symbols in the x- (y-) polarization, $p(t)$ is the pulse shape, and T is the symbol period, respectively.

Figure 1 depicts the basic model for a long-haul fiber-optic channel consisting of major transmission impairments including: 1)CD $H_{\mathrm{CD}}(\omega)$; 2) PMD (characterized by the first-order PMD matrix $H_{\mathrm{PMD}}(\omega)$]; and 3) amplified spontaneous emission (ASE) noise $z_x(t)$, $z_y(t)$ from inline optical amplifiers, which are typically modeled as complex circularly symmetric zero-mean additive white Gaussian random process. The PMD model shown here is known as first-order PMD in which the group delays of the two polarization modes are independent of frequency to the first order. The more general effect of all-order PMD are extensively studied in the literature but will be omitted here since the first-order model essentially captures the physics of PMD, and blind nonparametric adaptive DSP are typically used to compensate PMD regardless of its analytical model anyway. Practical laser imperfections include phase noise $\phi_t(t)$ and $\phi_r(t)$ of the transmitter laser and local oscillator (LO), respectively and frequency offset Δf between the lasers. Transmitter and LO phase noise are modeled as a Wiener processes [2] in which $\phi_{t(r)}(t_2) - \phi_{t(r)}(t_1)$ is a zero-mean Gaussian random variable with variance $2\pi\Delta v|t_2 - t_1|$ where Δv is the laser linewidth. In this case, the received signal is given by

$$\begin{bmatrix} E_{r,x}(t) \\ E_{r,y}(t) \end{bmatrix} = F^{-1}\left\{ H_{\mathrm{CD}}(\omega) H_{\mathrm{PMD}}(\omega) \begin{bmatrix} \tilde{E}_{t,x}(\omega) \\ \tilde{E}_{t,y}(\omega) \end{bmatrix} \right\} \cdot e^{\left(j2\pi\Delta f t + \phi_t(t) + \phi_r(t)\right)} + \begin{bmatrix} z_x(t) \\ z_y(t) \end{bmatrix}, \tag{2}$$

where $\tilde{E}_{t,x(y)}(\omega)$ is the Fourier transform of $E_{t,x(y)}(t)$ and F^{-1} denotes inverse Fourier transform. Note that the formulation here is an approximation as the transmitter phase noise $e^{j\phi t(t)}$ also propagates through the dispersive fiber channel, thus creating additional impairments. Nonetheless, this can be compensated by receiver DSP with no performance penalty. The fiber CD is a static impairment while PMD is time varying and is caused by random mechanical perturbations and temperature



[FIG1] A canonical channel model for polarization-multiplexed long-haul fiber-optic communication systems with digital coherent optical receivers. The transmitted signal in the $x-(y-)$ polarization is $E_{t,x}(t)$ and $E_{t,y}(t)$ and major transmission impairments include CD $H_{\mathrm{CD}}(\omega)$, PMD $H_{\mathrm{PMD}}(\omega)$, ASE noise $z_x(t)$, $z_y(t)$, as well as laser imperfection-induced impairments such as carrier frequency offset $e^{j2\pi\Delta ft}$ between transmitter laser and local oscillator and laser phase noise $e^{j\phi_t(t)}, e^{j\phi_r(t)}$. In this formulation, L, β_2, τ denote fiber length, group velocity dispersion coefficient, and differential group delay between two polarization modes, respectively, and θ, φ are angles relating the input signal's state of polarization to the principal states of polarization of the fiber.

fluctuations to the fiber at random locations. Several DSP techniques can be borrowed from the wireless communications community and applied to digital coherent optical receivers. Consequently, and as depicted in Figure 2, RF front-end corrections followed by resampling, CD compensation, TPR, polarization demultiplexing/PMD compensation, FOE, CPE and forward error correction (FEC) represent the standard digital coherent optical receiver DSP platform.

As we move from QPSK to higher-order QAM formats (e.g., 16-QAM) and commensurate baud rates, several impairments impose unique and extremely stringent constraints on system performance that are unforeseen in wireless/copper-wire systems. This situation has led to a series of new challenges in DSP research for long-haul coherent optical systems that become somewhat more fiber-optic specific including:

1) Current/future wireless LAN and cellular systems use OFDM with pilot subcarriers. OFDM has a higher spectral efficiency, requires simpler channel and timing estimation, and is ideal for frequency-selective channels. However, except for filtering limitations created by wavelength selective switches (WSSs) and/or DACs/analog-to-digital converters (ADCs), the amplitude transfer function of the fiber-optic channel is largely frequency-flat. Therefore, the fundamental advantage of OFDM cannot be leveraged to its full extent. While the debate between single-carrier and OFDM systems goes on, current research and commercial transponders for long-haul systems are focusing on single-carrier transmission systems.

2) The algorithms to appropriately compensate fiber nonlinearity and achieve information-theoretic capacities are highly sought after. Recent work, including digital back-propagation, improves transmission performance to some extent but the high-computational complexity prohibits their practical use at present. For DSP techniques to mitigate fiber nonlinearity, the interested readers are referred to other articles included in this issue for more details.

3) The oscillator phase noise for optical systems (laser phase noise) is much larger than that of wireless systems.

4) The sampling rate for optical systems is of the order of tens of giga-samples per second while current application-specific integrated circuit (ASIC) technologies support serial processing speeds of approximately 1 GHz [3]. Therefore, while massive algorithmic/hardware parallelization is required, this approach can exacerbate laser phase noise-induced impairments in feedforward as well as feedback algorithms. Algorithms and hardware structures need to be optimized to reduce such parallelization penalties.

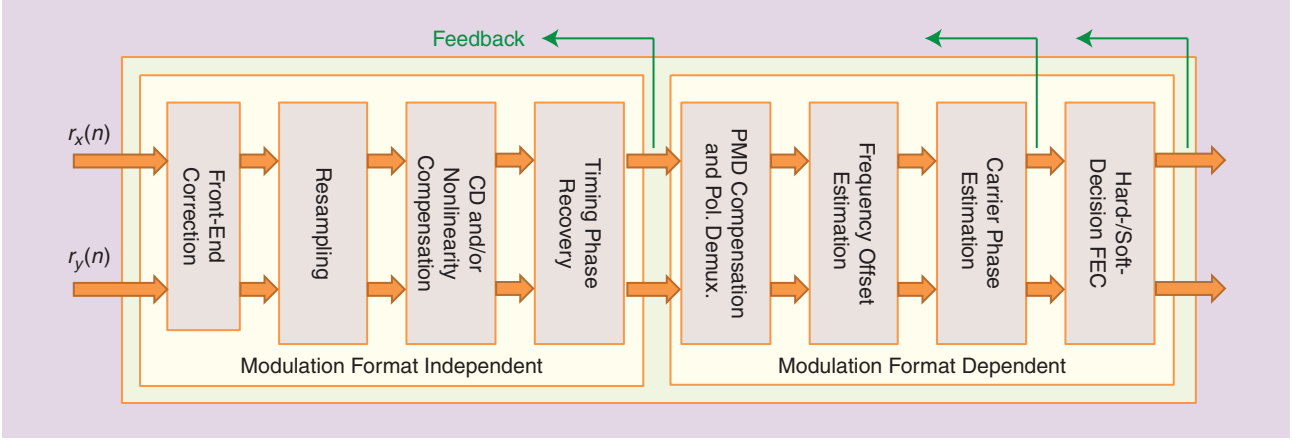
5) To meet transponder form factor and power consumption requirements, DSP algorithms must have low complexity. This constraint becomes more challenging as we aim to develop a universal receiver DSP platform applicable to multiple single-carrier modulation formats for adaptive/flexible transponders.

HIGH-PERFORMANCE, LOW-COMPLEXITY ADVANCED ADAPTIVE DSP TECHNIQUES HAVE BECOME MAJOR FOCUS AREAS OF RESEARCH FOR HIGHER-ORDER QAM FORMATS SUPPORTING 200 Gb/s PER CHANNEL AND BEYOND.

Owing to these unique challenges and over the past few years, high-performance, low-complexity advanced adaptive DSP techniques have become major focus areas of research for higher-order QAM formats supporting 200 Gb/s per channel and beyond.

CARRIER PHASE ESTIMATION FOR HIGH SPECTRAL EFFICIENCY TRANSMISSIONS

In this section we discuss CPE, which is an important component in coherent optical communications. After the preceding DSP that both compensates for other impairments and downsamples to the symbol rate, let the kth symbol in one polarization going into the CPE unit be $r_{\mathrm{CPE}}(k) = s_k \cdot e^{j\phi(k)} + z(k)$, where



[FIG2] A standard DSP subsystems configuration in a digital coherent optical receiver. Different variants of DSP algorithms proposed in the literature may combine multiple subsystems together through feedback signaling and/or training symbols.

$\phi(k) = \phi_t(k) + \phi_r(k)$ is the combined transmitter and receiver phase noise and $z(k)$ is the additive ASE noise. In wireless communication systems, phase synchronization is typically achieved by a phase-lock loop (PLL) employing a one-tap least-mean square (LMS) filter to implement decision-directed (DD) phase estimation in a feedback manner. Feedback techniques rely on calculating the phase estimate $\phi(k)$ from $r_{\mathrm{CPE}}(k)$ and the symbol decision $\hat{s}_k$, and use this information as the initial estimate for $\phi(k+L)$, where L denotes the total feedback delay. L is determined by the degree of parallelization P and the pipelining (number of steps in an algorithm). Unfortunately, in high-speed coherent optical communications systems, required parallelization may create large feedback delays and considerably limit the performance of feedback algorithms. One would need either feedforward algorithms that avoid feedback loops and/or the development of intelligent algorithmic and hardware structures that can somehow reduce P and/or L.

For feedforward CPE, the Viterbi-Viterbi phase estimation (VVPE) is most commonly used for QPSK systems [4]. VVPE raises a block of the received signals $r_{\mathrm{CPE}}(k)$ to the fourth power to remove the QPSK modulation and calculate the phase noise estimate by $\hat{\phi}(k) = (1/4)\arg\left\{\sum_{n=k-N}^{k+N} r_{\mathrm{CPE}}^4(n)\right\}$. For higher-order QAM, the most widely known feedforward CPE is the blind-phase search (BPS) algorithm [5]. BPS originates from general synchronous communication systems, and the basic idea is to rotate the input signals by M test carrier phase angles $\zeta_m = (m - 1/M)\cdot(\pi/2), m\epsilon\{1,2,\ldots,M\}$. Then, for a given ζ_m, one computes the squared distance between test symbols and closest constellation point over $2N$ consecutive symbols as the error signal

$$e_m = \sum_{n=k-N}^{k+N} |r_{\mathrm{CPE}}(n)\exp(-j\zeta_m) - \Gamma(r_{\mathrm{CPE}}(n)\exp(-j\zeta_m))|^2, \quad (3)$$

where $\Gamma(\cdot)$ denotes decision. Subsequently, the phase noise is estimated by searching for the phase ζ_m that minimizes e_m. BPS is blind, feedforward, and applicable to arbitrary QAM formats but it has large algorithmic and hardware complexity. Such complexity can be somewhat lowered by reducing the number of "test phases" using multistage approaches [6].

Another CPE approach for 16-QAM systems is to realize that the signals in the inner and outer circle of a 16-QAM constellation can actually be used in the standard VVPE as a coarse phase estimate. Identifying these symbols for VVPE is known as QPSK partitioning [7], [8]. A second-stage of maximum likelihood (ML) phase estimate can further refine the estimation accuracy. The overall QPSK partitioning + ML technique gives the same linewidth tolerance compared with single-stage BPS while the computation complexity can be reduced by a factor of five [8]. Various aspects of QPSK partitioning are further optimized to reduce complexity and improve linewidth tolerance. However, for 64-QAM signals and above, single- or two-stage BPS based-algorithms are still more advantageous and feasible at this point.

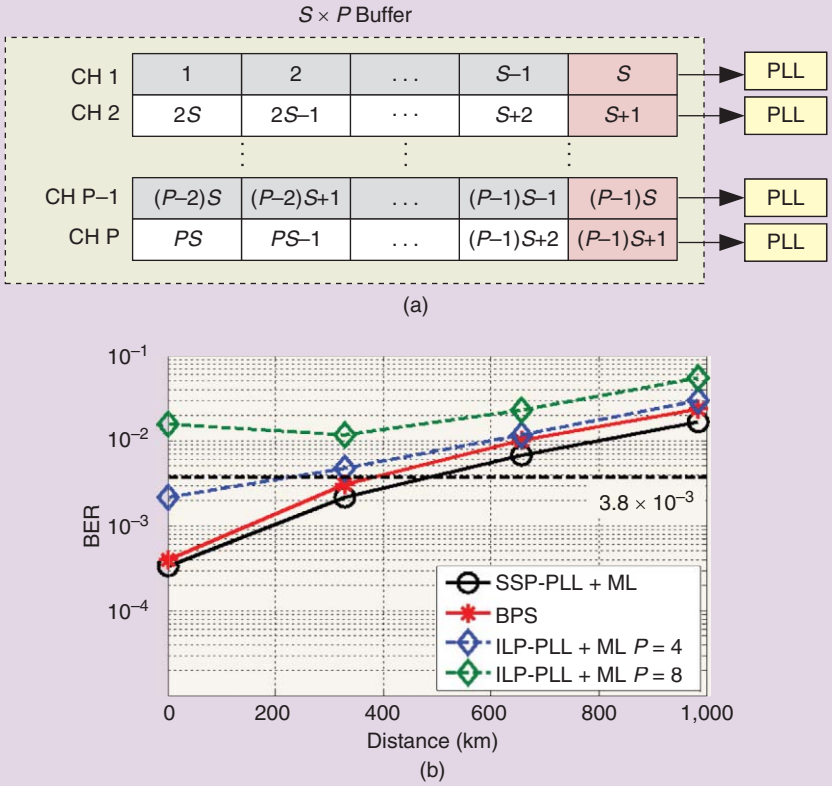
In addition to feedforward techniques, advanced feedback CPE based on decision feedback from a block of previous symbols has also been investigated in recent years [9], [10]. Essentially, multiple decision feedback produces independent phase noise estimates across the block of symbols that can be processed together to improve the estimation accuracy. These techniques show similar performance to BPS, and Tolmachev et al. have recently demonstrated a real-time multiplier-free multi-symbol delay detection (MSDD) at 25 Gbaud [11].

It should be noted that any CPE technique is susceptible to a hardware parallelization penalty regardless of whether it is feedforward or feedback. In [12], superscalar parallelization (SSP)-based CPE was proposed to alleviate its impact at a cost of a very large buffer. The basic idea is to avoid the conventional phase-noise parallelization penalty using a buffer to store the received symbols and then rearrange the order to have consecutive symbols in each block as shown in Figure 3. By doing so, the feedback delay is significantly reduced. Note that pilot symbols are required to initialize the PLL for each block and the pilot symbols can be shared between two blocks thus reducing the overhead. Moreover, polyblock parallelization [9] can also be used for serial-parallel data shuffling. We later proposed an improved SSP-based-PLL combined with a subsequent ML stage (denoted as SSP-PLL + ML) in [13], and experimentally compare its performance with a standard PLL + ML (also known as interleaving parallelization (ILP)-PLL + ML) and BPS and the results are shown in Figure 3(b) for 16-QAM transmission. With one external cavity laser (ECL) at the transmitter and one 2.6 MHz linewidth distributed feedback (DFB) laser at the receiver, the proposed SSP-PLL + ML enables a 25% increase in transmission distance with respect to BPS. In addition, the proposed SSP-PLL + ML can reduce the number of real multipliers, real adders, and slicers by a factor of 12 ~ 25, 18~36, and 16 ~ 32 compared with BPS, respectively. Compared to the original SSP-PLL in [12], where a block length (proportional to the required buffer size) larger than 512 symbols is used, in our SSP-PLL + ML we reduce the block length to 100 symbols.

Finally, most CPE techniques suffer from cycle slips in which the CPE outputs are phase rotated by integer multiples of $\pi/2$ and lead to catastrophic detection errors. To alleviate the problem of cycle slips, one can use differential encoding (DE) where information is encoded in the difference between neighboring bits/symbols. However, DE considerably increases the noise power and hence the optical signal-to-noise ratio (OSNR) penalty for systems using soft-decision-FEC (SD-FEC) codes. Alternative techniques include the insertion of pilot tones or pilot symbols at regular intervals so that an absolute phase reference is readily available for CPE. Recently, we proposed to examine the sliding average of twice estimated phase noise regardless of modulation format and choice of CPE and locate its minimum as a blind and modulation-format-independent cycle-slip detection and correction technique [14].

ADAPTIVE SIGNAL PROCESSING

In 100-Gb/s PM-QPSK systems, adaptive filtering techniques are mainly applied to equalize PMD and other



[FIG3] (a) The superscalar buffer structure for parallelized PLL. (b) BER versus distance for 28-Gbaud 16-QAM with one ECL at the transmitter and one 2.6-MHz linewidth DFB laser at the receiver.

transceiver imperfections such as impedance mismatching that are time-varying and difficult to explicitly measure. A popular adaptive technique is the LMS algorithm. As a brief review, for input vector $\mathbf{u}$ to an adaptive filter $\mathbf{W}$ with output

$$\mathbf{v} = \begin{bmatrix} w_{xx} & w_{xy} \\ w_{yx} & w_{yy} \end{bmatrix} \mathbf{u} = \mathbf{W}\mathbf{u}, \tag{4}$$

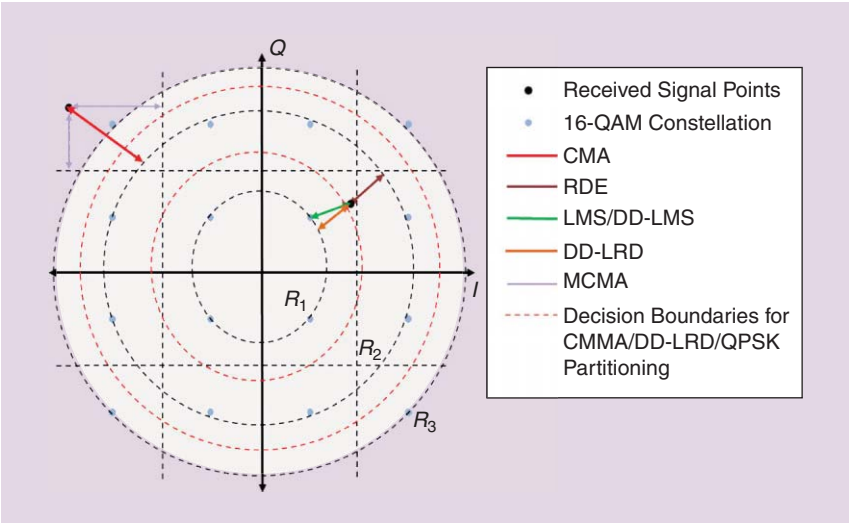
the filter $\mathbf{W}$ aims to minimize the mean squared error between the equalizer output $\mathbf{v}$ and the training symbols $\mathbf{s} = [s_x s_y]^T$. This is characterized by the cost function $J(\mathbf{W}) = \mathrm{E}[|\mathbf{v} - \mathbf{s}|^2]$ and the optimization is achieved through stochastic gradient descent methods. In particular, initializing the filter as $\mathbf{W}^{(0)}$, the $n+1$th iteration is given by $\mathbf{W}^{(n+1)} = \mathbf{W}^{(n)} - \mu(\partial J(\mathbf{W}))/(\partial \mathbf{W}^*) = \mathbf{W}^{(n)} - \mu(\mathbf{v}^{(n)} - \mathbf{s})\mathbf{u}^{(n)H}$, where μ is the step size. A larger step size speeds up convergence while a smaller step size allows better steady-state operation. The DD-LMS can be implemented by replacing $\mathbf{s}$ with symbol decisions $\hat{\mathbf{s}}$.

Despite its popularity, the LMS algorithm requires training symbols that reduce the net transmission rate and is sensitive to phase noise. In long-haul coherent systems, the constant modulus algorithm (CMA) [2] is typically used for PMD compensation and polarization demultiplexing as it is blind and insensitive to laser phase noise and frequency offset. In particular, the CMA aims to minimize the squared difference between the output power and the nominal power P_0. Such difference is characterized by the cost function $J(\mathbf{W}) = \mathrm{E}[(P_0 - |v_1|^2)^2 + (P_0 - |v_2|^2)^2]$ and the corresponding update algorithm is given by $\mathbf{W}^{(n+1)} = \mathbf{W}^{(n)} - \mu(\partial J(\mathbf{W}))/(\partial \mathbf{W}^*) = \mathbf{W}^{(n)} + \mu(P_0 - |\mathbf{v}^{(n)}|^2)\mathbf{v}^{(n)}\mathbf{u}^{(n)H}$.

CMA is subject to the singularity problem in which the two CMA outputs become identical. In this case, $\mathbf{W}$ will have rank one and zero determinant as opposed to being unitary (since PMD effects can be modeled as a unitary matrix). One way to overcome the singularity is to monitor its determinant $\det(\mathbf{W})$ and reset the filter coefficients when $\det(\mathbf{W}) \to 0$. A more common approach is to have a two-stage structure (either two stages in time or in hardware). In the first stage, w_{xx}, w_{xy} are updated until convergence is achieved. Then they are used to initialize w_{yx}, w_{yy} using the unitary property $w_{yx} = -w_{yy}^*, w_{yy} = w_{xx}^*$ and the whole $\mathbf{W}$ is further updated in the second stage until convergence.

As we move to higher-order QAM systems (e.g., 16-QAM), different variants/extensions of the CMA algorithm using cost functions that are tailored to multimodulus formats are proposed. These include the constant multimodulus algorithm (CMMA) [15], the radius-directed equalizer (RDE) [2], the DD-least radius distance (LRD) algorithm [16], and the modified-CMA (MCMA) algorithm; the cost functions and error functions of these algorithms are graphically illustrated in Figure 4 and Table 1. However, it should be noted that CMA will not completely fail for

16-QAM formats and above. The cost function will still converge to a nonzero minimum. In fact, for practical OSNR values, the ASE noise dominates the cost function $J(\mathrm{W})$ such that the benefit of CMMA/RDE is merely faster convergence at the expense of instability due to decision and classification errors of the received signal amplitudes. At present, a first stage of CMA for preconvergence followed by DD-LMS or LMS has become a common configuration in transmission experiments using 16-QAM and above [17]. For recent high baud rate transmission using 32- or 64-QAM systems, an additional LMS or DD-LMS is added after CPE to eliminate residual low bandwidth distortions caused by transmitter reflections [17]. This opens up new directions based on combining polarization demultiplexing, FOE, CPE, CMA, LMS, and permuting them differently to form various multistage configurations using hybrid blind/training-based adaptive signal processing. This is an active ongoing research area.



[FIG4] A graphical illustration and comparison between various adaptive signal processing techniques used in 16-QAM systems with equalizer output $\mathbf{v}$ and amplitude of decided symbols $\hat{R}$.

ENABLING FLEXIBLE TRANSPONDERS AND ELASTIC OPTICAL NETWORKS

It is envisioned that future Internet traffic will be more dynamic, unpredictable and heterogeneous as a result of the emergence of large content providers creating dynamic traffic demands across optical networks. Consequently research toward enabling flexible/adaptive transmission and/or EONs that maximize network efficiency has recently attracted a great deal of attention [18]. For example, research in impairment-aware network layer routing and spectrum assignment protocols as well as flexible reconfigurable optical add/drop multiplexers [19] enabling a switching granularity of 3.125 GHz allow one to move beyond the limitations of a standard 50-GHz grid. In terms of the physical layer, we suggested a universal software-defined bandwidth-variable transceiver (BVT) [20] that supports different carrier wavelengths, bit rates, symbol rates, modulation formats, coding, and transmission paths/distances across a network. Another imperative feature for flexible receivers is rapid and robust physical channel estimation for EONs.

TOWARD A UNIVERSAL DSP PLATFORM FOR ARBITRARY QAM TRANSMISSION

DSP for receiver front-end corrections, CD compensation, and TPR are essentially modulation format independent while CPE, FOE, and adaptive signal processing for compensating PMD and other transceiver imperfections are modulation format dependent. For example, CMMA or RDE are tailored to 16-QAM signals and above even though CMA is used for preconvergence regardless of modulation format. Also, current CPE techniques can range from the simple VVPE or DD-PLL for QPSK signals to complex multistage BPS+ML for 256-QAM systems. It would be ideal for EONs if a single universal DSP platform could adequately equalize various transmission impairments and recover data for arbitrary modulation formats.

For CPE, it can be argued that the BPS architecture is applicable to all modulation formats. However, similar to the majority of feedforward CPE techniques, BPS still requires a priori knowledge of the modulation format to compute the phase estimate. To

[TABLE 1] THE COST FUNCTIONS AND ERROR FUNCTIONS FOR VARIOUS ADAPTIVE SIGNAL-PROCESSING ALGORITHMS.

ALGORITHM	COST FUNCTION $J(\mathbf{w})$	ERROR FUNCTION	REMARKS
CMA	$E[(\lvert\mathbf{v}\rvert^2 - R_2^2)^2]$	$(R_2^2 - \lvert\mathbf{v}\rvert^2)\mathbf{v}$	SLOW CONVERGENCE BUT CAN BE USED AS PRECONVERGENCE FOR ALL FORMATS
CMMA	$E[(\lvert\lvert\mathbf{v}\rvert - R_1\rvert - R_2\rvert - R_3)^2]$	$\mathrm{sgn}(\mathbf{v})\mathrm{sgn}(\lvert\mathbf{v}\rvert - R_1)\mathrm{sgn}(\lvert\lvert\mathbf{v}\rvert - R_2\rvert - R_2)$	UNSTABLE WITH LOW OSNR
RDE	$E\left[\left(\lvert\mathbf{v}\rvert^2 - \lvert\hat{\boldsymbol{R}}\rvert^2\right)^2\right]$	$(\hat{R} - \lvert\mathbf{v}\rvert^2)\mathbf{v}$	UNSTABLE WITH LOW OSNR
DD-LMS	$E[(\mathbf{v} - \hat{\mathbf{s}})^2]$	$\mathbf{v} - \hat{\mathbf{s}}$	UNSTABLE WITH LARGE PHASE NOISE
DD-LRD	$E\left[\left(\lvert\mathbf{v}\rvert^2 - \lvert\hat{\mathbf{s}}\rvert^2\right)^2\right]$	$(\lvert\mathbf{v}\rvert^2 - \lvert\hat{\mathbf{s}}\rvert^2)\mathbf{v}$	FASTER CONVERGENCE THAN DD-LMS
MCMA	$E\left[\left(\mathrm{Re}\{\mathbf{v}\}^2 - R_2/\sqrt{2}\right)^2 + \left(\mathrm{Im}\{\mathbf{v}\}^2 - R_2/\sqrt{2}\right)^2\right]$	$\mathrm{Re}\{\mathbf{v}\}\left(\mathrm{Re}\{\mathbf{v}\}^2 - R_2/\sqrt{2}\right) + j\,\mathrm{Im}\{\mathbf{v}\}\left(\mathrm{Im}\{\mathbf{v}\}^2 - R_2/\sqrt{2}\right)$	LOW TOLERANCE TO FREQUENCY OFFSET

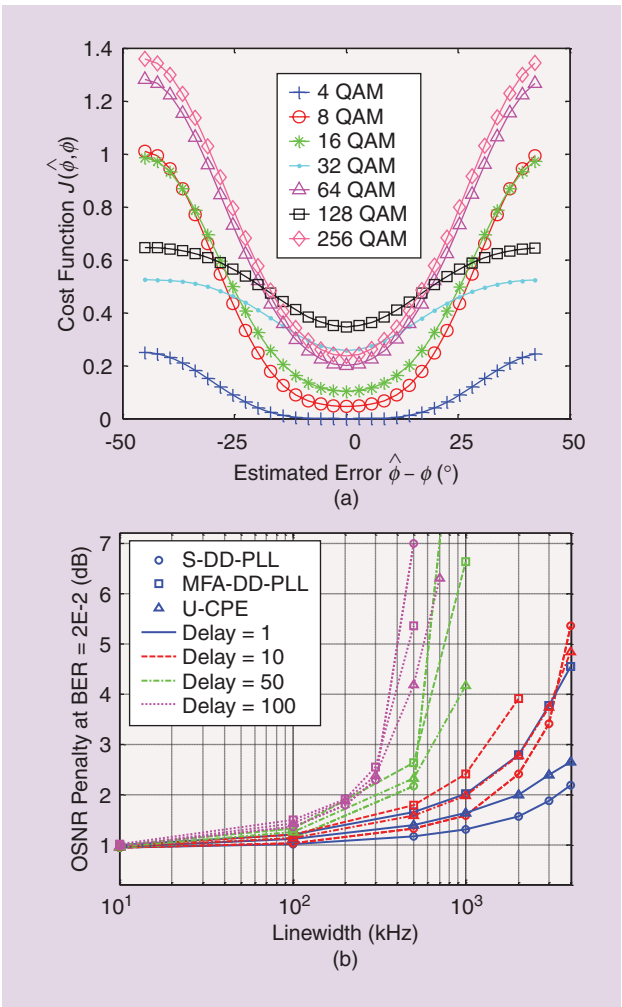
facilitate a blind and modulation-format-independent or universal-CPE (U-CPE), one seeks a cost function common to all modulation formats that is optimized when the phase estimate $\hat{\phi}$ approaches the true laser phase ϕ. For practical modulation formats, maybe the only common characteristic is that their constellation is somewhat square shaped. We make use of this insight [21] and study the real and imaginary part of the CPE input $r_{\mathrm{CPE}}(k)$ in a single polarization and consider the cost function

$$J(\hat{\phi},\phi) = \mathrm{E}\Big[\Big(\big(\mathrm{Re}\big\{r_{\mathrm{CPE}}(k)e^{-j\hat{\phi}}\big\}^2 - 0.5\big)^2 + \big(\mathrm{Im}\{r_{\mathrm{CPE}}(k)e^{-j\hat{\phi}}\}^2 - 0.5\big)^2\Big)^2\Big]. \quad (5)$$

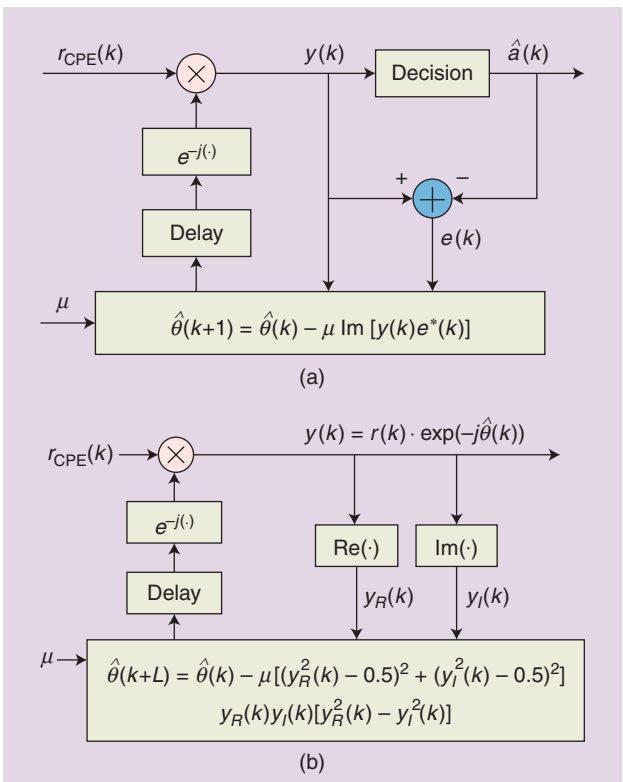
As shown in Figure 5(a), $(\hat{\phi},\phi)$ is minimized at $\hat{\phi}-\phi=0$ regardless of modulation format, thus illustrating that $J(\hat{\phi},\phi)$ can be a suitable cost function to realize U-CPE.

We compare U-CPE with 1) standard DD-PLL (S-DD-PLL) in which the modulation format is known for symbol decision, and 2) modulation format agnostic (MFA)-DD-PLL in which QPSK is assumed in the DD-PLL irrespective of the actual signal modulation format. The block diagram for S-DD-PLL and the update algorithm for U-CPE are shown in Figure 6. The OSNR penalties for different laser linewidths and feedback delays are shown in Figure 5(b) for 32 GBaud PM-16-QAM systems. U-CPE induces negligible penalties compared with S-DD-PLL and MFA-DD-PLL and, in fact, outperforms the other algorithms for large feedback delays. Nonetheless, we believe that, similar to CMA, U-CPE may be best suitable as a universal preconvergence algorithm followed by a modulation-format-specific or decision-directed CPE.

In addition to U-CPE, a universal-FOE (U-FOE) that is modulation format independent is also an integral part of a universal DSP platform for arbitrary QAM transmission. To this end, Nakagawa proposed to examine the received signal spectrum and use the asymmetric distribution of power with respect to zero frequency as a coarse estimate of frequency offset Δf [21]. To refine the estimate of Δf, one can raise the signal to the fourth power to remove the information for the QPSK signals and Δf can be estimated from the slope of the phase evolution over a block of symbols. In principle, this can also work for 16-QAM signals and above as long as the block length is long enough to average out the information remaining in the fourth-power of the



[FIG5] (a) Cost function $J(\hat{\phi},\phi)$ (without ASE noise) versus phase estimation error $\hat{\phi}-\phi$ for different modulation formats. (b) OSNR penalty versus laser linewidth for 32-GBaud PM-16-QAM systems using standard S-DD-PLL, U-CPE, and MFA-DD-PLL at an SD-FEC threshold of 2E-2, corresponding to an OSNR of 16.8 dB.



[FIG6] A block diagram of (a) S-DD-PLL (and MFA-DD-PLL) and (b) U-CPE. The "delay" block represents the total number of symbol delays L due to number of parallelization paths multiplied by pipeline delay of the algorithms.

higher-order signals. Consequently, the spectral asymmetry method can be combined with the fourth-power technique to form a blind and universal FOE unit. The implementation details and performance characterizations of such a scheme and other U-FOE proposals are yet to be better understood.

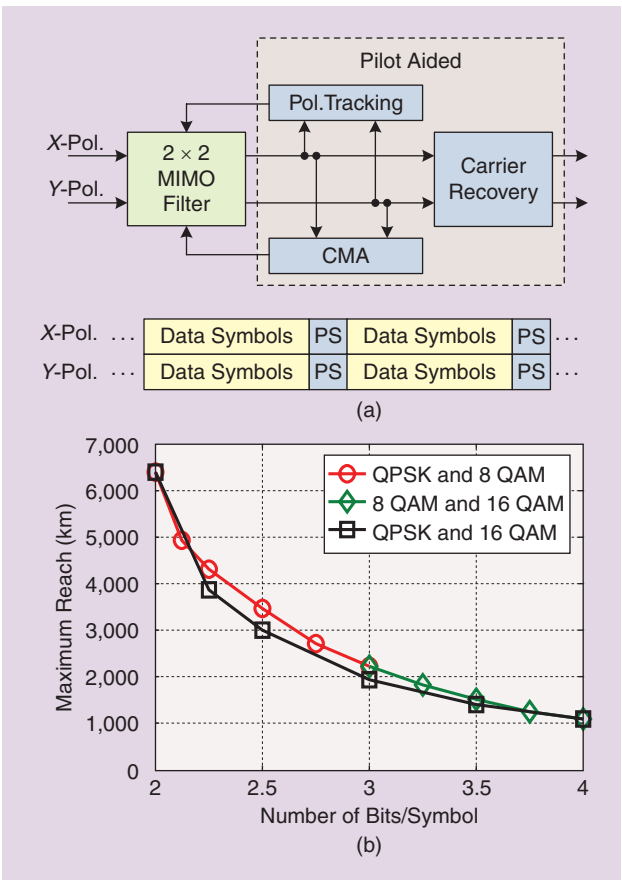
The U-CPE and U-FOE described above are blind algorithms. If pilot symbols (PSs) are allowed, universal DSP designs can be simpler and more efficient. In [23], we proposed a pilot-aided universal DSP, which consists of two stages at the receiver: 1) receiver initialization and 2) channel tracking. At the initialization stage, QPSK training symbols are sent for synchronization and estimation/compensation of channel parameters such as frequency offset. The tracking stage includes an adaptive 2×2 MIMO filter, carrier recovery unit, and periodic transmission of PS as shown in Figure 7(a). Each data frame contains 400 data symbols and four PSs with QPSK format, resulting in an overhead of 1%. A special pair of PSs $e^{j\pi/4}[11;1-1]$ is used to calculate the polarization rotation angle (described in a later section) for each two frames, and CMA is used on the rest of the six QPSK PSs for other channel drifting effects such as PMD. The same PSs are used in the carrier recovery including SSP-PLL+ML and frequency offset tracking.

We conducted Tb/s super channel experiments using different modulation formats including QPSK, 8-QAM and 16-QAM where the BERs of all subchannels are below the thresholds (7% hard-decision FEC and 20% soft-decision FEC). Various distances ranging from 240 km to 6,000 km are achieved for different formats and different FEC coding schemes, demonstrating the flexibility and universality of our proposed DSP platform.

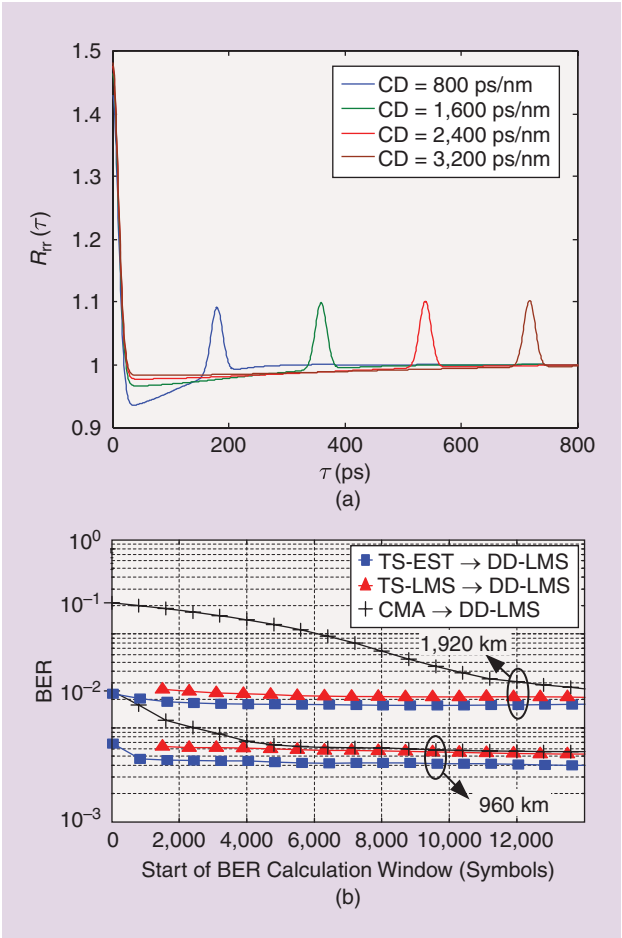
RATE-ADAPTIVE TRANSMISSION USING TIME-DOMAIN HYBRID QAM AND RATE-ADAPTIVE CODING

Conventional transponders employ individual modulation formats such as BPSK, QPSK, or 16-QAM, which have discrete spectral efficiencies and achievable transmission distances, resulting in limited flexibility that hinders overall network efficiencies. One approach to realize full flexibility in bit rates and/or spectral efficiencies is time-domain hybrid QAM (TDHQ) [23]. This is accomplished by interleaving two different modulation formats in the time domain and their relative ratio of occurrence determines the overall spectral efficiency and hence maximal reach. Recently, we investigated various combinations of two QAM formats including QPSK, 8QAM, and 16QAM, denoted as QPSK&8QAM, 8QAM&16QAM, and QPSK&16QAM. The average number of bits per symbol is determined by the ratio of the two formats. Note that the power ratio of the two hybrid formats should be adjusted to achieve the same bit-error ratio (BER), which is crucial for the FEC decoding. We conducted 28 Gbaud nonreturn-to-zero (NRZ) experiments deployed on a fixed 50-GHz grid. At the receiver, IQ-imbalance compensation, CD compensation, and frequency offset compensation are first applied. The data frame is synchronized followed by LMS-based adaptive MIMO equalization and DD-PLL as CPE. The frame synchronization is critical since the decision circuit needs to be switched to the current QAM format for the subsequent LMS and DD-PLL. Before the DD operation, 4,096 training symbols are used for the initialization of the MIMO filter coefficients.

The measured maximum reach at a BER of 4.6×10^{-3} [7% hard-decision forward error correction (FEC) threshold] versus the number of bits/symbol (or spectral efficiency) are shown in Figure 7(b). We realized continuous data rates from 112 Gb/s to 224 Gb/s and corresponding distances from 6,400 km to 1,100 km by varying the hybrid signals. The experiment demonstrates that TDHQ can be used to maximize spectral efficiencies for different link conditions in an EON. Moreover, no extra hardware is required with respect to a conventional DAC-based-flexible transmitter. In addition, four-dimensional (4-D) modulation, which fully exploit the 4-D space of the optical carrier, have attracted attention recently for their higher power efficiency compared to conventional QAM formats. Four-dimensional modulation can also be used in flexible transceivers as they achieve a higher resolution in spectral efficiency. Sjödin et al. [24] demonstrated that 128 set-partitioning (SP)-QAM can increase the transmission distance by around 50% in the expense of 12.5% spectral efficiency reduction compared to PM-16-QAM.



[FIG7] (a) A block diagram of the channel tracking stage and illustration of the transmitted frames. (b) Measured maximum reach versus the number of bits per symbol for 28 Gbaud time-domain hybrid QAM transmission.



[FIG8] (a) The autocorrelation function of received power waveform $R_{rr}(\tau)$ for 112-Gb/s PM-QPSK systems with 800, 1,600, 2,400, and 3,200 ps/nm of CD. The peak location can be used to determine link CD and is analytically shown to be insensitive to laser impairments, ASE noise, and first-order PMD effects. (b) Transient BER using TS-EST, TS-LMS, and CMA as preconvergence algorithms for a 224-Gb/s PM-16-QAM system.

In addition to varying modulation formats, rate-adaptive transmission can also be realized by rate-adaptive coding in which the FEC code rate is made adaptive according to link conditions such as link OSNR or BER before the decoder [25]. In particular, for a fixed modulation format and fixed symbol rate, one can use a nominal code and puncture it (remove parity bits of a codeword) to increase code rate or shorten it (deleting information bits of a codeword) to decrease code rate. In this case, the resulting code can span a range of code rates that is suitable for different link conditions, thus realizing flexibility in the overall transmission bit rate. This idea can be further extended to rate-adaptive modulation and coding in which both the code rate and modulation format can be made adaptive to link conditions.

FAST AND ROBUST PARAMETER ESTIMATION FOR FLEXIBLE TRANSMISSION

Enabling dynamic light path provisioning and automatically switched optical networks (ASONs) will require a digital coherent optical receiver able to quickly estimate the physical layer conditions of the channel. In principle, parameters such as link CD and PMD can be estimated using blind adaptive equalization techniques. However, for 28–32 Gbaud signals, an adaptive equalizer can handle around ± 200 ps/nm residual CD beyond which it will take too long to converge or may simply not converge at all. Therefore, separate efforts for CD estimation prior to data transmission is desirable. CD can be estimated by inserting training symbols or specific bit sequences [26] but the length of pilot symbols needs to grow with CD and may become prohibitive for links approaching 100,000 ps/nm of CD. On the other hand, blind nondata-aided (NDA) CD estimation is often required/preferred [27] for equalizer initialization, but they need to be insensitive to other impairments, independent of modulation format, and need to be fast.

We proposed using the autocorrelation of received signal power waveform $R_{rr}(\tau) = \mathrm{E}[(|E_{r,x}(t)|^2) + (|E_{r,y}(t)|^2) \cdot (|E_{r,x}(t+\tau)|^2 + |E_{r,y}(t+\tau|^2)]$ for fast and robust CD estimation [28]. Due to CD-induced intersymbol interference, Figure 8(a) shows that $R_{rr}(\tau)$ exhibits a peak whose location is analytically shown to depend on accumulated CD and is independent of amplifier noise, laser frequency offset, laser phase noise, first-order PMD, and modulation format. Thus, the peak location can be used as a fast and robust CD estimation technique. The peak in $R_{rr}(\tau)$ become less apparent for NRZ and Nyquist-shaped signals but additional high-pass filtering to the received signal can "preserve" the peak and ensure robust CD estimation [28]. Recently, Malouin et al. demonstrated fast CD estimation by examining the clock tone power without linear scanning a preset CD range with a fairly small number of symbols [29]. However, such a clock tone will be considerably suppressed by PMD effects and additional efforts are needed to improve the robustness of estimation. It should be noted that most of the proposed CD estimation techniques can be seen as different variants of calculating the CD induced-group delay between the upper-sideband and lower-sideband of the received signal and estimation accuracies will suffer for Nyquist shaped signals with sharp spectral roll-off.

Provided that CD is appropriately estimated and compensated, the next task of the receiver DSP is to estimate the 2×2 Jones channel matrix and perform polarization demultiplexing. We recently proposed a novel format transparent training symbols (TS)-based channel estimation (TS-EST) algorithm [30]. The TSs are used to estimate the angle between SOP and polarization axis (assuming PMD is absent), initialize the center taps of a 2×2 MIMO time-domain equalizer (TDE) and then steady-state DD-LMS can be started after as few as 40 TSs. Figure 8(b) compares the transient BER using TS-EST, TS-LMS, and CMA for a 224 Gb/s PM-16-QAM system with different transmission distances. It is obvious that the TS-EST algorithm achieves the same steady-state BER with a much faster convergence speed. Besides initialization, TS-EST is also suitable for steady-state tracking of fast polarization transients if TS are sent periodically. Note that while the TS-EST model neglects PMD effects, the TS-EST algorithm can tolerate DGD up to half the symbol

duration. This is because the initialized center taps from TS-EST are already reasonably close to the exact inverse Jones channel matrix such that the steady-state DD-LMS can adequately update all taps of the MIMO TDE that inherently compensates PMD impairments.

It should be noted that while channel estimations are imperative for digital optical coherent receivers in compensating various transmission impairments, the task of channel estimation itself became mathematically equivalent and in fact identical to optical performance monitoring (OPM) when implemented at digital optical coherent receivers. In this case, many OPM functions such as CD/PMD/power monitoring are "free by-products" since they can be readily extracted from filter taps of various DSP algorithms. Extending this argument, we believe that digital optical coherent receivers enable OPM to be aggressively used to supply, in real time, channel state information to levels above the physical layer (e.g., updating routing information and routing tables).

DIGITAL COHERENT RECEIVERS ENABLE 100-Gb/s TRANSMISSIONS, AND BASIC DSP BLOCKS HAVE BECOME FUNDAMENTAL AND REPRESENTATIVE FEATURES OF CURRENT OPTICAL TRANSPONDERS.

CONCLUSIONS AND OUTLOOK

Digital optical coherent receivers enable 100 Gb/s transmissions, and basic DSP blocks have become fundamental and representative features of current optical transponders. We are witnessing another era of development in optical transmission technologies in which advanced DSP techniques help enable higher spectral efficiency transmission approaching 1 Tb/s per channel as well as flexible transponders to maximize network efficiencies. Research on adaptive transponders supporting and maximizing the throughput under a wide set of network conditions continues to be a major focus area in digital coherent communications. Future optical networks might even evolve into a cognitive structure in which each transponder in the network is able to sense the availability of spectrum and formulate appropriate communication strategies in real time and collectively maximize overall network efficiencies.

AUTHORS

Alan Pak Tao Lau (eeaptlau@polyu.edu.hk) received the B.A.Sc. degree in engineering science and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Ontario, Canada, in 2003 and 2004, respectively. He received the Ph.D. degree in electrical engineering from Stanford University in 2008 and joined the Hong Kong Polytechnic University where he is currently an associate professor. His current research interests include various aspects of coherent fiber-optic communication systems and OPM. He is currently the principal investigator and/or coinvestigator of various government- and industry-funded research projects in optical communications. He is a reviewer and technical program committee member for various IEEE/Optical Society of America journals and international conferences in photonics and communications.

Yuliang Gao (ryancoolg@gmail.com) received the B.E. and M.E. degrees from Beijing University of Posts and Telecommunications, China, in 2008 and 2011, respectively. He joined the Photonics Research Centre and Department of Electrical Engineering at the Hong Kong Polytechnic University in 2010 where he is now working toward the Ph.D. degree. His research interest is mainly focused on the DSP techniques in coherent optical transmission systems.

Qi Sui (suiqi.sjtu@gmail.com) received the B.Eng. degree in electronic engineering from Shanghai Jiaotong University in 2007. He joined the Photonics Research Centre of the Hong Kong Polytechnic University in 2008. His current research interest include coherent optical communications and OPM.

Dawei Wang (wdw2008@gmail.com) received the B.S. degree in information engineering and the Ph.D. degree in optical engineering from Zhejiang University, Hangzhou, China, in 2008 and 2013, respectively. He joined the Photonic Research Center at The Hong Kong Polytechnic University in 2010 and 2012, respectively, working as a research assistant. He is now a research engineer at the Communication Technology Laboratory, Huawei Technologies Co., Shenzhen, China.

Qunbi Zhuge (qunbi.zhuge@mail.mcgill.ca) received the B.S. degree in optical engineering from Zhejiang University, Hangzhou, China, in 2009, and the M.E. degree in electrical and computer engineering in 2012 from McGill University, Montreal, Canada, where he is currently working toward the Ph.D. degree in the area of coherent optical transmission system. From April 2013 to September 2013, he was an intern at Ciena, Ottawa, Canada, working on the design of next-generation coherent optical transceiver. He has published over 45 papers and is a recipient of the 2013 IEEE Photonics Society Graduate Student Fellowships and 2013 SPIE Optics and Photonics Education Scholarship.

Mohamed H. Morsy-Osman (mohamed.osman2@mail.mcgill.ca) received the B.S. and M.S. degrees from Alexandria University, Egypt, in 2006 and 2009, respectively. In 2010, he joined the Department of Electrical and Computer Engineering, McGill University, Montreal, Canada, as a research assistant, where he is currently working toward the Ph.D. degree. His current research interests include DSP for high-speed optical transport networks, optical burst switched networks, optical code-division multiple-access networks, media access control protocols in optical networks, and wireless communications.

Mathieu Chagnon (mathieu.chagnon@mail.mcgill.ca) received an engineering physics degree with the Principal and Vice Chancellor's mention of excellence from Polytechnique Montreal in 2008 and an M.Eng. degree (dean's honors) from McGill University in 2010, where he is currently working toward the Ph.D. degree. His research interests are focused on DSP for optical coherent transceivers for high bit rate, long-haul optical transmission and novel 4-D modulation formats for passive nonlinear mitigation. He is also working on performance analysis of integrated photonics coherent receivers.

Xian Xu (xian.xu@mail.mcgill.ca) received an M.E. degree in electrical and computer engineering in 2010 from McGill University, Montreal, Canada, where he is currently working toward the Ph.D. degree. His research interests focus on DSP techniques for coherent optical transmissions, particularly, pulse-shaping techniques for optic fiber channel nonlinear mitigation and blind equalization techniques in the coherent receivers. He authored and coauthored over 30 journal and conference papers. He is a member of the IEEE Photonics Society and the recipient of the Fonds de Recherche du Québec—Nature et Technologies doctoral research scholarship.

Chao Lu (chao.lu@polyu.edu.hk) received the B.Eng. degree in electronic engineering from Tsinghua University, China, in 1985 and the M.Sc. and Ph.D. degrees from the University of Manchester, United Kingdom, in 1987 and 1990, respectively. From 1991 to 2006, he was with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, as lecturer, senior lecturer, and associate professor. From June 2002 to December 2005, he was seconded to the Institute for Infocomm Research, Agency for Science, Technology, and Research (A*STAR), Singapore, as program director and department manager. Since April 2006, he has been a professor in the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University. His research interests include optical communication systems and networks, fiber devices for optical communication, and sensor systems.

David V. Plant (david.plant@mcgill.ca) received the Ph.D. degree from Brown University (1989), was a research engineer at the University of California Los Angeles (1989–1993), and has been a professor at McGill University since 1993, where he holds a James McGill Professorship. He has published over 230 journal and conference papers and has one licensed patent. He has received five teaching awards and numerous other awards including the IEEE Photonics Society Distinguished Lectureship and the IEEE Microwave Theory and Techniques Society Microwave Prize. He is a Fellow of the IEEE, OSA, Engineering Institute of Canada, and the Canadian Academy of Engineering and currently holds a Killam Research Fellowship.

REFERENCES

[1] M. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb. 2004.

[2] S. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE Select. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1164–1179, Oct. 2010.

[3] A. Leven, N. Kaneda, and S. Corteselli, "Real-time implementation of digital signal processing for coherent optical digital communication systems," *IEEE Select. Topics Quantum Electron.*, vol. 16, no. 5, pp. 1227–1234, Sept./Oct. 2010.

[4] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with applications to burst digital transmission," *IEEE Trans. Inform. Theory*, vol. 29, no. 4, pp. 543–551, July 1983.

[5] T. Pfau, S. Hoffman, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for M-QAM constellations," *J. Lightwave Technol.*, vol. 27, no. 8, pp. 989–999, Apr. 2009.

[6] J. Li, L. Li, Z. Tao, T. Hoshada, and J. C. Rasmussen, "Laser-linewidth-tolerant feedforward carrier phase estimator with reduced complexity for QAM," *J. Lightwave Technol.*, vol. 29, no. 16, pp. 2358–2364, Aug. 2011.

[7] I. Fatadin, D. Ives, and S. J. Savory, "Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning," *IEEE Photon. Technol. Lett.*, vol. 22, no. 9, pp. 631–633, May 2010.

[8] Y. Gao, A. P. T. Lau, S. Yan, and C. Lu, "Low-complexity and phase noise tolerant carrier phase estimation for dual-polarization 16-QAM systems," *Opt. Express*, vol. 19, no. 22, pp. 21717–21729, Oct. 2011.

[9] N. Sigron, I. Tselniker, and M. Nazarathy, "Carrier phase estimation for optically coherent QPSK based on Wiener-optimal and adaptive multi-symbol delay detection (MSDD)," *Opt. Express,* vol. 20, no. 3, pp. 1981–2003, Jan. 2012.

[10] S. Zhang, P.-Y. Kam, C. Yu, and J. Chen, "Decision-aided carrier phase estimation for coherent optical communications," *J. Lightwave Technol.*, vol. 28, no. 11, pp. 1597–1607, June 2010.

[11] A. Tolmachev, I. Tselniker, M. Meltsin, N. Sigron, and M. Nazarathy, "Efficient multiplier-free FPGA demonstration of polar-domain multi-symbol-delay-detector (MSDD) for high performance phase recovery of 16-QAM," in *Proc. Optical Fiber Communications/National Fiber-Optic Engineers Conf. (OFC/NFOEC 2013)*, Paper OM2C.8.

[12] K. Piyawanno, M. Kuschnerov, B. Spinnler, and B. Lankl, "Low complexity carrier recovery for coherent QAM using superscalar parallelization," in *Proc. 2010 European Conf. Optical Communications*, Paper. We.7.A.3.

[13] Q. Zhuge, M. Morsy-Osman, X. Xu, M. E. Mousa-Pasandi, M. Chagnon, Z. A. El-Sahn, and D. V. Plant, "Pilot-aided carrier phase recovery for M-QAM using superscalar parallelization based PLL," *Opt. Express*, vol. 20, no. 17, pp. 19599–19609, Aug. 2012.

[14] Y. Gao, A.P.T. Lau, C. Lu, Y. Dai, and X. Xu, "Blind cycle-slip detection and correction for coherent communication systems," in *Proc. 2013 European Conf. Optical Communications*, Paper P3.16.

[15] X. Zhou, J. Yu and P. Magill, "Cascaded two-modulus algorithm for blind polarization demultiplexing of 114-Gb/s PDM-8-QAM optical signals," in *Proc. OFC/NFOEC 2009,* Paper OWG3.

[16] X. Xu, B. Chatelain, and D. V. Plant, "Decision directed least radius distance algorithm for blind equalization in a dual-polarization 16-QAM system," in *Proc. Optical Fiber Communications/National Fiber-Optic Engineers Conf. (OFC/NFOEC) 2012*, Paper OM2H.5.

[17] X. Zhou, L. E. Nelson, R. Isaac, P. Magill, B. Zhu, D. W. Peckham, P. Borel, and K. Carlson, "1200-km transmission of 50GHz spaced, 5x504-Gb/s PDM-32-64 hybrid QAM using electrical and optical spectral shaping," in *Proc. Optical Fiber Communications/National Fiber-Optic Engineers Conf. (OFC/NFOEC) 2012,* Paper OM2A.2.

[18] O. Gerstel, M. Jinno, A. Lord, and S. J. Yoo, "Elastic optical networking: A new dawn for the optical layer? *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s12–s20, Feb. 2012.

[19] T. A. Strasser and J. L. Wagener, *IEEE Select. Topics Quantum Electron.*, vol. 16, pp. 1150–1157, Sept.–Oct. 2010.

[20] M. Eiselt, B. Teipen, K. Grobe, A. Autenrieth, and J.-P. Elbers, "Programmable modulation for high-capacity networks" in *European Conf. Optical Communications (ECOC 2011)*, Paper Tu.5.A.5.

[21] Y. Gao, A. P. T. Lau, and C. Lu, "Modulation-format-independent carrier phase estimation for square M-QAM systems," *IEEE Photon. Technol. Lett.,* vol. 25, no. 11, pp. 1073–1076, June 2013.

[22] T. Nakagawa, M. Matsui, T. Kobayashi, K. Ishihara, R. Kudo, M. Mizoguchi, and Y. Miyamoto, "Non-data-aided wide-range frequency offset estimator for QAM optical coherent receivers," in *Proc. Optical Fiber Communications/National Fiber-Optic Engineers Conf. (OFC/NFOEC 2011)*, Paper OMJ.1.

[23] Q. Zhuge, M. Morsy-Osman, X. Xu, M. Chagnon, M. Qiu, and D. V. Plant, "Spectral efficiency-adaptive optical transmission using time domain hybrid QAM for agile optical networks," *J. Lightwave Technol.*, vol. 31, no. 15, pp. 2621–2628, Aug. 2013.

[24] M. Sjödin, P. Johannisson, J. Li, E. Agrell, P. A. Andrekson, and M. Karlsson, "Comparison of 128-SP-QAM with PM-16-QAM," *Opt. Express*, vol. 20, no. 8, pp. 8356–8366, Apr. 2012.

[25] G.-H. Gho, L. Klak, and J. M. Kahn, "Rate-adaptive coding for optical fiber transmission systems," *J. Lightwave Technol.*, vol. 29, no. 2, pp. 222–233, Jan. 2011.

[26] C. C. Do, A. V. Tran,C. Zhu, S. Chen, T. Anderson, D. Hewitt, and E. Skafidas, "Data-aided chromatic dispersion estimation for polarization multiplexed optical systems," *IEEE Photon. J.*, vol. 4, no. 5, pp. 2037–2049, Oct. 2012.

[27] R. A. Soriano, F.N. Hauske,N.G.Gonzalez, and I. T. Monroy, "Chromatic dispersion estimation in digital coherent receivers," *J. Lightwave Technol.*, vol. 29, no. 11, pp. 1627–1637, June 2011.

[28] Q. Sui, A. P. T. Lau, and C. Lu, "Fast and robust blind chromatic dispersion estimation using auto-correlation of signal power waveform for digital coherent systems," *J. Lightwave Technol.*, vol. 31, no. 2, pp. 306–312, Jan. 2013.

[29] C. Malouin, P. Thomas, B. Zhang, J. O'Neil, and T. Schmidt, "Natural expression of the best-match search Godard clock-tone algorithm for blind chromatic dispersion estimation in digital coherent receivers," in *Proc. OSA SPPCom*, June 2012, Paper SpTh2B.4.

[30] M. Morsy-Osman, M. Chagnon, Q. Zhuge, X. Xu, M. E. Mousa-Pasandi, Z. A. El-Sahn, and D. V. Plant, "Ultrafast and low overhead training symbol based channel estimation in coherent M-QAM single-carrier transmission systems," *Opt. Express*, vol. 20, no. 26, p. B171–B180, 2012.

[SP]

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

Lotfollah Beygi, Erik Agrell, Joseph M. Kahn, and Magnus Karlsson

Coded Modulation for Fiber-Optic Networks

Toward better tradeoff between signal processing complexity and optical transparent reach

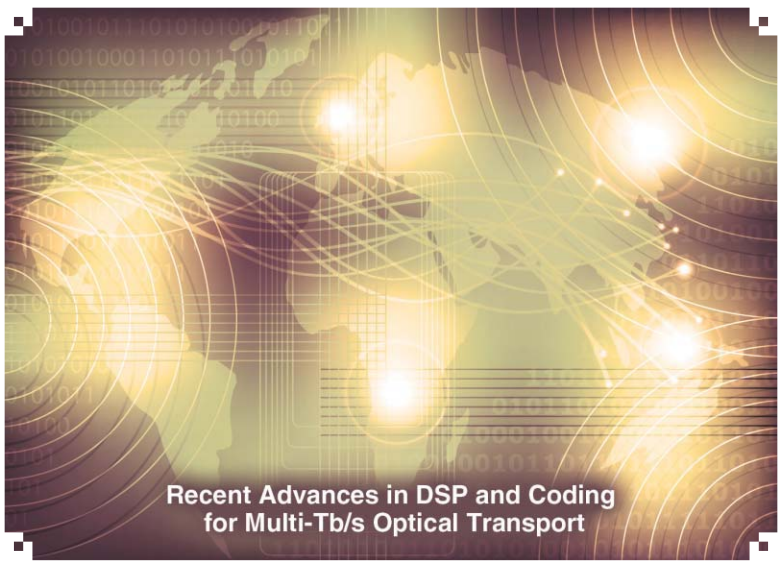


IMAGE LICENSED BY INGRAM PUBLISHING /JANMIKS

In this tutorial, we study the joint design of forward error correction (FEC) and modulation for fiber-optic communications. To this end, we use an information-theoretic design framework to investigate coded modulation (CM) techniques for standard additive white Gaussian noise (AWGN) channels and fiber-optic channels. This design guideline helps us provide a comprehensive overview of the CM schemes in the literature. Then, by invoking recent advances in optical channel modeling for nondispersion-managed links, we discuss two-dimensional (2-D) and four-dimensional (4-D) CM schemes. Moreover, we discuss the electronic computational complexity and hardware constraints of CM schemes for optical communications. Finally, we address CM schemes with signal shaping and rate-adaptation capabilities to accommodate the data transmission scheme to optical links with different signal qualities.

Digital Object Identifier 10.1109/MSP.2013.2290805
Date of publication: 12 February 2014

INTRODUCTION

The tremendous growth in the demand for high data rates in optical networks encourages exploiting the available resources in this medium more efficiently. Much effort has been devoted to quantifying fundamental limits of fiber-optic channels [1]–[3]. Indeed, the more severe signal-dependent nonlinear effect in fiber-optic channels, compared to wireline and wireless channels, makes the channel modeling and capacity analysis of these channels cumbersome. The recent progress in channel modeling [4]–[6] and capacity analysis [3] of fiber-optic channels have opened a new horizon in the design of data transmission schemes operating with higher spectral efficiencies than current systems. The transparent reach, i.e., the transmission distance of a fiber-optic link with no inline

electrical signal regenerators, is intimately related to the desired spectral efficiency, i.e., the number of information bits sent in each polarization per symbol period, as well as to the digital signal processing (DSP) complexity [depicted in Figure 1(a)]. For example, the larger the transparent reach is, the higher the DSP complexity gets, provided that the desired spectral efficiency is achievable for this transparent reach.

Joint coding and (multilevel) modulation schemes, so-called CM, have been investigated as means to provide higher coding gain to increase reach while maintaining acceptable complexity. The CM techniques [7] are known to be superior to conventional approaches using independent FEC and modulation in the sense of requiring less signal-to-noise ratio (SNR) for the same spectral efficiency. In fact, a CM scheme can exploit the four available dimensions of a fiber-optic link, i.e., two polarizations each consisting of in-phase and quadrature dimensions, with more flexibility than conventional schemes. In addition, the channel state information (CSI) can be taken into account in the design of a CM scheme, leading to a channel-aware CM scheme capable of adapting to different signal qualities in optically switched mesh networks with a dynamic or heterogeneous structure.

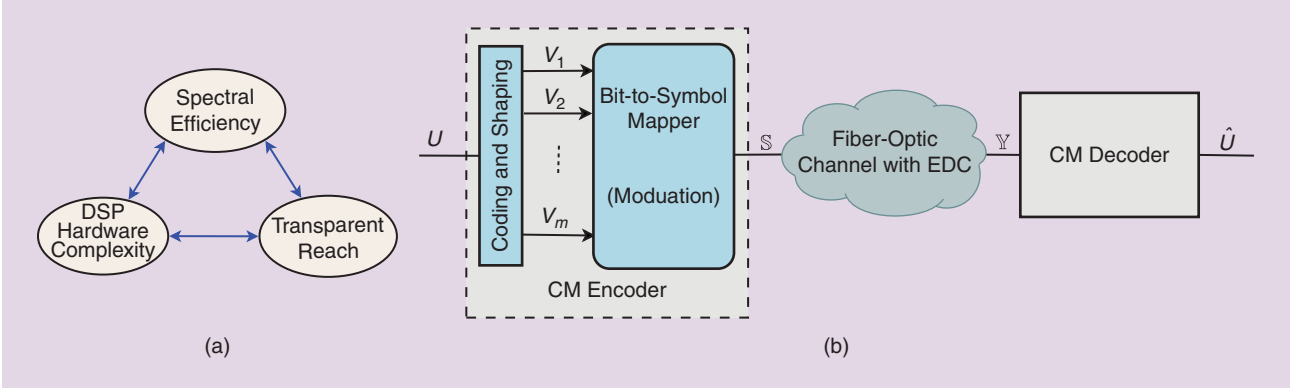
JOINT CODING AND (MULTILEVEL) MODULATION SCHEMES, SO-CALLED CM, ARE KNOWN TO BE SUPERIOR TO CONVENTIONAL APPROACHES USING INDEPENDENT FORWARD ERROR CORRECTION AND MODULATION, IN THE SENSE OF REQUIRING LESS SIGNAL-TO-NOISE RATIO FOR THE SAME SPECTRAL EFFICIENCY.

FIBER-OPTIC LINKS

Light is an electromagnetic wave, which can be modulated to convey information bits in fiber-optic links including N spans, each consisting of a single-mode fiber (SMF) and an erbium-doped fiber amplifier (EDFA). The electric field of the propagating signal experiences four types of impairments in these links: 1) signal attenuation, 2) AWGN noise added in each EDFA after amplifying the signal to compensate for the fiber loss, 3) frequency-dependent phase shift known as chromatic dispersion, and 4) intensity-dependent phase shift in the time domain, the so-called nonlinear Kerr effect. If the fiber is broken into sufficiently short segments, the chromatic dispersion and the nonlinear Kerr effect can be thought of as acting sequentially and independently. The propagation of light in these channels is described by the nonlinear Schrödinger equation. Due to the lack of analytical solutions and the complexity of numerical approaches, deriving the discrete-time statistics of such channels is, in general, cumbersome.

A fiber-optic link can compensate for the chromatic dispersion optically using an inline dispersion compensation fiber, leading to a dispersion-managed (DM) link, or electronically by an electronic dispersion compensation (EDC) unit in the receiver, resulting in a so-called non-DM link. Generally speaking, the high accumulated chromatic dispersion in a non-DM link turns the distribution of the electric field into Gaussian and consequently mitigates the nonlinear Kerr effect. Therefore, non-DM links outperform the widely used DM links for sufficiently large symbol rates and Gaussian or Nyquist pulses. The better performance of non-DM links has attracted a global interest in exploiting SMF links with EDC for next-generation optical networks.

A non-DM link including a CM encoder and decoder with EDC is depicted in Figure 1(b). As seen, the CM scheme first encodes the sequence of information bits U to m bit sequences $V_1, V_2, \ldots, V_m$. These m sequences are mapped to a sequence of symbols $\mathbb{S}$ from a 4-D constellation (at each time instant, a vector consisting of one bit from each m bit sequences is mapped to a 4-D symbol). A 4-D constellation can be constructed by a Cartesian product of two equal quadrature amplitude modulations (QAMs), which are used for independent data transmission over each polarization. The symbol sequence $\mathbb{S}$ is transmitted through a fiber-optic channel and received as the symbol sequence $\mathbb{Y}$ after the EDC.



[FIG1] **(a) The three main factors in the design of a CM scheme for fiber-optic links. (b) A fiber-optic link including a CM encoder and decoder with EDC (U and $\hat{U}$ are the transmitted and decoded information bit sequences, respectively).**

CHANNEL MODEL

Recently, a series of analytical models have been proposed for non-DM fiber-optic links [5], [6] with standard M-ary QAM (M-QAM) considering additive, Gaussian noise. The Gaussian noise model represents the received signal $\mathbb{Y}$ in a polarization-multiplexed (PM) fiber-optic channel with EDC as $\mathbb{Y} = \zeta\mathbb{S} + \mathbb{Z}$, where $\mathbb{S}$ is the transmitted PM signal, $\mathbb{Z}$ is a noise vector with a complex zero-mean circularly symmetric AWGN in each polarization, and ζ is a complex constant attenuation factor, which attenuates and rotates the transmitted symbol in each polarization. The variance of the zero-mean AWGN in each polarization is given by $\sigma^2 = N\sigma^2_{\text{ASE}} + \sigma^2_{\text{NL}}$, where $\sigma^2_{\text{NL}} = a_{\text{NL}}P^3$ is the variance of the noiselike interference, the so-called nonlinear noise, caused by the nonlinear Kerr effect, in which a_{NL} is a function of channel parameters and P is the average transmitted power. The term $N\sigma^2_{\text{ASE}}$ denotes the variance of the total amplified spontaneous emission (ASE) noise from the EDFAs over N amplifier spans. Finally, the SNR is defined as $|\zeta|^2P/\sigma^2$ for the non-DM system. Since the variance of the (nonlinear distortion) noise grows as the cube of the transmitted power, as shown in Figure 2(a), the system performance is eventually degraded at high transmitted power levels. This nonlinear behavior distinguishes these channels from classical AWGN channels. Clearly, there is an optimum power [shown by two stars in Figure 2(a)], which yields the minimum uncoded symbol error ratio (SER) or the maximum SNR after the EDC.

This optimum signal power is almost independent of the transparent reach, and the systems introduced in this article are assumed to operate at the optimal transmit power. A well-designed CM scheme allows for reliable data transmission with a higher uncoded SER, which leads to increasing the transparent reach. In this article, we consider only a single-channel system to keep the numerical simulation run time reasonable. However, the Gaussian noise model applies also to wavelength-division-multiplexing (WDM) systems, as long as one accounts for the entire optical signal spectrum as outlined in, e.g., [5]. According to this model for non-DM fiber-optic links, numerically and experimentally validated, including effects of interchannel nonlinearities in the WDM case only increases the variance of the AWGN. This leads to a reduction in the maximum transparent reach at which a given bit rate can be achieved, but the results will not change qualitatively.

QUALITY PARAMETERS

We will use three quality parameters to evaluate the performance of optical data transmission systems with hard- and soft-decision decoding, including FEC threshold, NCG, and gap to the AWGN channel capacity.

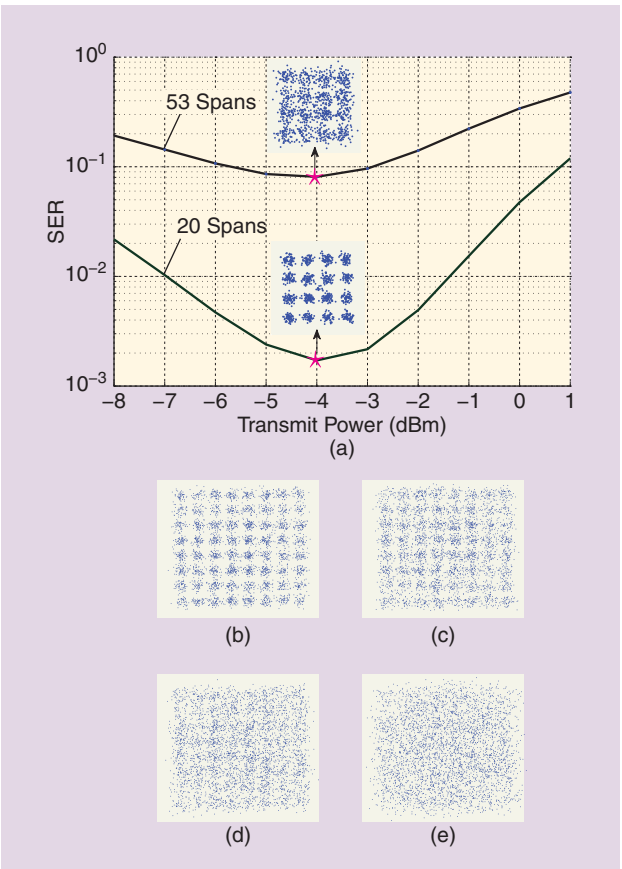
FEC THRESHOLD

Traditionally, due to the use of independent FEC and modulation together with hard-decision demodulation, the maximum bit-error ratio (BER) of a hard-decision demodulator (the input BER of the FEC decoder), the so-called FEC threshold, for obtaining the information BER of 10^{-15} at the output of the FEC decoder has been widely used as a metric for these channels. Often, the main goal of system designers was to meet the desired FEC threshold for an uncoded system.

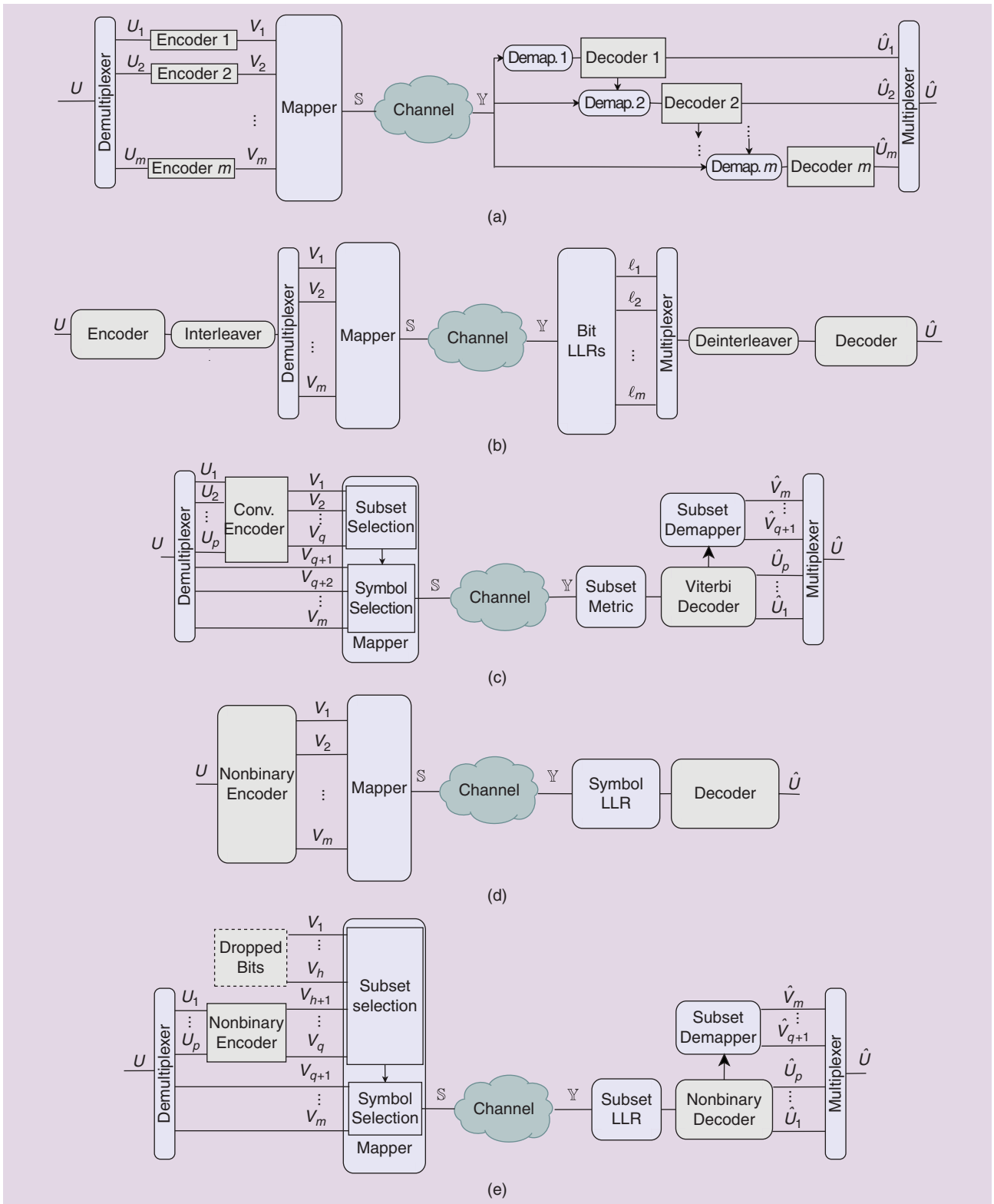
THE REDUCTION IN THE SNR REQUIREMENT RESULTING FROM ADDING CODING AT THE SAME INFORMATION BIT RATE AND THE SAME (LOW) INFORMATION BER FOR BOTH CODED AND UNCODED SYSTEMS IS CALLED THE NET CODING GAIN.

NET CODING GAIN

The reduction in the SNR requirement resulting from adding coding at the same information bit rate and the same (low) information BER for both coded and uncoded systems is called the net coding gain (NCG). The code rate of the coded system is $R = \eta/\eta_{\text{uncod}}$, where η_{uncod} and η are the spectral efficiencies of the uncoded and coded systems, respectively. The system coding overhead is defined as $\text{OH} = 1/R - 1$. The NCG is precisely



[FIG2] (a) The SERs of a nonlinear fiber-optic link with 20 and 53 spans together with the scatter plots of the received signals for a 16-QAM at the minimum SER, marked by two stars. The scatter plot of the received signal for a nonlinear fiber-optic link with 64-QAM operating (b) 6.5, (c) 4.5, (d) 2.5, and (e) 0 dB away from the AWGN channel capacity at a spectral efficiency of 5.5 bits per polarization. The values of the system parameters are given in Table 1.



[FIG3] **The block diagram of CM schemes: (a) MLCM, (b) BICM, (c) TCM, (d) nonbinary, and (e) polar nonbinary.**

defined as the gross coding gain scaled by the code rate of the coded system to compare the coded and uncoded systems at the same information bit rate [8]. The NCG of a system at a certain information BER can be expressed as $\text{NCG} = R\gamma_{\text{uncod}}/\gamma$, where γ_{uncod} and γ are the SNRs required to meet the desired BER for the given uncoded and coded systems, respectively.

[TABLE 1] SYSTEM PARAMETER VALUES.

SYMBOL RATE R_s	32 Gbaud
NONLINEARITY COEFFICIENT γ	1.4 W^{-1} km^{-1}
ATTENUATION COEFFICIENT α	0.2 dB/km
DISPERSION COEFFICIENT D	17 ps/nm/km
OPTICAL CENTER WAVELENGTH λ	1,550 nm
EDFA NOISE FIGURE F_n	5 dB
SPAN LENGTH L	80 km

GAP TO THE AWGN CHANNEL CAPACITY

The advent of CM schemes in fiber-optic communications with soft-decision decoding enables new evaluation techniques for these systems. For a system with a rate R, there is a minimum SNR γ (in dB) to obtain a BER of 10^{-15} at the output of the CM decoder, which is usually computed by numerical simulations. The gap $\Delta\gamma$ between γ and the minimum SNR obtained using the Shannon formula for an AWGN channel with the spectral efficiency η, i.e., $2^\eta - 1$, is a useful measure to compare different CM schemes. The AWGN capacity, although popular as a benchmark, may not represent the capacity of the nonlinear fiber-optic channel [3]. This gap, known as gap from AWGN capacity [9], can be expressed as $\Delta\gamma = \gamma - 10\log_{10}(2^\eta - 1)$ dB. In Figure 2(b)–(e), we have shown the scatter plots of the received signal for a non-DM fiber-optic link with ten, 15, 23, and 39 spans and the system parameters given in Table 1, operating at 6.5, 4.5, 2.5, and 0 dB, respectively, from the AWGN channel capacity.

A CM SCHEME CAN EXPLOIT THE FOUR AVAILABLE DIMENSIONS OF A FIBER-OPTIC LINK, I.E., TWO POLARIZATIONS EACH CONSISTING OF IN-PHASE AND QUADRATURE DIMENSIONS, WITH MORE FLEXIBILITY THAN CONVENTIONAL SCHEMES.

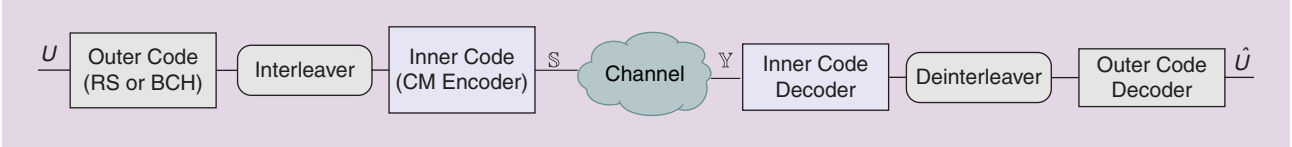
CM TECHNIQUES

Considering the bit-to-symbol mapper shown in Figure 1(b), the equivalent binary subchannels approach introduced in [10] can be applied to represent the mutual information (MI) between the channel input and the received signal after EDC as $I = \sum_{i=1}^{m} I_i$, where $I_i = I(V_i; \mathbb{Y} \mid V_1, \ldots, V_{i-1})$ is the conditional MI of subchannel i, provided that the transmitted bits of the subchannels $1, \ldots, i-1$ are given. The detection of the channel input bits is performed with a multistage decoder. An accurate channel model (see the section "Channel Model") is necessary to exploit this design framework. More precisely, this information-theoretic tool requires the signal statistics of the received signal $\mathbb{Y}$ from the channel. Clearly, the channel with input $\mathbb{S}$ and output $\mathbb{Y}$ can be modeled as m parallel subchannels with the inputs V_i, $i = 1, \ldots, m$ and the output $\mathbb{Y}$. An alternative parallel subchannel modeling approach is based on decoding the individual subchannels independently [10], which yields a sum rate of $\hat{I} = \sum_{i=1}^{m} \hat{I}_i$, in which $\hat{I}_i = I(V_i; \mathbb{Y})$. It can be shown that $I(V_i; \mathbb{Y}) \le I(V_i; \mathbb{Y} \mid V_1, \ldots, V_{i-1})$ [10], implying that $\hat{I} < I$. The gap between $\hat{I}$ and I strongly depends on the selected labeling of the constellation symbols. This gap is surprisingly small with Gray labeling. However, the multistage decoding technique is significantly superior to the parallel independent decoding for a finite-length code [10]. We explain below the three main categories of CM schemes, exploiting the equivalent subchannels for AWGN channels, as well as two CM schemes that are constructed from nonbinary component codes. They are all illustrated in Figure 3.

As shown in Figure 4, the CM schemes may be concatenated with an outer code to solve the problem of finding a coded scheme that has both a rapidly decreasing BER at moderate SNR, known as the waterfall region, and the possibility of reaching extremely low BERs without any error floor [11, Ch. 5]. As suggested in [8], one may use a capacity-approaching inner code, here realized by a CM scheme, to obtain BERs around 10^{-3}. Then the BER floor is suppressed using an outer code constructed based on classic codes with hard-decision decoding such as Reed–Solomon (RS) and Bose–Chaudhuri–Hocquenghem (BCH) codes to BERs acceptable for optical communications, e.g., 10^{-15}. The distributions of the received 2-D or 4-D symbols before decoding are computed using the noise variance given in the section "Channel Model."

MULTILEVEL CODED MODULATION

For an arbitrary modulation, the binary subchannels have in general different conditional MIs I_i. Hence, to approach the channel MI I, an unequal error protection technique, as depicted in Figure 3(a), is applied over the m binary subchannels. To this end, multilevel CM (MLCM) was designed consisting of m binary turbo [10] or low-density parity check (LDPC) [12] codes, originally introduced with classic block codes [13], each adapted to the conditional MI of the corresponding subchannels (I_i for channel i). MLCM has been shown to be a capacity-achieving scheme theoretically and through simulations [10] for AWGN. An interesting feature of MLCM is the possibility of exploiting a multistage decoder (MSD). As shown in Figure 3(a), the decoder of the first



[FIG4] The concatenation of an outer (RS or BCH) and inner (CM scheme) codes.

subchannel can decode the received bits independently of the other subchannels, then the second decoder uses the output from the first decoder to decode the bits received in the second subchannel, and so on for the rest of the subchannels. The MSD has lower complexity than the maximum-likelihood detector. An MLCM scheme was tailored in [14] for a memoryless nonlinear fiber-optic channel with RS component codes. In this paper, an unequal error protection scheme in the phase and radial direction of a 16-point ring constellation is exploited to minimize the block error rate of the system. For non-DM fiber-optic channels, two simplified MLCM schemes were introduced in [15] with staircase codes and LDPC codes, respectively. The subchannels are categorized in two groups in [15] and three groups in [16], to reduce the number of component codes.

BIT-INTERLEAVED CODED MODULATION

Zehavi [17] introduced bit-interleaved CM (BICM) as shown in Figure 3(b) simply by adding an interleaver between the encoder and the mapper to distribute the coded bits among different binary subchannels uniformly and exploit the diversity in the subchannels. In the BICM scheme, the subchannels are assumed to be independent and a simplified model using m independent decoders of the binary subchannels is used [10] with the MI $I(V_i; \mathbb{Y})$ for subchannel $i = 1, \ldots m$, in which each subchannel has no information from the input bits of the other subchannels. Usually, the binary decoder uses the log-likelihood ratios (LLRs) of the subchannels after deinterleaving to decode the received bits, where the LLR of bit v is defined as $\ln(\Pr(v = 1 \mid \mathbb{Y})/\Pr(v = 0 \mid \mathbb{Y}))$.

For channels such as wireless fast fading channels, the channel is unknown at the transmitter, and thus, the MIs of the subchannels are also unknown. BICM was originally proposed for fast-fading channels to exploit the diversity in binary subchannels [10]. BICM has been widely investigated in fiber-optic communications. For example, a comprehensive study of BICM for fiber-optic communications has been performed in [18] with different modulation formats. The performance of a BICM scheme is very sensitive to the type of the selected constellation labeling. Its performance is significantly degraded for a non-Gray labeling. To overcome this problem, one may exploit an iterative decoding between the 2-D or 4-D demapper (LLR calculation unit) and the binary code decoder [19].

TRELLIS CODED MODULATION

Ungerboeck [20] introduced a new type of binary labeling based on the set partitioning technique. The subchannels resulting from this labeling have ascending MI values. The early subchannels (with smaller indices) have lower MI values than the subchannels with indices close to m. The original version of trellis-CM (TCM), shown in Figure 3(c), splits the information bits into two groups of subchannels, where the group with smaller indices, the so-called "subset selection," is protected by a convolutional code, while the second group, denoted as "symbol selection," remains uncoded. Although this scheme can be decoded by MSD, Ungerboeck proposed a maximum likelihood decoder. The Viterbi decoder uses the subset metrics to decode the first group. The second group is decoded by a simple demapper within the decoded subset.

A capacity-approaching TCM scheme, known as turbo TCM, can be designed by replacing the convolutional code with a turbo code to decrease the gap from the Shannon limit for AWGN channels. Furthermore, multidimensional TCM was proposed in [21], which allows a higher spectral efficiency for a given signal constellation than one-dimensional (1-D) or 2-D TCM methods. In fiber-optic systems, TCM was proposed in [22] with an 8-point cubic polarization shift keying constellation. The simplest 4- and 16-state TCM schemes were applied to 8-point phase shift keying (PSK) and differential PSK in [23]. Finally, the concatenation of 2-D TCM with two different outer codes, RS and BCH codes, was studied in [24], which gives NCGs of 8.4 and 9.7 dB, respectively, at a BER of 10^{-13} for the AWGN channel.

CM SCHEME WITH A NONBINARY BLOCK CODE

The codewords of a nonbinary code are sequences of 2^q-ary symbols, each representing q bits. The code is constructed over a Galois field (GF) of order 2^q, denoted by $\mathrm{GF}(2^q)$. Binary codes can be considered as the simplest case of these codes, defined over GF(2) with two symbols zero and one. The binary subchannels can be encoded and decoded jointly using nonbinary codes, at the cost of increased complexity. As shown in Figure 3(d), the demapper computes symbol LLRs for each soft received symbol, retaining the MI between the subchannels compared to the independent bit LLR calculation in BICM. In fact, since symbol-wise decoding is used for a nonbinary scheme, its performance is not sensitive to the type of the selected constellation labeling and the decoding is performed with no iteration between the LLR calculation unit and the CM decoder.

Different types of nonbinary codes such as classic nonbinary codes, e.g., RS codes with a hard-decision decoding, or modern nonbinary LDPC and turbo TCM codes with a soft-decision decoding, can be used to construct the nonbinary CM schemes. Moderate-length (< 2,000 GF symbols) nonbinary LDPC codes have been widely proposed for fiber-optic communications [25], to approach the Shannon limit in AWGN channels. The nonbinary scheme can be used with both 2-D [25] and 4-D [16], [26] constellations.

POLAR NONBINARY CM SCHEME

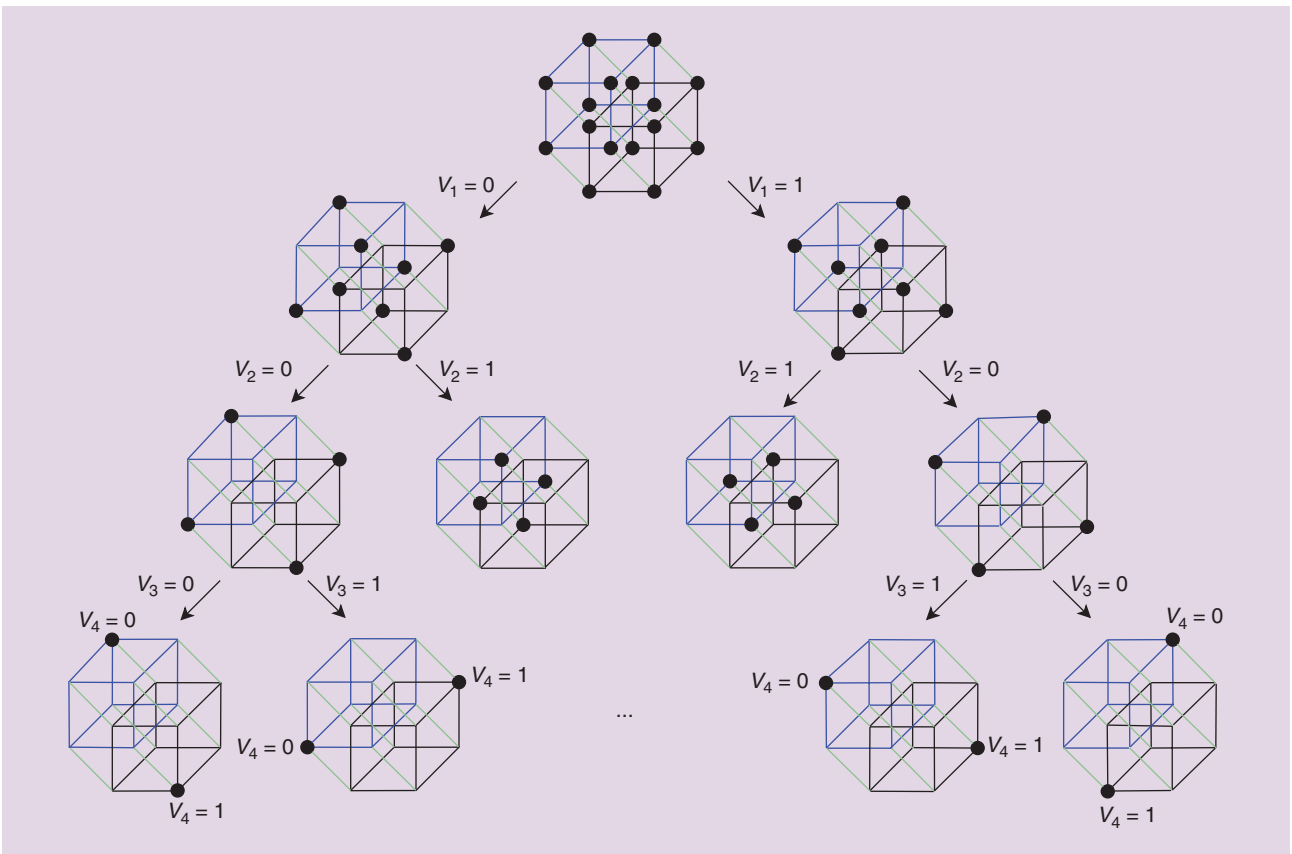
Although many techniques have been suggested to mitigate the computational complexity of nonbinary codes, the decoding complexity in the order of $O(q2^q)$, for a regular nonbinary LDPC code designed over $\mathrm{GF}(2^q)$, makes this scheme unrealistic for large ($\geq 2^7$ points) constellations [27]. To overcome this problem, a mapper, inspired by the polar coding technique [28], was devised [16] to categorize the binary subchannels into three groups: "bad," "intermediate," and "good" subchannels. The "bad" and "good" subchannels have MIs near zero and one, respectively, while the MIs of "intermediate" subchannels are between zero and one. Then, error protection using nonbinary

LDPC coding is performed solely over the "intermediate" subchannels. As shown in Figure 3(e), the "good" subchannels are left uncoded, whereas no information is transmitted on the "bad" subchannels denoted by dropped bits, which are fixed to zero and known to the receiver. Since the nonbinary encoder performs on the "intermediate" subchannels independently of the constellation size [16], the GF can have a lower order with this design than with the regular nonbinary scheme above, and consequently a CM scheme with a lower complexity is obtained. In this scheme, the bit-to-symbol mapper can be realized by a 4-D set partitioning technique illustrated using the bits $V_1, \ldots, V_4$ in Figure 5 for a PM-QPSK constellation [16].

2-D VERSUS 4-D CM SCHEMES

A CM scheme can exploit the available four dimensions in the signal space of a fiber-optic link either jointly as a 4-D channel or separately as two parallel 2-D channels. For the Gaussian noise model introduced in the section "Channel Model," these parallel channels are independent, as shown in [10], and one can get close to the MI of an AWGN channel using both 1-D and 2-D schemes. Although a 2-D CM scheme can achieve the MI of AWGN channels, a 4-D CM scheme has a better tradeoff between complexity and performance at the same spectral efficiency, as shown later in the performance analysis (see the section "Performance Analysis of 2-D and 4-D Schemes"). In fact, a 4-D scheme can provide more flexibility than 1-D or 2-D schemes, which facilitates exploiting rate adaptation and probabilistic shaping techniques. Here, we investigate 2-D and 4-D CM schemes with binary and nonbinary codes.

Classic and modern binary codes as well as their concatenations are used together with 2-D constellations such as QAM signals for constructing 2-D CM schemes. They are well investigated for fiber-optic communications and have been realized based on the three traditional CM schemes, i.e., MLCM [15], TCM [24], and BICM [18]. This group of CM schemes is capable of approaching the AWGN capacity provided that the block length is sufficiently large. For example, an NCG of 10.8 dB ($\Delta\gamma = 3$ dB) with 20.5% coding overhead is achieved with triple-concatenated codes, (4,608, 4,080) LDPC, (3,860, 3,824) BCH, and (2,040, 1,930) BCH using QPSK signals at a BER of 10^{-15} [8], where (n, k) denotes a block code with a codeword of length n bits and an input information vector of length k bits. As introduced in [25], the 2-D CM schemes can also be constructed using nonbinary codes. The (1,225, 1,088) LDPC code over GF(2^3) with 12.6% coding overhead provides an NCG of 9.4 dB ($\Delta\gamma = 2.3$ dB) at a BER of 10^{-10}. The improvement over the comparable binary (3,136, 2,800) LDPC code from the same family is 0.7 dB at a BER of 10^{-7}.



[FIG5] A 4-D set partitioning of a 16-ary 4-D constellation representing PM-QPSK. $V_4V_3V_2V_1$ represents the four bits in the binary labeling of the constellation [16].

CM schemes with 4-D constellations adopted from classical communication have been suggested for optical communications based on BICM. For example, a 4-D BICM scheme with two concatenated codes, an outer (992, 956) RS code and an inner (9,252, 7,976) LDPC code, can provide an NCG of 10.5 dB ($\Delta\gamma = 2.7$ dB) at a BER of 10^{-13} with an overall coding overhead of 20% and QPSK constellation [19]. In Figure 3(d) and (e), nonbinary codes are applied to 4-D CM schemes to improve the NCG of these systems, for example 0.29 dB, 1.17 dB, and 2.17 dB with 16-, 32-, and 64-point 4-D constellations, respectively, at a BER of 10^{-7} [26]. The nonbinary scheme in Figure 3(d) suffers from high complexity for constellations with a large number of symbols ($\geq 2^7$). The polar nonbinary CM scheme in Figure 3(e) decreases the complexity of the nonbinary CM schemes without performance degradation, by confining the required GF order of the nonbinary block code to a small number ($<2^7$ symbols), independent of the constellation size. Finally, it can be concluded that 4-D schemes may be more spectrally efficient than 2-D schemes at the same performance.

THE COMPLEXITY OF A CM SCHEME IS DOMINATED BY ITS TWO MAIN COMPONENTS: THE LLR CALCULATION FROM THE SOFT RECEIVED SYMBOLS AND THE ENCODER AND DECODER OF THE COMPONENT CODES.

HARDWARE REQUIREMENT AND DSP COMPLEXITY

The hardware requirements and electronic processing complexity of CM schemes play a crucial role for fiber-optic communications. Although the semiconductor technology is capable of providing ultra-high-speed analog-to-digital converters (ADCs) and massively parallelized DSP circuits, the system power consumption and hardware cost also need to be taken into account. In particular, since high-resolution ADCs and digital signal processors are costly for high-speed data transmission, the performance sensitivity of CM schemes to quantization errors has become an important factor in the design of these schemes [8]. The impact of quantization errors on the performance of a concatenated TCM scheme with two interleaved BCH outer codes was evaluated in [24], and it was shown that 4-bit quantization was sufficient to approach the infinite-precision performance to within 0.15 dB.

The complexity of a CM scheme is dominated by its two main components: the LLR calculation from the soft received symbols and the encoder and decoder of the component codes. To compute the LLR vector for a 4-D CM scheme, finding the closest 4-D symbol to the received vector among the constellation symbols requires approximately 4 times the computational complexity of finding the closest 1-D symbol in the constituent 1-D constellation, neglecting the three additions which may be needed to compute the 4-D minimum Euclidean distance from four 1-D minimum Euclidean distances [21]. This implies that one may compare the complexity of the receivers for CM schemes with different dimensions by taking into account solely the complexity of the component code decoders per dimension.

The complexity of LDPC and RS codes has been well studied in the literature. The computational complexity required per iteration of the fast Fourier transform sum-product algorithm in decoding a 2^q-ary regular nonbinary LDPC code designed over GF (2^q) is in the order of $O(J\rho q 2^q)$, where J and ρ are the number and weight of the rows of the parity-check matrix of the nonbinary LDPC code, respectively. This complexity is in the order of $O(q^2 2^q)$ for RS codes [11, Ch. 14]. Moreover, the number of iterations required for the convergence of LDPC iterative decoding also influences the complexity of the decoder of these codes.

PERFORMANCE ANALYSIS OF 2-D AND 4-D SCHEMES

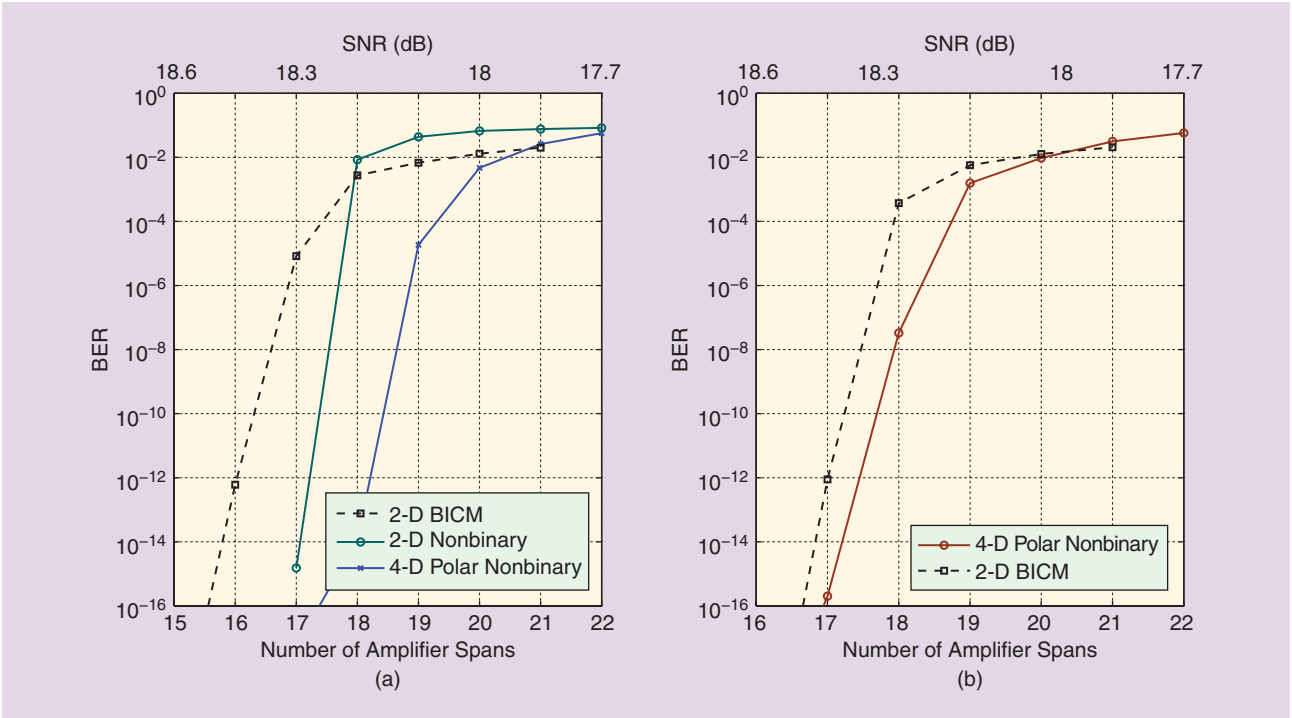
We compare the BER performance for three CM schemes: 2-D BICM, 2-D nonbinary CM, and 4-D polar nonbinary CM schemes, illustrated in Figure 3(b), (d), and (e), respectively. All schemes were designed with PM 64-QAM and an overall coding overhead of 21% over a single-channel non-DM fiber-optic link with the system parameters given in Table 1. The exploited LDPC codes were constructed based on finite fields [11, Ch. 11]. The numerical simulations of signal propagation in a non-DM fiber-optic link based on the Manakov equation are performed using the split-step Fourier method. Here, the schemes are compared based on two constraints: block length and complexity.

BLOCK-LENGTH-CONSTRAINED COMPARISON

Three systems are simulated with the same transmission block length consisting of inner and outer codes together with an interleaver as shown in Figure 4 for the following scenarios:

1) a 2-D BICM scheme with a (3, 21)-regular quasi-cyclic (a (γ,ρ)-regular quasi-cyclic LDPC code has γ nonzero elements in each column and ρ nonzero elements in each row of its parity-check matrix [11, Ch. 5]) binary (10,752, 9,236) LDPC inner code concatenated with a (1,016, 980) shortened RS outer code over GF(2^{10}), to bring down the output BER of the inner code from 2.2×10^{-4} to 10^{-15}
2) a 2-D nonbinary CM scheme with a (3, 9)-regular quasi-cyclic nonbinary (2,688, 2,309) LDPC inner code over GF(2^6) concatenated with a (970, 930) shortened RS code over GF(2^{10}), to bring down the output BER of the inner code from 1.9×10^{-4} to 10^{-15}
3) a 4-D polar nonbinary CM scheme with a (3, 9)-regular quasi-cyclic nonbinary (1,728, 1,162) LDPC inner code over GF(2^6) concatenated with a (963, 949) shortened RS code over GF(2^{10}), to bring down the output BER of the inner code from 1.5×10^{-5} to 10^{-15}.

The length of the interleaver between the inner and the outer code is 11 times the inner code length for the 2-D BICM and seven times the inner code length for the 2-D nonbinary CM schemes, resulting in coded block lengths of $11 \times 10,752 = 118,272$ and $7 \times 2,688 \times 6 = 112,896$ bits, respectively. The interleaver length is five times the inner code length for the 4-D polar nonbinary CM scheme, resulting in a coded block length of $5 \times 1,728 \times 12 = 103,680$ bits. Considering transmission of



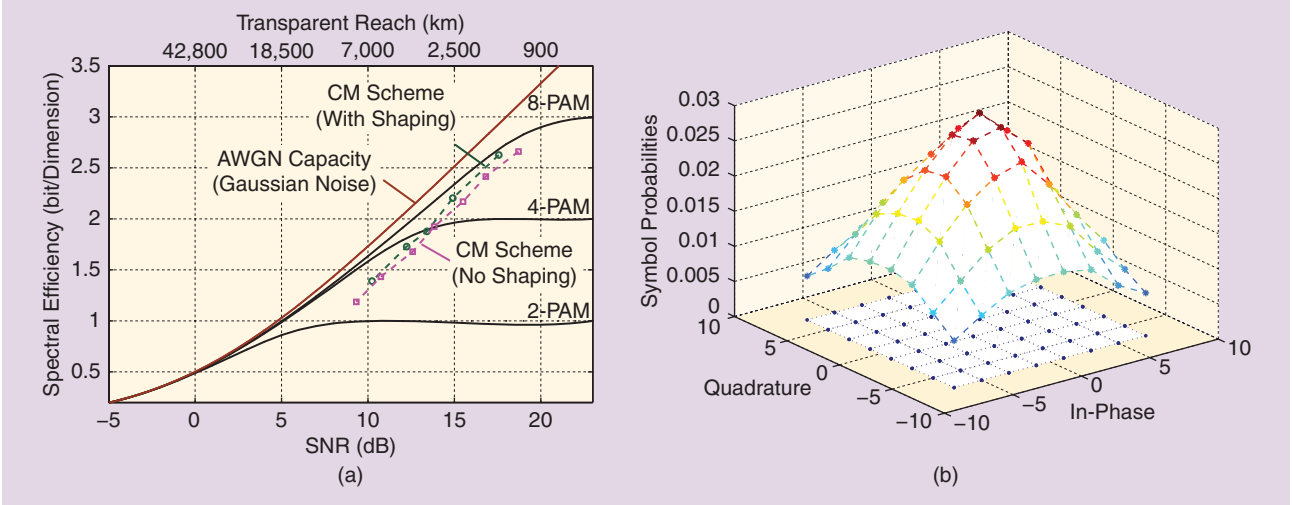
[FIG6] (a) The BER of three CM schemes with information-block-length-constraint. (b) The BER of 2-D and 4-D CM schemes with binary and nonbinary LDPC codes, respectively, and similar complexity. All the CM schemes use PM 64-QAM with 21% coding overhead and have therefore the same spectral efficiency.

12 bits by each 4-D symbol at 32 Gbaud, we obtain block lengths of 308, 294, and 270 ns for the 2-D BICM, 2-D nonbinary, and polar 4-D nonbinary schemes, respectively. According to the BER results shown in Figure 6(a), the polar 4-D nonbinary scheme is superior to the 2-D BICM and 2-D nonbinary CM schemes with nearly the same transmission block length.

COMPLEXITY-CONSTRAINED COMPARISON

We designed the following 2-D and 4-D schemes with similar complexities using the results provided in the section "Hardware Requirement and DSP Processing Complexity":

- a 2-D BICM scheme consisting of a (3, 21)-regular quasi-cyclic binary (16,128, 13,844) LDPC inner code concatenated



[FIG7] (a) The spectral efficiency per dimension versus the transparent reach and the SNR for a non-DM link with EDC. The CM scheme curves are based on the results given in [16] and the spectral efficiency for the Gaussian noise model is computed by $\log_2(1+\mathrm{SNR})/2$, where $\mathrm{SNR}=|\zeta|^2P/\sigma^2$. (b) The 2-D symbol probabilities of the probabilistically shaped 4-D CM scheme.

with a (1,015, 977) shortened RS outer code over GF(210), to bring down the output BER of the inner code from 2.3 × 10−4 to 10−15

- a 4-D polar nonbinary CM scheme consisting of a (3, 9)-regular quasi-cyclic nonbinary (1,152, 778) LDPC inner code over GF(26) concatenated with a (1,011, 995) shortened RS outer code over GF(210), to bring down the output BER of the inner code from 2.5 × 10−5 to 10−15.

As seen in Figure 6(b), the 4-D polar nonbinary scheme performs slightly better. Since the GF order can be kept fixed in this scheme, i.e., GF(2^6), independent of the constellation size, the 4-D scheme is superior to the 2-D scheme for large constellations.

SIGNAL SHAPING

Signal shaping in data transmission systems over AWGN channels refers to the manipulation of the symbol distribution to make it better approximate a Gaussian distribution [7]. Two types of shaping methods have been proposed for optical communications: probabilistic [15], [16] and geometric [29] shaping. Probabilistic shaping means changing the symbol probabilities for a standard constellation such as QAM, while geometric shaping implies changing the coordinates of the points in the constellation, which typically results in irregular (nonuniform) constellations. Two well-established probabilistic shaping methods, shell mapping and trellis shaping [7], have been applied to fiber-optic communications in [16] and [15], respectively. With probabilistic shaping, instead of having a uniform distribution for the input symbols, the symbols close to the origin of the constellation (with small amplitudes) are sent more often than the symbols far from the origin, as illustrated in Figure 7(b) for a 64-QAM with the shell mapping algorithm. Probabilistic shaping reduces the average transmitted power compared with a uniform distribution. Bearing in mind that the variance of the introduced nonlinear distortion is cubic with input power (see the section "Channel Model"), the system performance improves by performing probabilistic shaping as shown in Figure 7(a) [16].

RATE-ADAPTIVE CM SCHEMES

To improve the utilization of optical networks with dynamic or heterogeneous structure, the rate of the CM scheme can be adapted according to the CSI at the transmitter of each fiber-optic link. Two well-known choices for the CSI are the SNR, which is estimated after EDC, and the inner code BER, which is computed by a syndrome-based error estimator [9]. Rate-adaptive schemes have been investigated using multiple codes with different rates or a single fixed-rate code [9], [16], [30]. Different code rate can be constructed either separately or by puncturing or shortening a single mother code. For example, a rate-adaptive nonbinary scheme with six nonbinary LDPC codes was proposed in [30] to provide a transmission bit rate between 100 Gb/s and 300 Gb/s in steps of 26.67 Gb/s at a fixed symbol rate. In a more practical scenario, a rate-adaptive BICM scheme was proposed exploiting six combinations of binary LDPC and RS codes together with three modulations formats [9].

The method based on multiple codes with different rates is demanding in terms of hardware and thus costly to implement. A 4-D scheme with a flexible structure can perform rate adaptation with a single component code rather than using a different code for each rate. The 4-D scheme shown in Figure 3(e) was used in [16] to devise a rate-adaptive scheme with a single fixed-rate encoder. In this scheme, the number of bits in the different "good" and "bad" groups introduced in the polar CM scheme in the section "Polar Nonbinary CM Scheme" are adjusted according to the CSI such that the number of "intermediate" bits is always the same. Since the mapper is solely a simple look-up table, the rate adaptation is straightforward to implement. As shown in Figure 7(a), the rate-adaptive CM scheme using a single nonbinary code with probabilistic shaping can achieve $\Delta\gamma < 3$ dB for transparent reaches from 17 × 80 to 112 × 80 km.

TWO WELL-KNOWN CHOICES FOR THE CSI ARE THE SNR, WHICH IS ESTIMATED AFTER EDC, AND THE INNER CODE BER, WHICH IS COMPUTED BY A SYNDROME-BASED ERROR ESTIMATOR.

SUMMARY

To utilize the available resources in an optical network efficiently, the tradeoff between spectral efficiency, DSP hardware complexity, and transparent reach needs to be optimized for different links in the network. Joint coding and modulation schemes offer more freedom to exploit the available four dimensions in these channels than traditional independent FEC and modulation techniques. As discussed, a CM scheme can operate over a link with larger transparent reach than conventional schemes but with the same complexity (or even lower), for a wide range of spectral efficiencies.

Among the CM schemes discussed for AWGN channels, specifically, MLCM, BICM, TCM, nonbinary, and polar nonbinary schemes, MLCM is not attractive for fiber-optic communications because of its large number of component codes. The main bottleneck of nonbinary schemes is the decoding complexity, making it an unrealistic solution for large constellations. A better tradeoff between DSP complexity and transparent reach of 4-D CM schemes makes them superior to 2-D schemes. Finally, a 4-D CM scheme provides more flexibility than 1-D and 2-D CM schemes, which facilitates its combination with signal shaping techniques as well as rate adaptation methods with no need for multiple component codes.

AUTHORS

Lotfollah Beygi (beygil@chalmers.se) received his Ph.D. degree from Chalmers University of Technology, Göteborg, Sweden, in 2013. He was with the R&D division of Zaeim Electronic Industries from 2003 to 2008. Currently, he is with Qamcom Research and Technology AB. His main research interests include CM and DSP for fiber-optical and wireless communications.

Erik Agrell (agrell@chalmers.se) received his Ph.D. degree in information theory in 1997 from Chalmers University of Technology, Sweden. From 1997 to 1999, he was a postdoctoral researcher with the University of California, San Diego, and the University of Illinois at Urbana-Champaign. In 1999, he joined the faculty of Chalmers University of Technology, where he has been a professor in communication systems since 2009. In 2010, he cofounded the Fiber-Optic Communications Research Center at Chalmers, where he leads the signals and systems research area. His research interests are in information theory, coding theory, and digital communications, and his favorite applications are found in optical communications. He was the publications editor of *IEEE Transactions on Information Theory* from 1999 to 2002 and has been an associate editor of *IEEE Transactions on Communications* since 2012. He is a recipient of the 1990 John Ericsson Medal, 2009 ITW Best Poster Award, 2011 GlobeCom Best Paper Award, 2013 CTW Best Poster Award, and 2013 Chalmers Supervisor of the Year Award. He is a Senior Member of the IEEE.

Joseph M. Kahn (jmk@ee.stanford.edu) is a professor of electrical engineering at Stanford University. His research addresses communication and imaging through optical fibers, including modulation, detection, signal processing, and spatial multiplexing. He received the A.B. and Ph.D. degrees in physics from the University of California (U.C.), Berkeley, in 1981 and 1986, respectively. From 1987 to 1990, he was with AT&T Bell Laboratories, Crawford Hill Laboratory, in Holmdel, New Jersey. He was on the electrical engineering faculty at U.C. Berkeley from 1990 to 2003. In 2000, he cofounded StrataLight Communications, which was acquired by Opnext, Inc. in 2009. He received the National Science Foundation Presidential Young Investigator Award in 1991. He is a Fellow of the IEEE.

Magnus Karlsson (magnus.karlsson@chalmers.se) received his Ph.D. degree in 1994 from Chalmers University of Technology, Gothenburg, Sweden. Since 1995, he has been with the Photonics Laboratory at Chalmers, as assistant professor and, since 2003, as a professor in photonics. He has authored or coauthored over 240 scientific journal and conference contributions in the areas of nonlinear optics and fiber-optic transmission, and he cofounded the Fiber-Optic Communications Research Center at Chalmers in 2010. He has served on the technical committee for the Optical Fiber Communication Conference and currently is on the technical program committees for the European Conference of Optical Communication and the Asia Communications and Photonics Conference. He has been an associate editor of *Optics Express* since 2010. He received the CELTIC Excellence Award in 2011, Best Paper Award at GlobeCom 2011, and was appointed Fellow of the Optical Society of America in 2012.

REFERENCES

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightwave Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[2] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the non-linear Shannon limit," *J. Lightwave Technol.*, vol. 28, no. 4, pp. 423–433, Feb. 2010.

[3] E. Agrell. (2012). On monotonic capacity–cost functions. [Online]. Available: http://arxiv.org/abs/1108.0391

[4] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *J. Lightwave Technol.*, vol. 30, pp. 2011–2024, June 2012.

[5] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightwave Technol.*, vol. 30, no. 24, pp. 3857–3879, Dec. 2012.

[6] L. Beygi, E. Agrell, P. Johannisson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.

[7] G. D. Forney, Jr., and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998.

[8] F. Chang, K. Onohara, and T. Mizuochi, "Forward error correction for 100 G transport networks," *IEEE Commun. Mag.*, vol. 48, no. 3, pp. S48–S55, Mar. 2010.

[9] G.-H. Gho and J. M. Kahn, "Rate-adaptive modulation and low-density parity-check coding for optical fiber transmission systems," *J. Optical Commun. Netw.*, vol. 4, no. 10, pp. 760–768, Oct. 2012.

[10] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1361–1391, July 1999.

[11] W. E. Ryan and S. Lin, *Channel Codes: Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[12] J. Hou, P. Siegel, L. Milstein, and H. Pfister, "Capacity-approaching bandwidth-efficient coded modulation schemes based on low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 49, no. 9, pp. 2141–2155, Sept. 2003.

[13] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inform. Theory*, vol. 23, pp. 371–377, May 1977.

[14] L. Beygi, E. Agrell, P. Johannisson, and M. Karlsson, "A novel multilevel coded modulation scheme for fiber optical channel with nonlinear phase noise," in *Proc. IEEE Global Communication Conf.*, Dec. 2010.

[15] B. P. Smith and F. R. Kschischang, "A pragmatic coded modulation scheme for high-spectral-efficiency fiber-optic communications," *J. Lightwave Technol.*, vol. 30, no. 13, pp. 2047–2053, July 2012.

[16] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Rate-adaptive coded modulation for fiber-optical communications," *J. Lightwave Technol*, to be published.

[17] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873–884, May 1992.

[18] I. B. Djordjevic, M. Arabaci, and L. L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightwave Technol.*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.

[19] H. Bülow and E. Masalkina, "Coded modulation in optical communications," in *Proc. Optic Fiber Communication Conf.*, Mar. 2011.

[20] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[21] L. F. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Trans. Inform. Theory*, vol. 33, no. 4, pp. 483–501, July 1987.

[22] S. Benedetto, G. Olmo, and P. Poggiolini, "Trellis coded polarization shift keying modulation for digital optical communications," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1591–1602, Feb./Mar./Apr. 1995.

[23] H. Zhao, E. Agrell, and M. Karlsson, "Trellis-coded modulation in PSK and DPSK communications," in *Proc. Eur. Conf. Exhibition Optics Communication*, Sept. 2006, paper We3.P.93.

[24] M. Magarini, R.-J. Essiambre, B. E. Basch, A. Ashikhmin, G. Kramer, and A. J. de Lind van Wijngaarden, "Concatenated coded modulation for optical communications systems," *IEEE Photon. Technol. Lett.*, vol. 22, no. 16, pp. 1244–1246, Aug. 2010.

[25] I. Djordjevic and B. Vasic, "Nonbinary LDPC codes for optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2224–2226, Oct. 2005.

[26] M. Arabaci, I. Djordjevic, L. Xu, and T. Wang, "Four-dimensional non-binary LDPC-coded modulation schemes for ultra-high-speed optical fiber communication," *IEEE Photon. Technol. Lett.*, vol. 23, no. 18, pp. 1280–1282, Sept. 2011.

[27] D. Declercq and M. Fossorier, "Decoding algorithms for nonbinary LDPC codes over GF(q)," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 633–643, Apr. 2007.

[28] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[29] I. B. Djordjevic, H. G. Batshon, L. Xu, and T. Wang, "Coded polarization-multiplexed iterative polar modulation (PM-IPM) for beyond 400 Gb/s serial optical transmission," in *Proc. Optic Fiber Communication Conf.*, Mar. 2010, paper OMK2.

[30] M. Arabaci, I. B. Djordjevic, L. Xu, and T. Wang, "Nonbinary LDPC-Coded modulation for high-speed optical fiber communication without bandwidth expansion," *IEEE Photon. J.*, vol. 4, no. 3, pp. 728–734, June 2012.

[SP]

[Ivan B. Djordjevic, Milorad Cvijetic, and Changyu Lin]

Multidimensional Signaling and Coding Enabling Multi-Tb/s Optical Transport and Networking

[Multidimensional aspects of coded modulation]

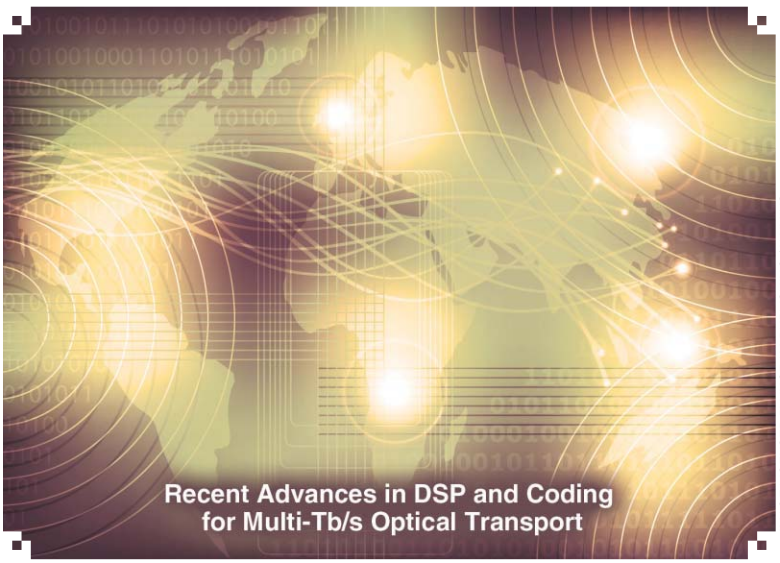


IMAGE LICENSED BY INGRAM PUBLISHING/JANMIKS

The design of next-generation optical transmission systems and networks should address the concerns with respect to a limited bandwidth of information infrastructure, high energy consumption, as well as the need to support the network heterogeneity and demand for an elastic and dynamic bandwidth allocation. To address these concerns simultaneously, we propose an adaptive, software-defined, low-density parity check (LDPC)-coded multiband approach that involves spatial-multiple-input, multiple-output (MIMO) and an all-optical orthogonal frequency-division multiplexing (OFDM) scheme since it can enable energy efficient high-bandwidth delivery with fine granularity and elastic model of bandwidth utilization. The modulation is based on multidimensional signaling to improve the tolerance to fiber nonlinearities and imperfect compensation of channel impairments and has a hybrid nature with both electrical and optical degrees of freedom employed. Optical degrees of freedom include spatial and

Digital Object Identifier 10.1109/MSP.2013.2275989
Date of publication: 12 February 2014

1053-5888/14/$31.00©2014IEEE

polarization modes in optical fibers supporting spatial-division multiplexing (SDM), while electrical degrees of freedom are based on $2M$ orthogonal basis functions. The adaptive coding has been performed by partial reconfiguration of the corresponding parity-check matrix. The proposed scheme is suitable for the conveyance of the information over optical fibers with bit rates exceeding 10 Tb/s. At the same time, the multitude of degrees of freedom will enable finer granularity and elasticity of the bandwidth, the features essential for next generation networking.

INTRODUCTION AND MOTIVATION

The exponential trend of Internet traffic growth is expected to continue, which will place the enormous capacity demand on the optical networking infrastructure. The standardization efforts to define high-speed Ethernet for optical networking are under way with the 100-Gb/s Ethernet (100 GbE) standard adopted few years ago [1], while 400 GbE and 1-Tb/s Ethernet (1 TbE) standardization should be completed in the near future [2]. The terabit Ethernet became a synonym for future transmission and networking. There are several technologies that can potentially be used to achieve bit rates of 1 Tb/s and beyond, such as optical time-division multiplexing (OTDM) [3], multiband OFDM [4]–[8], including so-called superchannel OFDM structure [6] or SDM over few-core fibers (FCFs) and/or few-mode fibers (FMFs) [4], [7]–[14]. These technologies have also been evaluated as enablers for envisioned 4 Tb/s- and 10 Tb/s-based optical transmission. The task to support terabit-based transmission is a complex one since it should take into account the limited bandwidth of information infrastructure, the need for reduction of energy consumption, and heterogeneity and dynamic nature of the optical networks.

THE HISTORICAL TREND IS TO SHIFT TO THE MULTIDIMENSIONAL CHARACTER OF OPTICAL TRANSMISSION WITH A NUMBER OF PARAMETERS THAT CAN BE ARRANGED IN DIFFERENT WAYS TO INCREASE THE CAPACITY AND RESILIENCE TO IMPAIRMENTS.

Historically, the on-off keying (OOK) modulation format has been dominant in optical fiber communications for more than three decades until multilevel modulation schemes coupled with coherent detection were revisited approximately seven years ago. Also, dense wavelength division multiplexing (DWDM) multichannel transmission with defined carrier wavelength spacing has been introduced to increase the overall capacity and to serve as a networking tool [14]. Transmission capacity and system performance has been gradually improved by usage of M-QAM modulation scheme. The first step was to apply quadrature phase-shift keying (QPSK = 4-QAM) modulation in combination with polarization-division multiplexing (PDM-QPSK). This scheme has been dominant in all application scenarios for 100 GbE. The total capacity can be doubled if 16-QAM format is considered instead of PDM-QPSK, which is already proven in some of experiments [15]. The employment of DWDM systems with PDM-M-QAM schemes was considered with either 50 or 25-GHz optical channel spacing. In parallel, there has been a consideration of more advanced forward error correction (FEC) schemes, which eventually led to the adoption of so-called soft-decision schemes, such as LDPC-based ones.

To increase bit rate per channel beyond 100 Gb/s, a multicarrier (subcarrier) channel design is proposed, which is known as a supercarrier channel concept [6]. The subcarriers in the channel are considered as an entity to be routed together, although the different arrangement with bandwidth partition of the channel can also be considered. The coherent optical (CO)-OFDM or Nyquist-based WDM have been introduced as a building block of superchannels [6], [16], [17]. Both CO-OFDM and Nyquist-WDM are based on the concept that signal overlapping either in the spectral domain or time domain is allowed as long as the orthogonality between subcarriers is preserved, as proven in lab or field experiments [13]. The superchannel design has also been a strong stimulus for a flexible grid channel spacing concept, which is evolution from the ITU-T-based fixed wavelength grid [2].

The SDM in SDM fibers (FMF, FCF, FMF-FCF) represents an additional way to increase the optical transmission capacity, with a number of experiments, such as [9], that prove its viability. In some cases, the FCF and FMF can be combined and the total number of spatial modes can exceed 20 [13]. It is important to mention that SDM fibers permit generation of the OAM (orbital angular momentum) modes [4], [18], a special category of spatial modes with azimuthal angular dependence.

The historical trend is to shift to the multidimensional character of optical transmission with a number of parameters that can be arranged in different ways to increase the capacity and resilience to impairments. In addition, the multidimensional character has always been a solid foundation for exploring a more advanced networking concept [14], [19]–[20].

In this article, we consider a further improvement of a multidimensional concept of transmission and networking by exploring the available degrees of freedom in a more effective way. We propose the use of hybrid multidimensional coded modulation (CM), employing both electrical and optical degrees of freedom, as an attractive candidate to address above constraints in a simultaneous manner. The optical degrees of freedom include the polarization and spatial modes in SDM fibers (both FMF and FCF). The electrical degrees of freedom we propose would include the orthogonal prolate spheroidal wave (OPSW) functions. The electrical degrees of freedom are used as the basis functions for multidimensional signaling. The use of multidimensional signaling brings two key advantages as compared to the conventional PDM-M-QAM schemes, including the following: 1) the Euclidean distance among signal constellation points for the same symbol energy can be increased as compared to two-dimensional (2-D) constellations, and 2) the nonlinear interaction of polarization modes in single mode fibers (SMFs) applications and nonlinear interaction among spatial modes in SDM fibers can be compensated for by using the reduced-complexity LDPC-coded turbo equalization [14], [21]. Notice that in our approach, all slowly varying channel impairments, including fiber nonlinearities, are

compensated for on the transmitter side, rather than receiver side. The reduced-complexity LDPC-coded turbo equalization scheme is then therefore responsible for an uncompensated random component of mode coupling. So far, the orthogonal subcarriers have been used as basis functions, which limits the number of orthogonal subcarriers that can be used as basis functions since the bandwidth expands as the subcarrier index increases [4], [8]. On the other hand, the use of modified orthogonal polynomials provides limited flexibility in terms of bit duration × bandwidth product [12]. The signal space in SMF applications is four-dimensional (4-D) and has been studied mainly in context to either improve spectral/power efficiency of conventional PDM-QAM [22], [23] or improve tolerance to fiber nonlinearities [24], rather than in the context of multi-Tb/s optical transport. The hybrid CM concept has already been considered by authors in [4] and [8], in which the electrical basis functions were based on complex orthogonal subcarriers. The increase in number of subcarriers, in such a scheme, reduces the overall spectral efficiency, which is not the case when OPSW functions are used since the occupied bandwidth is independent from the number of basis functions that are employed. This scheme represents generalization of polarization-multiplexed/spatial-division multiplexed schemes, which can be considered just as special cases of the proposed scheme. The scheme can be used to solve the "optical networks capacity crunch" problem, by using the commercially available electronics.

To address the constraints of heterogeneity and elasticity in optical networks, we envision the employment of simultaneous rate adaptation and signal constellation size selection to optimize the channel capacity of the observed transmission link. Accordingly, to meet the target bit-error rate (BER) requirement, the error correction strength is dynamically adjusted based on the optical channel conditions. The proposed adaptive software-defined LDPC-coded multidimensional scheme can be used in combination with the spectral-spatial multiplexing in a MIMO fashion to enable 10 + Tb/s serial optical transport. In such a scheme, the multidimensional signal constellations generated by the OSCD algorithm [23] are used in combination with properly designed irregular quasicyclic (QC)-LDPC codes.

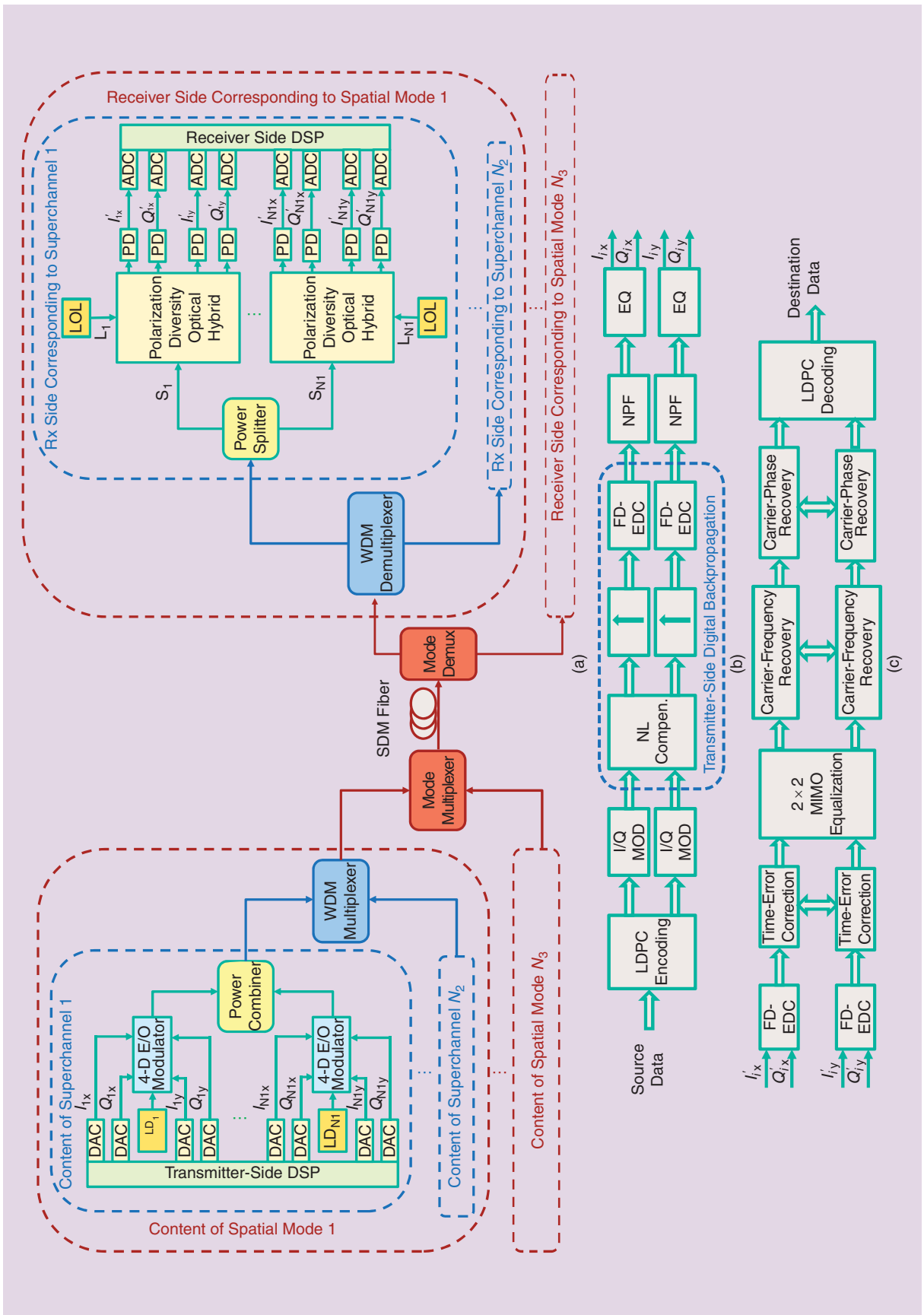
The main goal in this article has been to further explore the following innovative approaches:

- the adaptive spectral-spatial MIMO coded multidimensional scheme that can simultaneously address the high data rate demands, high energy consumption, and heterogeneity problems in current optical networking infrastructure
- the multidimensional constellations originated from an optimum signal constellation design algorithm have been used instead of the multidimensional ones obtained as the Cartesian product of lower-dimensionality constellations used in [4] and [8]
- the OPSW functions have been used as electrical basis functions instead of orthogonal subcarriers used in [4] and [8]
- a new class of irregular binary (BI) QC-LDPC codes, derived from pairwise-balanced designs (PBDs), is used instead of the conventional large-girth QC-LDPC code design used in [4] and [8]
- the proposed class of irregular BI QC-LDPC codes has been generalized into a nonbinary (NB) LDPC-code design
- the mode-multiplexed NB LDPC-CM scheme, a particular instance of proposed spectral-spatial MIMO scheme, is evaluated, for the first time, in transmission scenario over FMFs
- the dynamic allocation of spectrum with finer granularity enabled by the proposed spectral-spatial scheme.

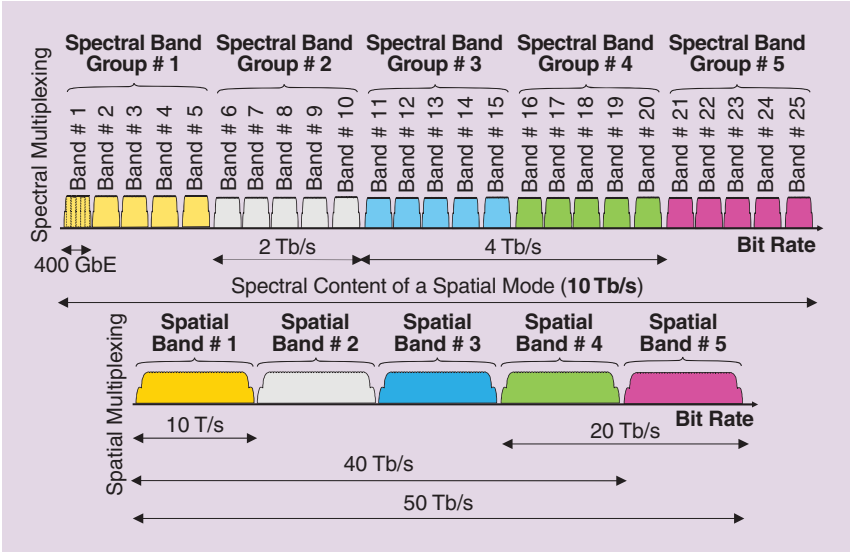
MULTIDIMENSIONAL SPECTRAL-SPATIAL PROCESSING AND DYNAMIC OPTICAL NETWORKING

A generic superchannel conventional transponder employing both WDM and mode-division multiplexing (MDM) is shown in Figure 1(a). The term *superchannel* is used to denote collection of optical carriers modulated and multiplexed together at the transmitter side. However, this term can also be used in a more specific case to denote a multiband optical OFDM signal, with center frequencies of the bands being mutually orthogonal, such as in the case described in [5] and [6]. As we see from Figure 1(a), the superchannel content is generated by applying 4-D signals (two phase channels and two polarizations) on a single optical carrier (or better to say an optical subcarrier) and corresponding 4-D modulator, being composed of two I/Q modulators, one polarization beam splitter and one polarization beam combiner, with details given in [4]. The signals of all optical subcarriers within the superchannel are power combined and led to a WDM multiplexer that also collects the outputs from all superchannels residing within the same spatial mode. The WDM multiplexer covers the entire band (C-band or L-band) and presents the second tear in the spectral arrangements of the content of individual spatial modes. The spatial modes are combined by spatial multiplexer and directed to the input of SDM fiber links. The process at the receiving side is just opposite with the spatial demultiplexing taking place first, followed by a two-tear spectral decomposing, which all results in selection of individual optical subcarriers loaded with 4-D electrical signals. Further processing of the 4-D signals is done in a conventional coherent detection scheme with balanced optical receivers, which is followed by the ADC and DSP processing blocks. The details of the supercarrier transmitter side DSP and receiver side DSP blocks are provided in Figure 1(b) and (c), respectively. The channel equalization can be successfully implemented by separating slow-varying channel impairments, such as chromatic dispersion, from time-varying channel impairments such as polarization mode dispersion (PMD). Slow-varying channel impairments can be precompensated at the transmitter side, while the time-varying channel impairments are compensated at the receiver side. The chromatic dispersion is an exception since it causes a huge memory effect, thus it makes sense to split the complexity of frequency-domain electronic dispersion compensation (FD-EDC) between transmitter and receiver.

In accordance with this interpretation, at the transmitter side we additionally perform nonlinear compensation (NLC), Nyquist prefiltering (NPF), and equalization (EQ) for imperfections of the transmitter. The NLC and EDC blocks can be combined into a transmitter-side digital backpropagation module, as shown in Figure 1(b).



[FIG1] (a) A generic conventional superchannel WDM/MDM scheme, (b) the configuration of transmitter side DSP blocks corresponding to the ith component of superchannel ($i = 1, \ldots, N_1$), and (c) the configuration of receiver side DSP blocks. N_1 denotes the number of frequency-locked laser signals within the superchannel; N_2 denotes the number of superchannels; N_3 denotes the number of spatial modes. LD: laser diode, E/O: electro-optical, PD: photodetector, ◧: up-sampling block, LOL: local oscillator laser.



[FIG2] The conceptual description of spectral-spatial processing enabling up to 50-Tb/s serial optical transport.

At the receiver side, we perform the correction of timing errors, adaptive 2×2 MIMO equalization of PMD effects and polarization rotation, carrier frequency recovery, and carrier phase recovery, all in addition to the compensation of the receiver portion of chromatic dispersion. The scheme from Figure 1 provides finer granularity as either single band or multiple bands can easily be simultaneously selected. By frequency locking of WDM channels, the superchannel signals can further be combined into even larger entities. This conventional scheme is already multidimensional as it employs both the electrical basis functions (specifically in-phase and quadrature signals) and optical basis functions, two polarization states and N spatial modes. By replacing the two electrical basis functions (in-phase and quadrature channels) with the set of $2M$ OPSW functions, the corresponding scheme becomes $4MN$-dimensional scheme, which is the subject of the investigation in the section "Hybrid Coded-Modulation Employing Electrical and Optical Basis Functions."

The conceptual diagram of multidimensional spectral-spatial processing enabling up to 50 Tb/s serial optical transport is illustrated in Figure 2. The frame is flexible and can support various future Ethernet rates (10 Tb/s and beyond), which have been already envisioned [8], [14], but not yet standardized. The scheme is organized into five band groups with center frequencies being orthogonal to each other. Each spectral component caries 400 Gb Ethernet (400 GbE), while each spectral group carries 2 Tb/s traffic. We can employ a three-step hierarchical architecture with a building block being a 400 Gb/s signals originating either from 4×100 GbE channels, 10×40 GbE, or one 400 GbE channel, respectively. Also, several optical subcarriers of all-optical OFDM scheme can be used to create a superchannel structure. Next, 400 GbE spectral slots are arranged in spectral band groups to enable 2 Tb/s traffic. By combining two spectral band-groups, the scheme can enable 4 Tb/s traffic. We assume that five spectral band groups (and possibly more) can be aligned along the optical spectrum as a content of the spatial mode. The second layer is related to spectral-division multiplexing, resulting in 10 Tb/s aggregate data rate per spatial mode. By combining two (or four) spatial modes, the scheme is compatible with envisioned 20 Tb (40 Tb) data rate speeds. Finally, the fiber link layer is implemented by combining the signals from spatial modes to achieve 50 Tb/s optical transmissions. The proposed scheme represents one of the possible combinations since there is a flexibility of how the spectral-spatial arrangement is done. We should also notice that we assumed that $\sim25\%$ of the line bit rate is occupied by applied FEC schemes. Clearly, the concept of spectral band is equivalent to conventional superchannel concept, while the spectral band group is equivalent to combining the superchannels shown in Figure 1(a) by using a WDM multiplexer, providing that the wavelength channels in WDM multiplexers are frequency locked as described in [5]. The aggregate data rate of the proposed scheme is given as $2mRR_sN_1N_2N_3$, where m denotes number of bits per symbol, factor two denotes two polarizations, R is the code rate, and R_s is the symbol rate. Parameters $N_i\,(i=1,2,3)$ have already been defined in Figure 1. For instance, by setting N_i parameters as follows: $N_1=5$, $N_2=5$, and $N_3=8$; the aggregate data rate of 32-ary 4-D constellation is given by $2mRR_sN_1N_2N_3=2\times\log_2 32\times0.8\times31.25\ \text{GS/s}\times5\times5\times8=50$ Tb/s.

The spectral efficiency of this example is 64 bits/s/Hz, which is higher than the spectral efficiency of schemes based on orbital angular momentum [18].

It is clear that the spectral-spatial multiplexing scheme can also be considered as a foundation of spectral-spatial interworking as we proposed in [20]. The plurality of spatial and spectral degrees of freedom introduced multiple dimensions for networking, which can also be supported by elastic and dynamic modulation and coding described earlier. The conceptual scheme of the multidimensional optical network from [20] is presented in Figure 3. We can clearly recognize from Figure 3 that the building blocks for spatial mode components (multiplexers and demultiplexers, amplifiers, and optical add-drop multiplexers) are the key elements for an efficient implementation of multidimensional networking in the future.

As we see from Figure 3, the spectral slots can be interchanged among the spatial modes carried by SDM fibers or to serve for local connections (drop ports). These functions are done by using spatial mode multiplexers/demultiplexers and tunable optical filters (the tunability is performed in both the central wavelength and the spectral bandwidth). The third component that would be required to maximize the overall throughput is a spectral convertor. We should mention that the employment of spatial mode demultiplexers at the drop side assume that mode coupling effect should be addressed properly. The simplest option will be when there is no need to have any MIMO processing at the add/drop site, which is case when FCFs are used since crosstalk level is more

than 20 dB below the signal level. As for FCF, an optimized design of at most six modes supporting a weak mode coupling should be considered for this purpose. If there is no MIMO processing at the add/drop site, the numbers of such sites will be limited (with possibly just one site in case of FMF employment). We should also keep in mind that the FEC strength can be adjusted with respect to perceived drop side, thus helping with the crosstalk issue (please see the sections "Hybrid Coded-Modulation Employing Electrical and Optical Basis Functions" and "Quasi-Cyclic LDPC Codes and Their Related Families"). On the other side, we expect that MIMO processing technique for SDM fibers will gradually mature with more compact and efficient DSP blocks, which means that we should not exclude the option of having MIMO processing even at the add/drop sites.

AN OPTICAL NETWORKING CONCEPT BASED ON SPATIAL AND SPECTRAL ENTITIES, APPLIED WITH TRANSMISSION TECHNOLOGIES, OFFERS AN EFFICIENT SPECTRAL UTILIZATION WITH BANDWIDTH PROCESSING THAT WOULD INCREASE THE NETWORK THROUGHPUT AND ENHANCE ITS ELASTIC PROPERTIES.

An optical networking concept based on spatial and spectral entities, applied with transmission technologies described next, offers an efficient spectral utilization with bandwidth processing that would increase the network throughput and enhance its elastic properties.

HYBRID CODED-MODULATION EMPLOYING ELECTRICAL AND OPTICAL BASIS FUNCTIONS

As already mentioned in the previous section, an adaptive $4MN$-dimensional hybrid CM scheme can be considered as a generalization of conventional WDM/MDM scheme, shown in Figure 1. The adaptive software-defined QC-LDPC-coded multiband OFDM with spatial MIMO scheme that we propose is shown in Figure 4(a). For an easier explanation, only a single polarization state is shown, with no details with respect to synchronization and clock recovery circuits. (Please refer to Figure 1 for the exact location of synchronization and clock recovery circuits.) The independent adaptive irregular QC-LDPC-coded data streams are written into $m_i \times n\,(i \in \{x, y\})$ block-interleaver [see Figure 4(b)]. The m_i bits from the block-interleaver are taken column-wise and used to select the coordinates of $2M$-dimensional signal constellation (employing $2M$ electrical basis functions). The configuration of corresponding modulator is provided in Figure 4(c). The even (odd) coordinates of $2M$-dimensional signal-constellation after up-sampling are passed through corresponding discrete-time (DT) pulse-shaping filters of impulse responses $h_m(n) = \Psi_m(nT)$, whose outputs are combined together into a single real (imaginary) data stream representing in-phase (quadrature) signal. After digital-to-analog conversion (DAC), the corresponding in-phase and quadrature signals are used as inputs to the I/Q modulator. The $2M$-dimensional modulator, shown in Figure 4(b), generates the signal constellation points given as

$$s_i = \sum_{d=0}^{2M-1} \psi_{i,d} \Psi_d, \tag{1}$$

where $\psi_{i,d}$ denotes the dth coordinate $d = 0, 1, 2, ..., 2M-1)$ of the i th signal-constellation point, and the set $\{\Psi_d\}$ denotes the set of basis functions (M basis functions correspond to in-phase channel, while other M basis functions correspond to quadrature channel). For instance, the pulse duration of the OPSW functions stays unchanged for all orders, while the bandwidth stays almost the same regardless of the order value. Specifically, the OPSW functions are simultaneously time-limited to symbol duration T_s and bandwidth-limited to band Ω, and can be obtained as solutions of the following integral equation [12], [25]:

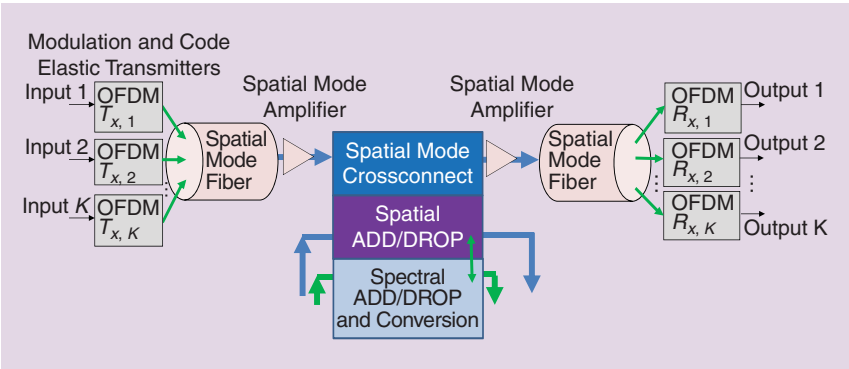
$$\int_{-T_s/2}^{T_s/2} \Psi_n(u) \frac{\sin \Omega(t-u)}{\pi(t-u)} du = \mu_n \Psi_n(t), \quad \mu_n \in (0,1], \tag{2}$$

where the coefficient μ_n is related to the energy concentration in the interval $[-T_s/2, T_s/2]$. The OPSW functions satisfy the double-orthogonality principle

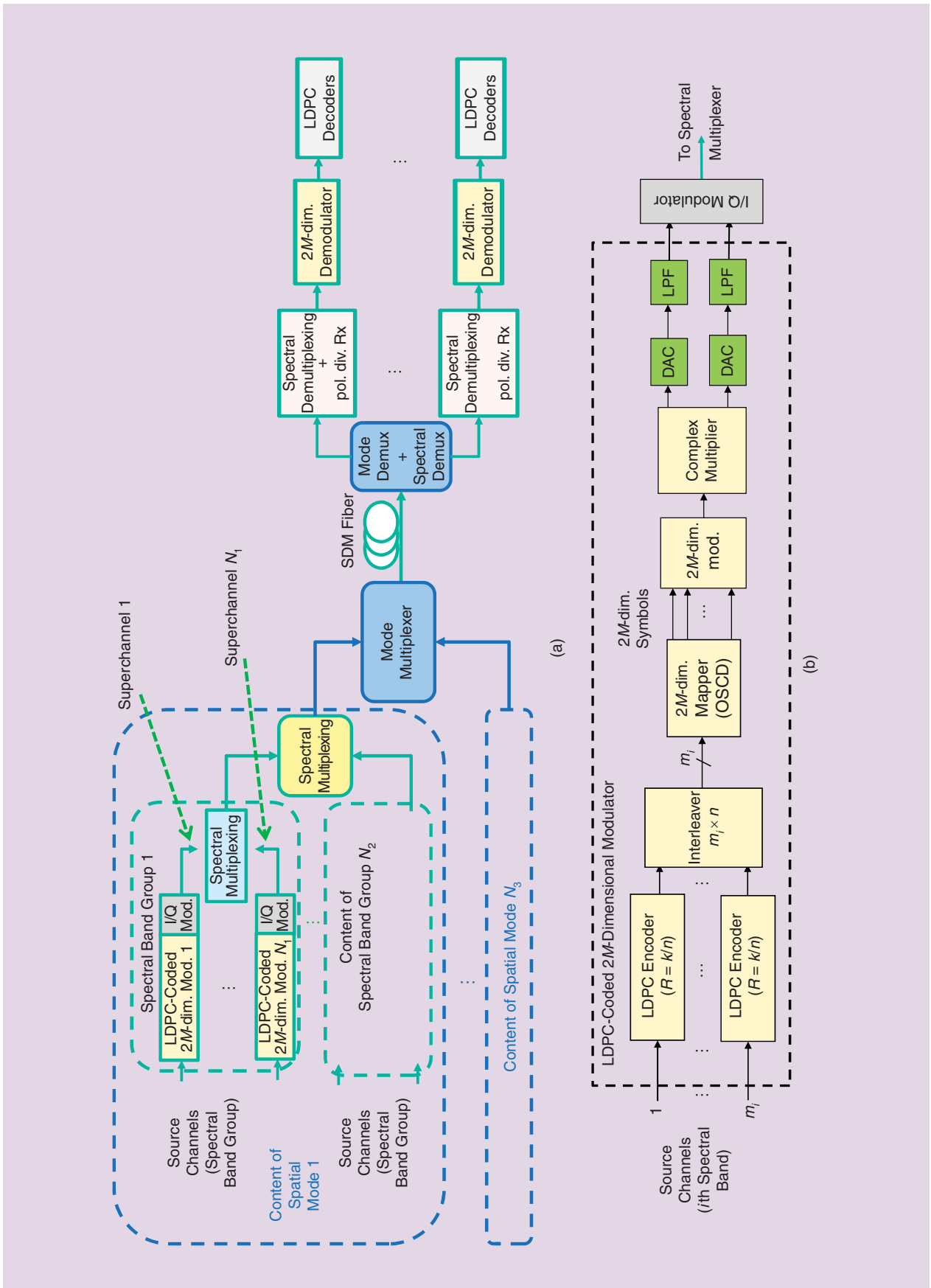
$$\int_{-T_s/2}^{T_s/2} \Psi_n(u)\Psi_m(u)du = \mu_n \delta_{nm}, \int_{-\infty}^{\infty} \Psi_n(u)\Psi_m(u)du = \delta_{nm}, \tag{3}$$

and, as such, these functions are well suited for the optical telecommunication applications. Alternatively, the set of M complex orthogonal polynomials [12] can be used instead of a set of $2M$ real basis functions.

The band selection within the band group is performed by complex multiplication with the $\exp(j2\pi f_n kT)$-term (T is the sampling interval), as shown in Figure 4(b), where f_n is the center frequency of the nth band in the band group. Such obtained signals are initially spectrally multiplexed to create the spectral band group. The spectral multiplexing can be achieved by the complex



[FIG3] **The structure of multidimensional spectral-spatial networking.**

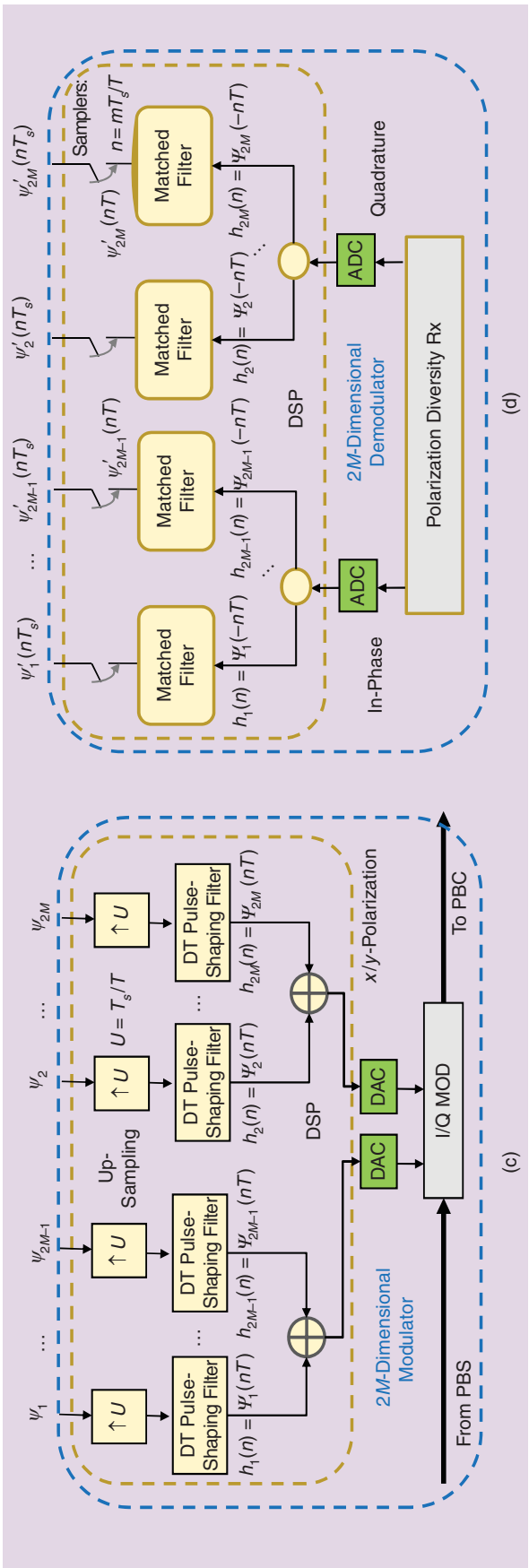


[FIG4] (a) A block diagram of an adaptive software-defined coded multiband optical-OFDM with spatial-MIMO. (b) Details of the adaptive LDPC-coded $2M$-dimensional modulator.

multiplication [to be performed in the electrical domain as shown in Figure 4(b)] of the corresponding $2M$-dimensional signals by $\exp[j2\pi(f_c + f_n)kT]$, where f_c is the central frequency of the cth spectral band group, and by combining them by a power coupler. Alternatively, the all-optical OFDM approach, similar to that from [6] and [26], can be used for both spatial bands and spatial band groups multiplexing. The corresponding spectral band-group signals are then coupled into SDM fibers by a mode-multiplexer, as shown in Figure 4(a). To facilitate the demodulation process, both the central frequencies of bands within the band group and frequencies among the band groups are properly chosen so that the principle of orthogonality is satisfied.

On the receiver side, after mode-demultiplexing, every mode projection is forwarded to the conventional polarization-diversity receiver, which provides the projections along the basis functions in both polarizations (and in-phase/quadrature channels). Each projection (in-phase/quadrature in either polarization) represents an M-dimensional electrical signal. Two M-dimensional projections (corresponding to x-/y-polarizations) are passed through analog-to-digital conversion (ADC) blocks and complex multiplier by $\exp[-j2\pi(f_c + f_n)kT]$, and used as inputs to corresponding matched filters with impulse responses $h_m(n) = \Psi_m(-nT)$, as indicated in Figure 4(d).

Finally, the resampled outputs represent projections along the corresponding basis functions, and these projections are used as inputs to the multidimensional a posteriori probability (APP) demapper, which calculates symbol log-likelihood ratios (LLRs). We iterate the extrinsic information between LDPC decoders and the APP demapper until convergence is achieved, or until a predetermined number of iterations has been reached. To compensate for the mode-coupling, optical MIMO detection principles described in [14] are used. If NB LDPC coding operating over $\mathrm{GF}(q_i = 2^{m_i})$ is used instead, then on the encoder side m_i BI LDPC encoders and a block-interleaver can be replaced by a single NB LDPC encoder. Similarly, multiple BI LDPC decoders on the receiver side can be replaced by a single NB LDPC decoder. This approach simplifies the transceiver design, whose overall complexity could be lower than that with the block-interleaved coded-modulation, as discussed in [27]. Specifically, in the links without in-line chromatic dispersion compensation, the distribution of samples upon compensation of chromatic dispersion and nonlinear phase distortions' compensation is Gaussian-like, as shown in the experimental study [28]. In this situation, the global iterations between the NB LDPC decoder and APP demapper lead to a negligible performance improvement. When mixed-domain fast Fourier transform-based q-ary sum-product algorithm is used for decoding, the decoder complexity is comparable or lower than that of BI block-interleaved coded-modulations, as explained in [27] (and references therein). Furthermore, the symbol LLRs can be calculated simply by



[FIG4 CONTINUED] (c) Details of a $2M$-dimensional modulator. (d) Configuration of $2M$-dimensional demodulator for $i \in \{x, y\}$ (i refers to the polarization modes): N_1 is the number of bands within the spectral band group; N_2 is the number of spectral band groups; N_3 is the number of spatial modes; and $2M$ denotes the number of electrical basis functions.

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

$$\mathrm{LLR}(s_j) = -\sum_{d=0}^{D-1} (r_d - s_{j,d})^2 / 2\sigma^2, \quad (4)$$

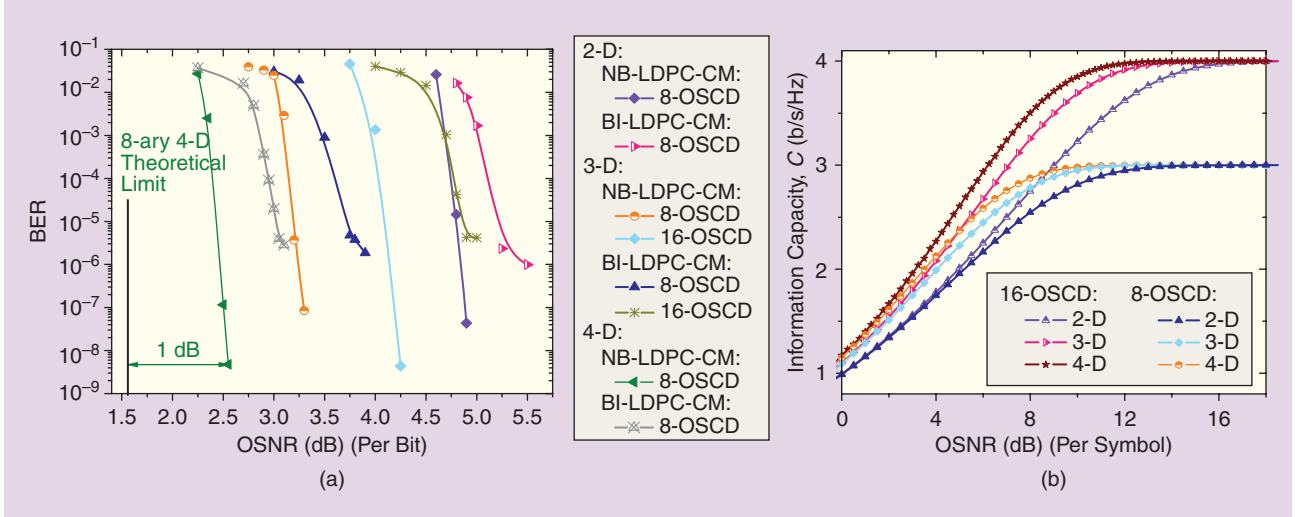
where r_d and $s_{j,d}$ denote the dth coordinates of received signal and symbol s_j, respectively, while σ^2 is the variance of an equivalent noise upon coherent detection. Therefore, even for a large number of coordinates the symbol LLRs are easy to calculate. When in-line chromatic dispersion compensation is applied, the resulting distribution is non-Gaussian in nature, and in this case it is more suitable to use the SDM of $2M$-dimensional signals instead of the full D-dimensional signaling, where $D = 4MN$.

TO PROVIDE SEAMLESS INTEGRATED TRANSPORT PLATFORMS THAT CAN SUPPORT HETEROGENEOUS NETWORKING, WE NEED TO CHANGE THE ERROR CORRECTION STRENGTH DEPENDING ON THE CHANNEL CONDITIONS.

As mentioned earlier, the rate-adaptive coding is an important component of this scheme due to a dynamic nature and different properties of individual different lightwave paths that can be established in optical networks. To provide seamless integrated transport platforms that can support heterogeneous networking, we need to change the error correction strength depending on the channel conditions. The code-rate adaptation in quasicyclic LDPC code design [4], [14], [27] can be performed by selecting different number of block-rows in corresponding parity-check matrix (please see the section "Quasi-Cyclic LDPC Codes and Their Related Families"). The code rate adaptation is performed by a partial reconfiguration of decoder by changing the size of corresponding block-submatrix of $\boldsymbol{H}$-matrix and/or by varying the number of employed block-rows while keeping codeword length intact. The aggregate data rate of the proposed scheme is given as $2(m_1+m_2)RR_sN_1N_2N_3$, where the factor two comes from two polarizations, R denotes the code rate, R_s denotes the symbol rate, N_1 denotes the number of bands within the spectral group, N_2 denotes the number of spectral groups, and N_3 denotes the number of spatial modes. The OSCD [23] signal constellation of size Q is decomposed in two subconstellations of sizes Q_i $(i = 1, 2)$ such that $Q = Q_1 + Q_2$. The first (second) incoming $m_1(m_2)$ bits are used to select the points from $Q_1 = 2^{m1}$-ary ($Q_2 = 2^{m2}$-ary) subconstellation. Two subconstellations are associated with in-phase and quadrature channels. The purpose of the constellation decomposition is to add more flexibility to the adaptive modulation and coding as the constellation sizes that are not power of two can be used, which allows a finer granulation of the aggregate data rate. Additionally, better immunity to in-phase-quadrature imbalance is obtained. The similar strategy is applicable in PDM systems. For instance, the aggregate data rate of 24-ary constellation is

$$\begin{aligned} 2(m_1+m_2)RR_sN_1N_2N_3 &= 2\times(1\log_2 16 + \log_2 8) \\ &\quad \times 0.8 \times 36\ \mathrm{GS/s} \times 5 \times 5 \times 5 \\ &= 50.4\ \mathrm{Tb/s}. \end{aligned}$$

For NB LDPC coding, two independent LDPC codes per each subconstellation have been used. To demonstrate the high potential of the proposed scheme, we have performed the Monte Carlo simulations, with results summarized in Figure 5(a). Both irregular NB and BI LDPC CMs are studied, designed as described in the next section. The OSCD-based modulation schemes are observed. As we see, BI-LDPC-CM schemes exhibit a known error floor phenomenon [14], [27], while the NB-LDPC-CM schemes do not exhibit



[FIG5] (a) BERs of irregular NB QC-LDPC-coded PDM spectrally multiplexed OSCDs. $N_1 = 5$, $N_2 = 4$, and channel symbol rate $R_s = 28.58$ GS/s. The BER results are obtained by averaging over all bands. The irregular QC-LDPC codes of code rate 0.875, length 27,616, girth-8, and average column-weight of three are used in simulations. (b) Information capacities of multidimensional OSCDs (per single polarization). All LDPC results for BI codes are based on (7), while these for NB codes are based on (8).

any error floor and significantly outperform corresponding BI counterparts. We have also calculated the information capacities for 8-ary and 16-ary OSCD-modulation schemes with results presented in Figure 5(b). The 8-ary 4-D coded OSCD-modulation scheme is about 1 dB away from the information capacity limit. (For the definition of OSNR, please refer to [4].) For BI codes, the number of decoder iterations is set to 25, while the number of global (APP demapper-LDPC decoders') iterations to three, resulting in a total of 75 iterations. On the other hand, for NB code, the number of decoder iterations is set to 75, while the number of global iterations is zero. (Particularly, for NB codes, the global iterations result in negligible improvement in ASE noise dominated scenario.) The number of iterations can be reduced two times for the same BER performance, by using the layer decoding, as discussed in [29].

QUASI-CYCLIC LDPC CODES AND THEIR RELATED FAMILIES

Let us analyze several attractive classes of LDPC codes that are suitable for multidimensional optical communications, which are quasi-cyclic (QC) [21], [27], convolutional [29], staircase [30], and spatially coupled [31] LDPC codes. As we will show below, the various classes of these LDPC codes can be derived from QC-LDPC code design, but allowing for greater flexibility in avoiding the cycles of short length as compared to QC-LDPC code design itself. On the other hand, similar flexibility can be obtained from irregular QC-LDPC codes in which irregular structure can be precisely controlled. One such irregular QC-LDPC code design, based on combinatorial objects known as PBDs, is described in [12].

The parity-check matrix of QC-LDPC codes can be represented by [21]

$$H=\begin{bmatrix} I & I & I & \cdots & I \\ I & P^{S[1]} & P^{S[2]} & \cdots & P^{S[c-1]} \\ I & P^{2S[1]} & P^{2S[2]} & \cdots & P^{2S[c-1]} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ I & P^{(r-1)S[1]} & P^{(r-1)S[2]} & \cdots & P^{(r-1)S[c-1]} \end{bmatrix}, \quad (5)$$

where I is $B\times B$ (B is a prime number) identity matrix, P is $B\times B$ permutation matrix given by $P=(P_{ij})_{B\times B}$, $p_{i,i+1}=p_{B,1}=1$ (zero otherwise), and where r and c represent the number of block-rows and block-columns, respectively. The powers of permutation matrices denoted as $S[i]; i=0,1,\ldots,c-1$; should be chosen appropriately so that cycles of short length (in particular four and six) are avoided. Starting from the BI LDPC parity-check check matrix, the NB parity-check check matrix can either be obtained by randomly assigning higher (GF) nonzeros elements to "1" positions in BI check matrix or in a structured fashion as described below.

Let the template H-matrix of (3, 5) regular QC code be given as

$$H=\begin{bmatrix} I & I & I & I & I \\ I & P^{S[1]} & P^{S[2]} & P^{S[3]} & P^{S[4]} \\ I & P^{2S[1]} & P^{2S[2]} & P^{2S[3]} & P^{2S[4]} \end{bmatrix}.$$

By rearranging the H-matrix as

$$H=\begin{bmatrix} I & & & & & & & \\ P^{S[1]} & P^{S[2]} & & & & & & \\ P^{2S[1]} & P^{2S[2]} & P^{2S[3]} & & & & & \\ & I & I & I & & & & \\ & & P^{S[3]} & P^{S[1]} & P^{S[2]} & & & \\ & & & P^{2S[1]} & P^{2S[2]} & P^{2S[3]} & & \\ & & & & I & I & & \\ & & & & & P^{S[3]} & & \\ & & & & & & \ddots & \end{bmatrix},$$

we will obtain the convolutional LDPC code of period four. The corresponding spatially coupled LDPC code is now given as

$$H=\begin{bmatrix} I & I & & \\ P^{S[1]} & P^{S[2]} & I & I \\ P^{2S[1]} & P^{2S[2]} & P^{S[3]} & P^{S[4]} \\ & & P^{2S[3]} & P^{2S[4]} \end{bmatrix}.$$

Accordingly, multiple spatially coupled LDPC codes can be designed from the same template-QC-LDPC code. The staircase-LDPC codes can be designed in similar fashion, which produces

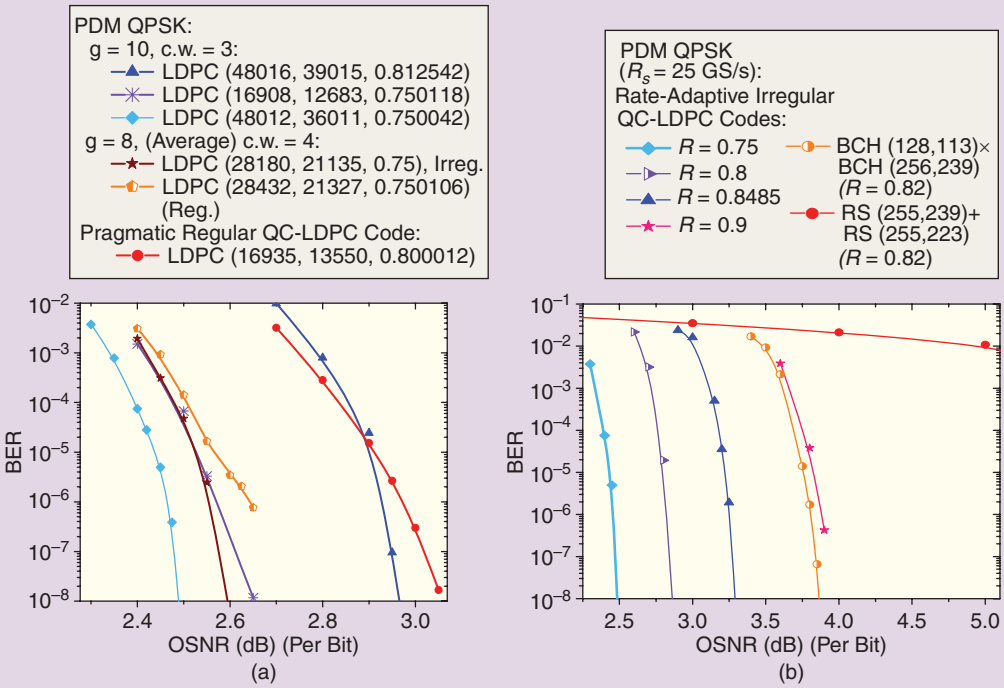
$$H=\begin{bmatrix} I & I & I & I & & & \\ & I & P^{S[1]} & P^{S[2]} & P^{S[3]} & & \\ & & & I & P^{2S[1]} & P^{2S[2]} & P^{2S[3]} \end{bmatrix}.$$

Therefore, convolutional, spatially coupled and staircaselike LDPC codes are just different arrangements derived from the QC-LDPC codes. However, the powers of permutation matrices should be properly chosen so that cycles of short lengths are avoided. These powers could be different for various classes of codes discussed above. For instance, the algorithm from [21] can be used to avoid the cycles of length $2k\,(k=2,3,\ldots)$. We add an integer at a time from the set $\{0,1,\ldots,B-1\}$ (not used before) to the initial set S and check if the equation,

$$S[i_1]j_1+S[i_2]j_2+\cdots+S[i_k]j_k+S[i_1]j_2+S[i_2]j_3+\cdots +S[i_k]j_1 \bmod B=0, \quad (6)$$

describing the cycle of length $2k$, is satisfied. If so, we will remove that integer from the set S and continue our search with another integer from the set $\{0,1,\ldots,B-1\}$ until we exploit all elements from $\{0,1,\ldots,B-1\}$. The code rate R of QC-code design is lower-bounded by $R\geq 1-r/|S|$ and the codeword length is $|S|B$, where $|S|$ denotes the cardinality of set S.

Let us now describe the irregular QC-LDPC code design derived from PBDs that we introduced in [12]. We assume that the set V of size υ represents elements, called points. Let any subset of points be called a block. Then, a pairwise balanced design PBD (υ, K, δ) is defined as a collection of blocks of different sizes taken from the set K, in such way that every pair of points is contained in δ of the blocks. As an illustration, the PBD(6, {2, 3}, 1) is given as collection of blocks {{0, 1, 3}, {1, 2, 4}, {2,3,5}, {0, 2}, {1, 5}, {3, 4}}. The parity-check matrix of irregular QC-LDPC, derived from PBDs, is given by (7), shown in the box at the bottom of the page, where I is $B\times B$ identity matrix and P is $B\times B$ permutation matrix defined above. In (7), $\{b_{ij}\}$ are the points of the ith block in PBD$(r, K, 1)$, in which the



[FIG6] The BER performance of (a) BI irregular QC-LDPC codes and (b) BI irregular rate adaptive length-48,000 LDPC code.

largest size of block k in set of sizes K satisfies the inequality $k \le r$. Finally, the $\mathcal{I}(b_{ij})$ denotes the indicator function, which has the value 1 for the existing point within the ith block, and 0 for the nonexisting point. Therefore, only those submatrices for which indicator function is 1 will be preserved from the template, which is a regular QC-LDPC code design. Given the fact that PBDs have regular mathematical structure that can be algebraically described, the irregular QC-LDPC codes derived from PBDs have the complexity comparable or lower to that of the regular QC-LDPC code design. The irregular QC-LDPC code derived from PBD(6, {2, 3},1) has the following form:

$$H = \begin{bmatrix} I & 0 & 0 & I & 0 & 0 \\ I & P^{S[1]} & 0 & 0 & P^{S[4]} & 0 \\ 0 & P^{2S[1]} & P^{2S[2]} & P^{2S[3]} & 0 & 0 \\ I & 0 & P^{3S[2]} & 0 & 0 & P^{3S[5]} \\ 0 & P^{4S[1]} & 0 & 0 & 0 & P^{S[5]} \\ 0 & 0 & P^{5S[2]} & 0 & P^{5S[4]} & 0 \end{bmatrix}.$$

Since both the identity matrix and the power of permutation matrix have a single one per row, the block size of the ith block from PBD determines ith block-column weight. In the previous example, the first three block-columns have column-weight three, while the last three have the column-weight two.

Equation (7) (shown in the box at the bottom of the page), which defines the PBD-based irregular QC-LDPC code design, can easily be generalized to NB LDPC codes and becomes (8) (shown in a box on the top of the next page), where α^i are nonzero elements of GF(q), while the other parameters have the same meaning as stated above. For the same example used above, i.e., for PBD(6, {2, 3}, 1), the corresponding parity-check matrix will be given by

$$H = \begin{bmatrix} \alpha^0 I & 0 & 0 & \alpha^3 I & 0 & 0 \\ \alpha^5 I & \alpha^0 P^{S[1]} & 0 & 0 & \alpha^3 P^{S[4]} & 0 \\ 0 & \alpha^5 P^{2S[1]} & \alpha^0 P^{2S[2]} & \alpha^1 P^{2S[3]} & 0 & 0 \\ I & 0 & \alpha^2 P^{3S[2]} & 0 & 0 & \alpha^2 P^{3S[5]} \\ 0 & P^{4S[1]} & 0 & 0 & 0 & \alpha\, P^{S[5]} \\ 0 & 0 & P^{5S[2]} & 0 & \alpha^2 P^{5S[4]} & 0 \end{bmatrix}.$$

The results of simulations of the QC-LDPC codes are shown in Figure 6, for polarization-division multiplexed QPSK of aggregate rate of at least 100 Gb/s and 50 iterations in min-sum-plus-correction-term algorithm. In Figure 6(a), a comparison of irregular QC-LDPC codes, derived from PBDs, is evaluated against the corresponding regular QC-LDPC codes. The girth-8 regular QC-LDPC(28432,21327) code of column-weight four exhibits the error floor phenomenon. On the other hand, the corresponding irregular

$$H = \begin{bmatrix} \mathcal{I}(b_{00})I & \mathcal{I}(b_{10})I & \mathcal{I}(b_{20})I & \dots & \mathcal{I}(b_{c-1,0})I \\ \mathcal{I}(b_{01})I & \mathcal{I}(b_{11})P^{S[1]} & \mathcal{I}(b_{21})P^{S[2]} & \dots & \mathcal{I}(b_{c-1,1})P^{S[c-1]} \\ \mathcal{I}(b_{02})I & \mathcal{I}(b_{12})P^{2S[1]} & \mathcal{I}(b_{22})P^{2S[2]} & \dots & \mathcal{I}(b_{c-1,2})P^{2S[c-1]} \\ \dots & \dots & \dots & \dots & \dots \\ \mathcal{I}(b_{0,r-1})I & \mathcal{I}(b_{1,r-1})P^{(r-1)S[1]} & \mathcal{I}(b_{2,r-1})P^{(r-1)S[2]} & \dots & \mathcal{I}(b_{c-1,r-1})P^{(r-1)S[c-1]} \end{bmatrix}, \quad (7)$$

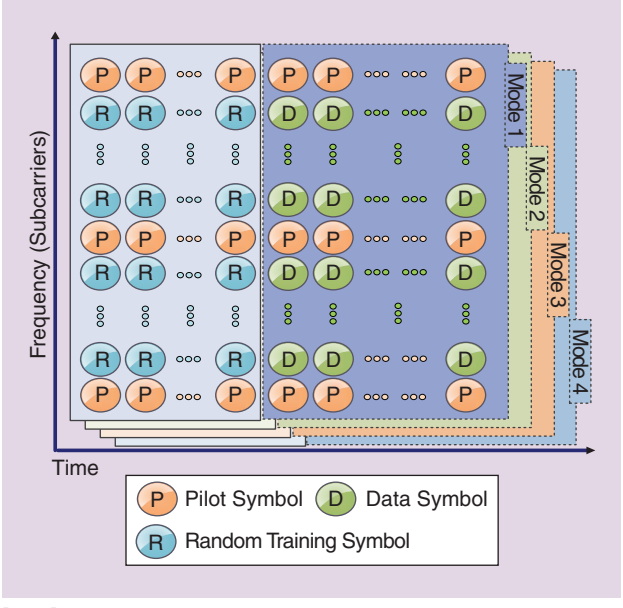
$$H = \begin{bmatrix} \alpha^0 \mathcal{I}(b_{00})I & \alpha^1 \mathcal{I}(b_{10})I & \cdots & \alpha^{c-1} \mathcal{I}(b_{c-1,0})I \\ \alpha^{c-1} \mathcal{I}(b_{01})I & \alpha^0 \mathcal{I}(b_{11}) P^{S[1]} & \cdots & \alpha^{c-2} \mathcal{I}(b_{c-1,1}) P^{S[c-1]} \\ \alpha^{c-2} \mathcal{I}(b_{02})I & \alpha^{c-1} \mathcal{I}(b_{12}) P^{2S[1]} & \cdots & \alpha^{c-3} \mathcal{I}(b_{c-1,2}) P^{2S[c-1]} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{c-(r-1)} \mathcal{I}(b_{0,r-1})I & \alpha^{c-r+2} \mathcal{I}(b_{1,r-1}) P^{(r-1)S[1]} & \cdots & \alpha^{c-r} \mathcal{I}(b_{c-1,r-1}) P^{(r-1)S[c-1]} \end{bmatrix}, \quad (8)$$

QC-LDPC(28180, 21135) does not exhibit such a floor while providing the net coding gain (NCG) of 9.4 dB at BER of 10^{-8} and expected NCG (obtained by extrapolation) of 12.4 dB at BER of 10^{-15}. Accordingly, this irregular LDPC code performs only 0.09 dB worse than the longer girth-10 regular QC-LDPC(48012,36011) code, while outperforming regular girth-10 LDPC(16908,12683) code (of column-weight-3). Moreover, this irregular QC-LDPC code outperforms pragmatic QC-LDPC(16935,13550) code [21] by 0.46 dB. In Figure 6(b), the proposed QC-LDPC code design is used for rate-adaptation, for code rates 0.75–0.9. The QC-LDPC code design, given by (7), of length 48,000 is used. It is evident that QC-LDPC codes of rates ranging from 0.75–0.85 outperform (soft-decision) turbo-product code of rate 0.82, and significantly outperform concatenated RS code of rate 0.82.

THE LS ESTIMATION IN OPTICAL MIMO-OFDM SYSTEMS APPEARS SUFFICIENT FOR THE SHORT-REACH AND MEDIUM-REACH APPLICATIONS, WHILE THE USE OF LMMSE CHANNEL ESTIMATION IS NECESSARY FOR THE LONG-HAUL APPLICATIONS.

COHERENT OPTICAL OFDM OVER FMFs IN THE PRESENCE OF NONLINEAR EFFECTS

Let us now demonstrate the possibility of long-haul transmission with the aggregate bit rate exceeding 1 Tb/s by using the mode-multiplexed NB LDPC-coded OFDM over FMFs, as initially suggested in [32]. This scheme represents one particular instance of a hybrid CM scheme described in the section "Hybrid Coded-Modulation Employing Electrical and Optical Basis Functions." Both the impacts of nonlinear effects and nonlinear interaction between spatial modes have been included through the nonlinear Schrödinger equation model [14], this time generalized for the FMF case. To analyze the feasibility of the this system, we study the propagation of the mode-multiplexed OFDM signals over targeted length of 3,000 km of FMF link by applying the Monte Carlo simulation. Two different ways of compensation of degradations in FMF have been evaluated, mainly the least square (LS) and the least minimum mean-square error (LMMSE) estimation schemes [14]. Both of them are used in combination with the channel interpolation based on piece-wise linear interpolation. The complexity of LMMSE algorithm is higher than that of the LS one since it requires the use of training sequences to estimate the correlation matrices and OSNR. In addition, the matrix inversion operation is required. The LMMSE OFDM arrangement is shown in Figure 7. On the other hand, the improvement brought by the LMMSE scheme is substantial. Accordingly, the LS estimation in optical MIMO-OFDM systems appears sufficient for the short-reach and medium-reach applications, while the use of LMMSE channel estimation is necessary for the long-haul applications.



[FIG7] **The OFDM symbol arrangement.**

The channel estimation can be performed by inserting the pilot tones into all subcarriers of OFDM frames with a specific insertion period [14]. The estimate of pilot signals, based on the least-square (LS) technique, is given by

$$\hat{H}_{p,\mathrm{LS}} = [H_{p,\mathrm{LS}}(0), H_{p,\mathrm{LS}}(1), \ldots, H_{p,\mathrm{LS}}(N_p - 1)]^T = \left[\frac{Y_p(0)}{X_p(0)}, \frac{Y_p(1)}{X_p(1)}, \ldots, \frac{Y_p(N_p-1)}{X_p(N_p-1)}\right]^T, \quad (9)$$

where $X_p(i)$ and $Y_p(i)$ are transmitted and received symbols of pilot i $(i = 0, 1, \ldots, N_p - 1)$. To improve the accuracy of pilot estimation, the LMMSE estimator is a preferable option. The representation of pilot subcarriers for the MMSE estimator is given by

$$\hat{H}_{p,\mathrm{LMMSE}} = R_{H_p, H_p}\left(R_{H_p, H_p} + \frac{\xi}{\mathrm{SNR}_p} I\right)^{-1} \hat{H}_{p,\mathrm{LS}}, \quad (10)$$

where $\hat{H}_{p,\mathrm{LS}}$ is the least-square estimate of as shown in (9), $\mathrm{SNR} = \mathrm{E}\{|X_p(k)|^2\}/\sigma_n^2$ is the signal-to-noise ratio for each mode estimated after LS compensation, and $\xi = \mathrm{E}\{|X_p(k)|\}^2 \mathrm{E}\{|(1)/(X_p(k))|^2\}$ is a constant dependent on the signal constellation, while the covariance matrix is defined as $R_{H_p,H_p} = E\{H_p H_p^{\dagger}\}$. (E denotes the expectation operator and † is Hermitian transposition operation.) By transmitting sufficiently long

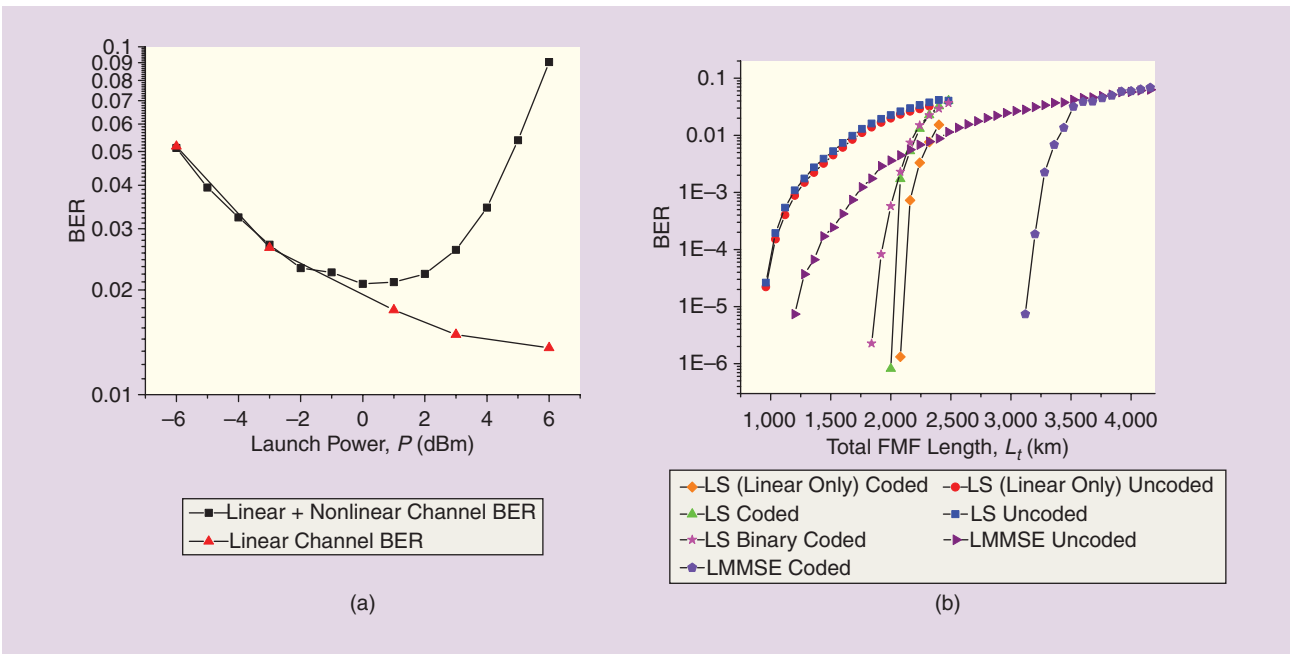
training OFDM frame sequences before data OFDM frames, we can estimate the covariance matrix. The channel transfer function of data subcarriers is interpolated based on the LS or LMMSE estimate of the pilot signals as shown in Figure 7. The preamble of OFDM symbols, containing both the pilot and random training symbols, is used to estimate all relevant parameters (SNR, cross-correlation function, and chromatic dispersion effects). The data OFDM frame contains pilot subcarriers needed for an estimation of the mode coupling. Notice that chromatic dispersion compensation operation as a slow-varying channel impairment can be completely moved to the transmitter side to reduce the overall complexity at receiver side. In the pure linear channel model, the ASE noise is the only limiting factor in addition to the mode-coupling. However, when the impact of nonlinear effects is included, an optimum launch power will be recognized, as illustrated in Figure 8(a). The optimum launch power, which is around 0.5 dBm, presents a tradeoff between (ASE) noise and fiber nonlinearity impacts. Additional details of this scheme, as well as OFDM system parameters, can be found in [32], but in that scheme only linear effects have been considered. On the other hand, in this article we evaluate the system performance in the presence of nonlinear effects as well.

Figure 8(b) shows the simulation results corresponding to LS and LMMSE estimation methods for different FMF lengths. As we can see, there is about 1,000-km gap in reach between LS and LMMSE estimation. Therefore, it is possible to achieve a transmission distance of 3,000 km of FMF (at BER of 10^{-5}) with bit rates exceeding 1 Tb/s when four OFDM bands are used. By using NB LDPC coding, we can extend the transmission distance for an additional 200 km. We believe that it is possible to extend the transmission distance beyond 3,000 km by using the DFT-spread-based approaches, having better tolerance to fiber nonlinearities.

A MULTIDIMENSIONAL APPROACH OPENS UP THE NEW POSSIBILITIES FOR DYNAMIC AND ELASTIC NETWORKING BY USING BOTH SPECTRAL AND SPATIAL ENTITIES AS TOOLS.

CONCLUSIONS

We have proposed and analyzed the key aspects of the hybrid multidimensional coded-modulation that employs both electrical and optical degrees of freedom to simultaneously address the key constraints in next-generation optical transmission systems and networks with bit rates exceeding 1 Tb/s per optical channel. In the proposed scheme, the orthogonal prolate spheroidal wave functions have been used as the electrical basis functions, while polarization and spatial modes have been used as the optical basis functions. In addition, we analyze the performance of quasi-cyclic LDPC codes that are best suited for multidimensional optical communications and have shown their comparative advantages. This scheme represents a generalization of conventional WDM-space-division multiplexed superchannel transmission scheme. Many of existing schemes proposed for multi-Tb/s serial transport so far, including 4-D SMF-based transport, space-division multiplexing, WDM superchannel optical transport, WDM-space-division multiplexed superchannel transmission, Nyquist-WDM, just to mention few, are just particular instances of the proposed hybrid coded-modulation scheme. We have also illustrated the cases where the proposed hybrid scheme can serve as an enabler for ultra-high speed bit rates exceeding 10 Tb/s. Finally, we have outlined that a multidimensional approach opens up the new possibilities for dynamic



[FIG8] **(a) The BER versus launch power (after 2,000 km of FMF). (b) The LDPC-coded OFDM BER performance versus total transmission distance.**

and elastic networking by using both spectral and spatial entities as tools.

AUTHORS

Ivan B. Djordjevic (ivan@email.arizona.edu) is a tenured associate professor in the Department of Electrical and Computer Engineering (College of Engineering) of the University of Arizona, with a joint appointment in the College of Optical Sciences. Prior to joining the University of Arizona, he was with the University of Bristol and University of the West of England, Bristol, United Kingdom; Tyco Telecommunications, Eatontown, United States; National Technical University of Athens, Greece; and State Telecommunication Company, Nis, Serbia. In 2013, he was also with TU Darmstadt, Germany. He is an author/coauthor of four books, over 300 international journal/conference publications, and 20 U.S. patents. He is an associate editor for three journals.

Milorad Cvijetic (milorad@optics.arizona.edu) is a professor in the College of Optical Sciences at the University of Arizona. For more than 30 years, he has been directly involved in state-of-the-art R&D in several of the world's leading telecommunication laboratories. He has published numerous technical papers, authored/coauthored four books, and has been the author/coauthor of 12 U.S. patents, all related to optical communications and networking.

Changyu Lin (lcy@email.arizona.edu) is a Ph.D. student in the Department of Electrical and Computer Engineering of the University of Arizona. He is an author of several journal papers and conference papers. His research interests include OFDM and MIMO signal processing for multi-Tb/s optical transport networks and nonbinary LDPC coded modulation for optical communications.

REFERENCES

[1] P. Winzer, "Beyond 100G Ethernet," *IEEE Commun. Mag.*, vol. 48,no. 7, pp. 26–30, July 2010.

[2] S. Gringery, B. Basch, and T. J. Xia, "Technical considerations for supporting bit rates beyond 100 Gb/s," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 21–30, Feb. 2012.

[3] H. G. Weber, S. Ferber, M. Kroh, C. Schmidt-Langhorst, R. Ludwig, V. Marembert, C. Boerner, F. Futami, S. Watanabe, and C. Schubert,, "Single channel 1.28 Tbit/s and 2.56 Tbit/s DQPSK transmission," *Electron. Lett.*, vol. 42, no. 3, pp. 178–179, Feb. 2006.

[4] I. B. Djordjevic, "Spatial-domain-based hybrid multidimensional coded-modulation schemes enabling multi-Tb/s optical transport," *J. Lightw. Technol.*, vol. 30, no. 14, pp. 2315–2328, 2012.

[5] Y. Ma, Q. Yang, Y. Tang, S. Chen, and W. Shieh, "1-Tb/s single-channel coherent optical OFDM transmission over 600-km SSMF fiber with subwavelength bandwidth access," *Opt. Exp.*, vol. 17, no. 11, pp. 9421–9427, May 2009.

[6] S. Chandrasekhar, X. Liu., B. Zhu, and D. W. Peckham, "Transmission of a 1.2-Tb/s 24-carrier no-guard-interval coherent OFDM superchannel over 7200-km of ultra-large-area fiber," in *Proc. ECOC'09*, paper PD2.6, 2009.

[7] X. Liu, S. Chandrasekhar, A. H. Gnauck, P. J. Winzer, S. Randel, S. Corteselli, A. R. Chraplyvy, R. W. Tkach, B. Zhu, T. F. Taunay, and M. Fishteyn, "Digital coherent superposition for performance improvement of spatially multiplexed coherent optical OFDM superchannels," *Opt. Exp.*, vol. 20, no. 26, pp. B595-B600, Dec. 2012.

[8] I. B. Djordjevic, T. Liu, and T. Wang, "On the adaptive software-defined LDPC-coded multidimensional spatial-MIMO multiband generalized OFDM enabling beyond 10 Tb/s optical transport," *IEEE Photonics J.*, vol. 5, no. 1, Paper 7200207, Feb. 2013.

[9] R. Ryf, S. Randel, N. K. Fontaine, M. Montoliu, E. Burrows, S. Chandrasekhar, A. H. Gnauck, C. Xie, R. Essiambre, P. Winzer, R. Delbue, P. Pupalaikis, A. Sureka, Y. Sun, L. Gruner-Nielsen, R. V. Jensen, and R. Lingle, "32-bit/s/Hz spectral efficiency WDM transmission over 177-km few-mode fiber," in *Proc. OFC/NFOEC 2013*, Paper PDP5A.1.

[10] E. Ip, M.-J. Li, Y. Huang, A. Tanaka, E. Mateo, W. Wood, J. Hu, Y. Yano, and K. Koreshkov, "146 $\lambda \times 6 \times$ 19-Gbaud wavelength and mode-division multiplexed transmission over 10 × 50-km spans of few-mode fiber with a gain-equalized few-mode EDFA," in *Proc. OFC/NFOEC 2013*, Paper PDP5A.2.

[11] X. Chen, A. Li, J. Ye, A. Amin, and W. Shieh, "Reception of mode-division multiplexed superchannel via few-mode compatible optical add/drop multiplexer," *Opt. Exp.*, vol. 20, no. 13, pp. 14302–14307, June 2012.

[12] I. B. Djordjevic, "On the irregular nonbinary QC-LDPC-coded hybrid multidimensional OSCD-modulation enabling beyond 100 Tb/s optical transport," *J. Lightw. Technol.*, vol. 31, no. 16, pp. 2969–2975, Aug. 2013.

[13] D. Qian, E. Ip, M. Huang, M. Li, A. Dogariu, S. Zhang, Y. Shao, Y. Huang, Y. Zhang, X. Cheng, Y. Tian, P. Ji, A. Collier, Y. Geng, J. Linares, C. Montero, V. Moreno, X. Prieto, and T. Wang, "1.05 Pb/s transmission with 109 b/s/Hz spectral efficiency using hybrid single and few-mode cores," in *Proc. 2012 Frontiers Optics/Laser Science. XXVIII (FiO/LS)*, Paper FW6C.3, Rochester, NY.

[14] M. Cvijetic and I. B. Djordjevic, *Advanced Optical Communication Systems and Networks*. Boston, MA; London, England: Artech House, 2013.

[15] T. J. Xia, G. Wellbrock, A. Tanaka, M. Huang, E. Ip, D. Qian, Y. Huang, S. Zhang, Y. Zhang, P. Ji, Y. Aono, S. Murakami, and T. Tajima, "High capacity field trials of 40.5 Tb/s for LH distance of 1,822 km and 54.2 Tb/s for regional distance of 634 km," in *Proc. OFC/NFOEC 2013*, Paper PDP5A.4, 2013.

[16] R. Schmogrow, M. Winter, M. Meyer, D. Hillerkuss, S. Wolf, B. Baeuerle, A. Ludwig, B. Nebendahl, S. Ben-Ezra, J. Meyer, M. Dreschmann, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold "Real-time Nyquist pulse generation beyond 100 Gbit/s and its relation to OFDM," *Opt. Exp.*, vol. 20, no. 1, pp. 317–337, Jan. 2012.

[17] J. Yu and X. Zhou, "Ultra-high-capacity DWDM transmission system for 100G and beyond," *IEEE Comm. Mag.*, vol. 8, pp. 57–64, Mar. 2010.

[18] H. Huang, G. Xie, Y. Yan, N. Ahmed, Y. Ren, Y. Yue, D. Rogawski, M. Tur, B. Erkmen, K. Birnbaum, S. Dolinar, M. Lavery, M. Padgett, and A. E. Willner, "100 Tbit/s free-space data link using orbital angular momentum mode division multiplexing combined with wavelength division multiplexing," in *Opt. Fiber Communication Conf./Nat. Fiber Optic Engineers Conf. 2013*, paper OTh4G.5.

[19] T. Morioka, Y. Awaj, R. Ryf, P. Winzer, D. Richardson,and F. Poletti,, "Enhancing optical communications with brand new fibers," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 40–50, Feb. 2012.

[20] M. Cvijetic, I. B. Djordjevic, and N. Cvijetic, "Dynamic multidimensional optical networking based on spatial and spectral processing," *Opt. Exp.*, vol. 20, no. 1, pp. 317–337, 2012.

[21] I. B. Djordjevic, M. Arabaci, and L. Minkov, "Next generation FEC for high-capacity communication in optical transport networks," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3518–3530, Aug. 2009.

[22] H. Buelow, T. Rahman, F. Buchali, W. Idler, and W. Kuebart, "Transmission of 4-D modulation formats at 28 Gbaud," in *Opt. Fiber Communication Conf./Nat. Fiber Optic Engineers Conf. 2013*, paper JW2A.39.

[23] T. Liu and I. B. Djordjevic, "On the optimum signal constellation design for high-speed optical transport networks," *Opt. Exp.*, vol. 20, no. 18, pp. 20396–20406, Aug. 2012.

[24] Y. Zhang, M. Arabaci, and I. B. Djordjevic,, "Evaluation of four-dimensional nonbinary LDPC-coded modulation for next-generation long-haul optical transport networks," *Opt. Exp.*, vol. 20, no. 8, pp. 9296–9301, Sept. 2012.

[25] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty V: The discrete case," *Bell Syst. Tech. J.*, vol. 57, pp. 1371–1430, June 1978.

[26] D. Hillerkuss, R. Schmogrow, T. Schellinger, M. Jordan, M. Winter, G. Huber, T. Vallaitis, R. Bonk, P. Kleinow, F. Frey, M. Roeger, S. Koenig, A. Ludwig, A. Marculescu, J. Li, M. Hoh, M. Dreschmann, J. Meyer, S. Ben Ezra, N. Narkiss, B. Nebendahl, F. Parmigiani, P. Petropoulos, B. Resan, A. Oehler, K. Weingarten, T. Ellermeyer, J. Lutz, M. Moeller, M. Huebner, J. Becker, C. Koos, W. Freude, and J. Leuthold, "26 Tbit/ s line-rate super-channel transmission utilizing all-optical fast Fourier transform processing," *Nat. Photonics*, vol. 5, pp. 364–371, May 2011.

[27] M. Arabaci, I. B. Djordjevic, R. Saunders, and R. M. Marcoccia, "Polarization-multiplexed rate-adaptive non-binary LDPC-coded multilevel modulation with coherent detection for optical transport networks," *Optic Engineers*, vol. 18, no. 3, pp. 1820–1832, 2010.

[28] Y. Zhao, J. Qi, F. N. Hauske, C. Xie, D. Pflueger, and G. Bauch, "Beyond 100G optical channel noise modeling for optimized soft-decision FEC performance," in *Proc. Optical Fiber Communication Conf. Exposition (OFC/NFOEC), 2012* and *Nat. Fiber Optic Engineers Conf.*, Paper OW1H.3, pp. 4–8, Mar. 2012.

[29] D. Chang, F. Yu, Z. Xiao, N. Stojanovic, F. N. Hauske, Y. Cai, C. Xie, L. Li, X. Xu, and Q. Xiong, "LDPC convolutional codes using layered decoding algorithm for high speed coherent optical transmission," in *Proc. OFC/NFOEC 2012*, Paper OW1H.4, 2012.

[30] B. P. Smith, A. Farhood, A. Hunt, and F. R. Kschischang, "Staircase codes: FEC for 100 Gb/s OTN," *J. Lightw. Technol.*, vol. 30, no. 1, pp. 110–117, Jan. 2012.

[31] K. Sugihara, Y. Miyata, T. Sugihara, K. Kubo, H. Yoshida, W. Matsumoto, and T. Mizuochi, "A spatially coupled type LDPC code with an NCG of 12 dB for optical transmission beyond 100 Gb/s," in *Proc. OFC/NFOEC 2013*, Paper OM2B.4, 2013.

[32] C. Lin, I. B. Djordjevic, D. Zou, M. Arabaci, and M. Cvijetic, "Non-binary LDPC coded mode-multiplexed coherent optical OFDM 1.28 Tbit/s 16-QAM signal transmission over 2000 km of few-mode fibers with mode dependent loss," *IEEE Photonics J.*, vol. 4, no. 5, pp. 1922–1929, 2012.

[SP]

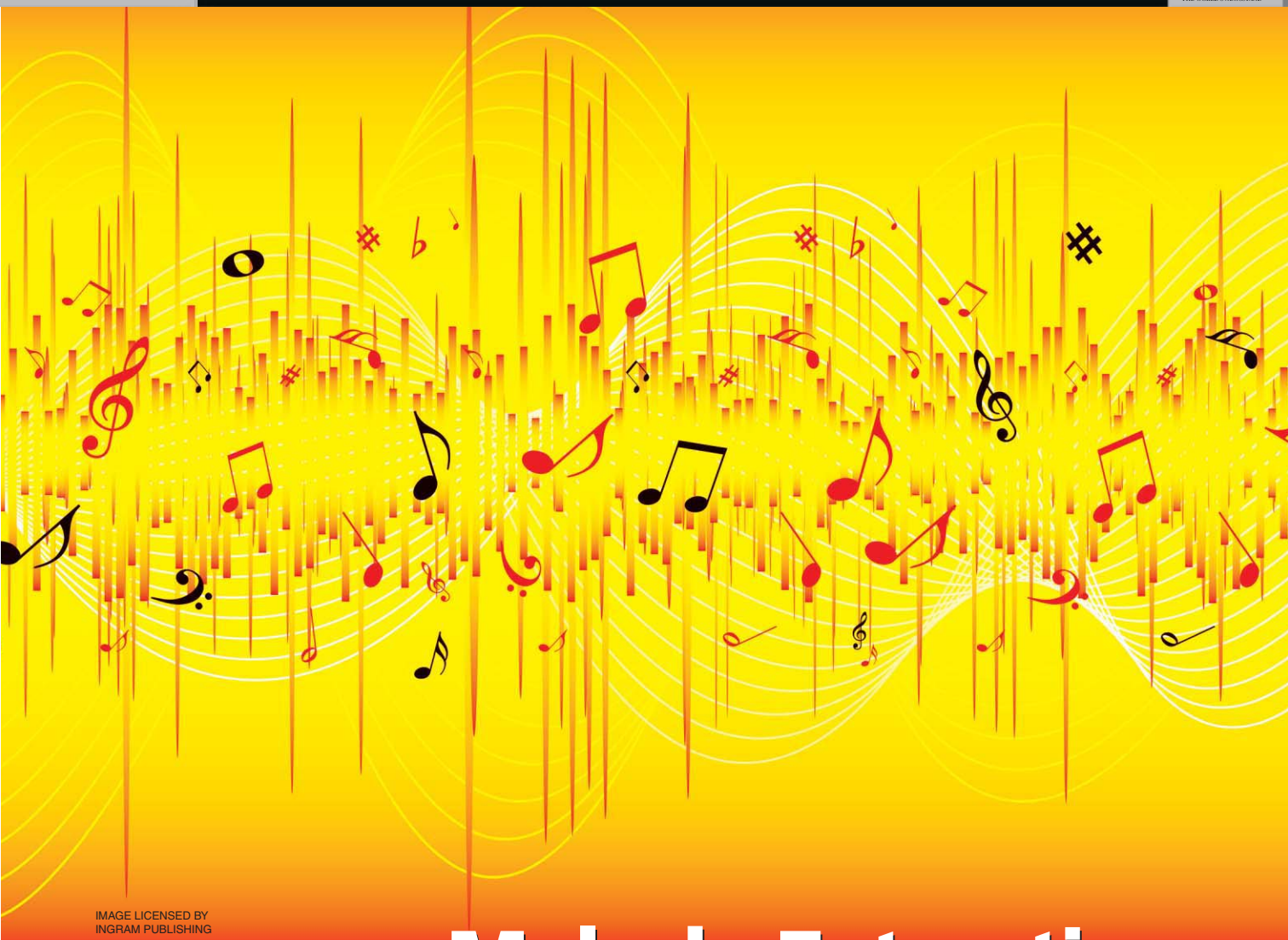


IMAGE LICENSED BY INGRAM PUBLISHING

Melody Extraction from Polyphonic Music Signals

[Approaches, applications, and challenges]

[Justin Salamon, Emilia Gómez, Daniel P.W. Ellis, and Gaël Richard]

Melody extraction algorithms aim to produce a sequence of frequency values corresponding to the pitch of the dominant melody from a musical recording. Over the past decade, melody extraction has emerged as an active research topic, comprising a large variety of proposed algorithms spanning a wide range of techniques. This article provides an overview of these techniques, the applications for which melody extraction is useful, and the challenges that remain. We start with a discussion of "melody" from both musical and signal processing perspectives and provide a case

Digital Object Identifier 10.1109/MSP.2013.2271648
Date of publication: 12 February 2014

1053-5888/14/$31.00©2014IEEE

study that interprets the output of a melody extraction algorithm for specific excerpts. We then provide a comprehensive comparative analysis of melody extraction algorithms based on the results of an international evaluation campaign. We discuss issues of algorithm design, evaluation, and applications that build upon melody extraction. Finally, we discuss some of the remaining challenges in melody extraction research in terms of algorithmic performance, development, and evaluation methodology.

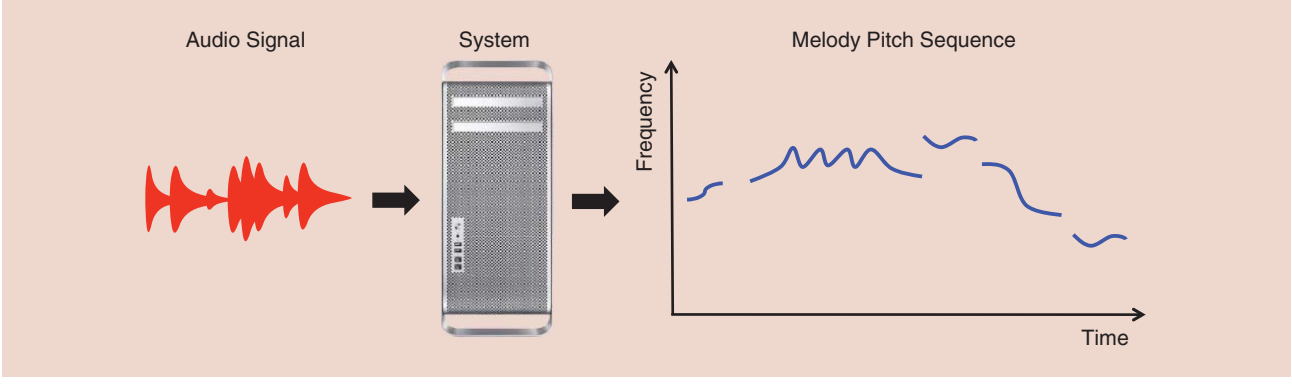
INTRODUCTION

Music was the first mass-market industry to be completely restructured by digital technology starting with the compact disc and leading to today's situation where typical consumers may have access to thousands of tracks stored locally on their smartphone or music player, and millions of tracks instantly available through cloud-based music services. This vast quantity of music demands novel methods of description, indexing, searching, and interaction. Recent advances in audio processing have led to technologies that can help users interact with music by directly analyzing the musical content of audio files. The extraction of melody from polyphonic music signals is such a technology and has received substantial attention from the audio signal processing and music information retrieval (MIR) research communities. Known as *melody extraction, audio melody extraction, predominant melody extraction, predominant melody estimation,* or *predominant fundamental frequency estimation*, the task involves automatically obtaining a sequence of frequency values representing the pitch of the dominant melodic line from recorded music audio signals (Figure 1).

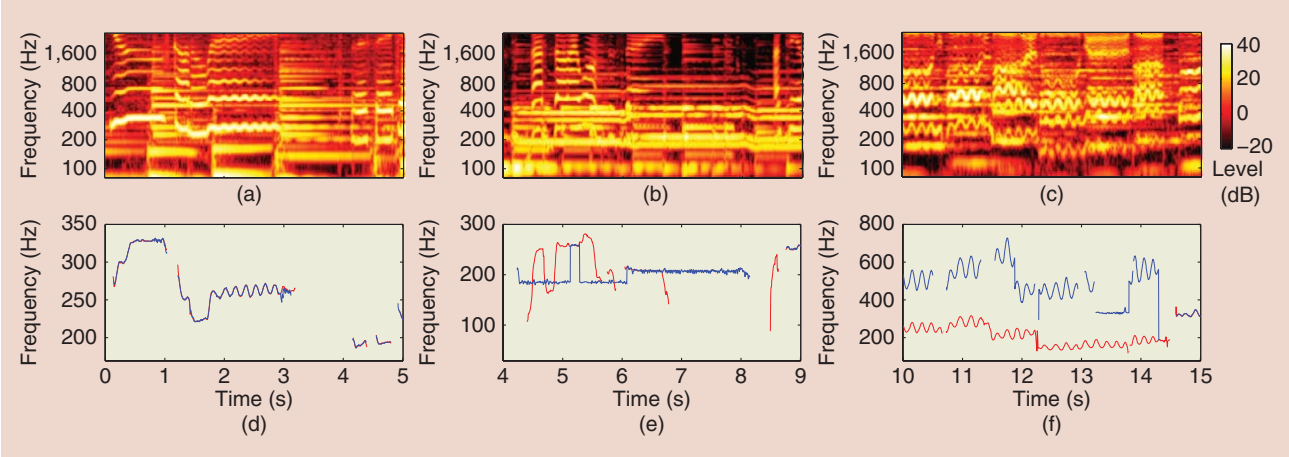
OVER THE PAST DECADE, MELODY EXTRACTION HAS EMERGED AS AN ACTIVE RESEARCH TOPIC, COMPRISING A LARGE VARIETY OF PROPOSED ALGORITHMS SPANNING A WIDE RANGE OF TECHNIQUES.

Music transcription, i.e., converting an audio signal into a description of all the notes being played, is a task that can usually be achieved by a trained student of music and has long been a topic of computational research. It has, however, proven to be very difficult due to the complex and deliberately overlapped spectral structure of musical harmonies. In one of the earliest works in the field, Masataka Goto pointed out that many interesting music tasks, such as melody-based retrieval or melody line suppression for karaoke, could be achieved with a much more limited transcription that recovered only a single melody line as the "strongest" pitch in the likely melody range at any time [1]. This idea was picked up by Emilia Gómez, Beesuan Ong, and Sebastian Streich, who put together a melody extraction task as part of the Audio Description Contests associated with the 2004 International Conference on Music Information Retrieval (ISMIR), organized by the Music Technology Group at Pompeu Fabra University, Barcelona [2]. This activity was followed by the Music Information Retrieval Evaluation eXchange (MIREX) evaluation campaign for MIR technologies [3] and has in subsequent years resulted in a series of well-organized international evaluations with broad participation, described in the section "Algorithm Overview: 2005 to Date."

To frame the technical task of melody extraction, we should start by examining the musicological concept of "melody," which ultimately relies on the judgement of human listeners [2] and will therefore tend to vary across application contexts (e.g., symbolic melodic similarity [4] or music transcription [5]). Centuries of musicological study [6] have resulted in no clear consensus regarding the definition of "melody" but, faced with the need for a common interpretation, the MIR community has opted for simplified, pragmatic definitions that result in a task amenable to signal processing. One popular definition [2] holds that "the melody is the single (monophonic) pitch sequence that a listener might reproduce if asked to whistle or hum a piece of polyphonic music, and that a listener would recognize as being the essence of that music when heard in comparison." This definition is still open to a considerable degree of subjectivity, since different listeners might hum different parts after listening to the same song (e.g., lead vocals versus guitar solo). In practice, research has focused on what we



[FIG1] Melody extraction: obtaining a sequence of frequency values representing the pitch of the melody from the audio signal of polyphonic music.



[FIG2] **Case study examples: (a)–(c) show the log-frequency spectrogram of three excerpts in the genres of (a) vocal jazz, (b) pop, and (c) opera. Parts (d)–(f) show the extracted melody [16] (blue) and ground truth (red) for each excerpt, respectively.**

term *single source predominant fundamental frequency estimation*. That is, the melody is constrained to belong to a single sound source throughout the piece being analyzed, where this sound source is considered to be the most predominant instrument or voice in the mixture. While the subjective element can not be completely eliminated even in this definition (for instance, how do we define predominant?), the problem is avoided in practice by working with musical material that contains a clear lead singer or instrument. Thus, our modified task definition becomes "single source predominant fundamental frequency estimation from musical content with a lead voice or instrument." While this definition is too limited to encompass everything one might consider as melody, its solution would nonetheless lead to extremely powerful technologies. Note that we have used the term *fundamental frequency* (henceforth f_0) to refer to the physical property most closely related to the perceptual property of pitch [7]. Still, the terms *pitch* and f_0 are often used interchangeably in the melody extraction literature, and for the sake of readability we shall do the same here. The final musical term we must define is "polyphonic music." Although musicology draws distinctions between monophonic, homophonic, heterophonic, and polyphonic musical textures, in this article "polyphonic" is simply used to refer to any type of music in which two or more notes can sound simultaneously, be it on different instruments (e.g., voice, guitar, and bass) or a single instrument capable of playing more than one note at a time (e.g., the piano).

With these definitions of melody and polyphony, it becomes easier to define melody extraction as a signal processing challenge: given a recording of polyphonic music, we want to automatically estimate the sequence of f_0 values that corresponds to the pitch of the lead voice or instrument. Furthermore, we must estimate the time intervals when this voice is not present in the mixture (known as the "voicing detection" problem). For a human listener, this task might seem almost trivial—many of us can sing the melodies of our favorite songs even without any musical training. Those with musical training can even transcribe a melody into musical notation. However, when we try to automate this task, it turns out to be highly challenging. The complexity of the task is mainly due to two factors: first, a polyphonic music signal is composed of the superposition of the sound waves produced by all instruments in the recording, and much of the time these instruments play simultaneously. When considering the spectral content of the signal, the frequency components of different sources superimpose making it very hard to attribute specific energy levels in specific frequency bands to the notes of individual instruments. This is further complicated by mixing and mastering techniques which can add reverberation (thus blurring note onsets and offsets and increasing the overlap of sound sources) or apply dynamic range compression (thus reducing the difference between soft and loud sources, increasing interference). Second, even after we obtain a pitch-based representation of the audio signal, we still need to determine which pitch values belong to the predominant melody and which are merely accompaniment. The challenge is illustrated in Figure 2, which displays the spectrograms of three polyphonic excerpts (a)–(c) and the target melody sequence [(d)–(f), in red] together with the estimate (in blue) of a melody extraction algorithm (see the next section).

As we discuss in the section "Software and Applications," melody extraction has many potential applications, including query-by-humming (QBH) (searching for a song by singing or humming part of its melody) and cover song identification (detecting whether two recordings are different renditions of the same musical piece) [8], [9], genre classification (automatically sorting your music collection based on genre) [10], music desoloing for the automatic generation of karaoke accompaniment [11], and singer characterization [12]. It also has a wide range of applications in computational musicology and ethnomusicology, such as music transcription [13], intonation analysis [14], and automatic melodic motif and pattern analysis [15]. Determining the melody of a song could also be used as an intermediate step toward the derivation of other semantic labels from music signals. Finally, melody extraction also has a variety of uses outside

FOR MORE INFORMATION

- Adobe Audition: http://www.adobe.com/products/audition/html
- Melodyne: http://www.celemony.com/cms
- SMSTools: http://mtg.upf/edu/technologies/sms
- Wavesurfer: http://www.speech.kth.se/wavesurfer
- LabROSAmelodyextract2005: http://labrosa.ee.columbia.edu/projects/melody/
- FChT: http://iie.fing.edu/uy/investigacion/grupos/gpa/fcht.html
- separateLeadStereo: http://www.durrieu.ch/research/jstsp2010.html
- IMMF0salience: https://github.com/wslihgt.IMMF0salience
- Vamp audio analysis plug-in system: http://www.vamp-plugins.org
- MELODIA: http://mtg.upf.edu/technologies/melodia
- Melody Extraction for Music Games: http://www.idmt.fraunhofer.de/en/Service_Offerings/products_and_technologies/m_p/melody_extraction.html
- SoundHound: http://www.soundhound.com
- ADC2004 and MIREX05 data sets: http://labrosa.ee.columbia.edu/projects/melody/
- MIR-1K data set: https://sites.google.com/site/unvoicedsoundseparation/mir-1k
- RWC pop data set: http://staff.aist.go.jp/m/goto/RWC-MDB/
- Audio Melody Extraction Annotation Initiative: http://ameannotationinitiative.wikispaces.com

the realm of research, such as electroacoustic composition and music education. Melody extraction technologies are beginning to be incorporated into professional music production tools such as Adobe Audition and Melodyne (see "For More Information").

CASE STUDY

To better understand the challenges of melody extraction and the types of errors afflicting melody extraction algorithms, we start with a closer look at the actual melody extraction results for some musical excerpts. For conciseness, we limit ourselves to one state-of-the-art algorithm [16], but the types of errors we observe (and the challenges they represent) are common to all methods.

Figure 2 shows the output of the algorithm for three excerpts in the genres of vocal jazz [(d)], pop music [(e)], and opera [(f)]. In (a)–(c), we display a log-frequency spectrogram of each excerpt, showing the complex pattern of harmonics associated with these polyphonic musical signals. Plots (d)–(f) display the final melody line estimated by the algorithm (blue) overlaid on top of the ground truth annotation (red).

Before we can interpret different types of errors in the plots, it is useful to know what a correct extraction looks like, provided in Figure 2(d). We see that the blue (estimated) and red (ground truth) melody sequences overlap almost perfectly, and there are practically no frames where only one sequence is present. The perfect overlap means the pitch estimation of the algorithm is correct. The fact that there are no frames where only one sequence is present indicates we have not made any voicing detection mistakes—a red sequence on its own would mean we wrongly estimated the frame as unvoiced when the melody is actually present. A blue sequence on its own would mean a case of voicing false alarm, i.e., a frame where we mistakenly included some other pitched source in the melody when the melody is in fact not present in that frame. In (d), we see that the algorithm correctly estimates the pitch of the lead singer while excluding the notes of the piano chord played between seconds three and four.

In Figure 2(e), we provide an example that contains both pitch errors (seconds four to seven) and voicing errors (seconds seven to nine). The excerpt is taken from a pop song whose arrangement includes a lead singer, guitar accompaniment, and backing vocals. Here the source of both types of errors are the backing vocals, who sing a stable pitch in the same range as the melodic line of the lead singer. As a result, the algorithm mistakenly tracks the backing vocals, resulting in a wrong pitch estimate (up to the seventh second) followed by a voicing false alarm, since the backing vocals continue after the lead singer has paused.

Finally, in Figure 2(f), we provide an example where the algorithm makes octave errors. In this excerpt, taken from an opera aria sung by a male singer, the pitch class of the melody is correctly estimated but in the wrong octave (one octave above the actual pitch of the singer). Here the octave errors most likely stem from the actual singing technique used by the singer. Unlike pop or jazz singers, classical singers are trained to produce a highly resonant sound (allowing them to be heard over the orchestra). In the low frequencies this resonance results in the second harmonic often having a larger amplitude than the fundamental frequency, and in the high frequencies the appearance (especially in male singers) of a clear formant around 3 kHz (the "singer's formant") [17]. Combined, these phenomena can cause the algorithm to give more weight to $2f_0$ than to f_0 (f_0 being the correct fundamental frequency), as seen in the spectrogram in Figure 2(c) between seconds ten and 12. The increased salience at double the true f_0 combined with the relatively low pitch range of the melody (algorithms often

bias the tracking against low frequencies) results in the algorithm tracking the melody one octave above the correct pitch, thus producing the observed octave errors.

ALGORITHM OVERVIEW: 2005 TO DATE

Melody extraction is strongly linked to pitch (fundamental frequency) estimation, which has a long research tradition. Early approaches for pitch estimation in music dealt with the estimation of the f_0 of monophonic music recordings and were adopted from the speech processing literature [18]. Since then, various approaches specifically tailored for f_0 estimation in monophonic music signals have been proposed [19]. More recently, algorithms have also been proposed for estimating the f_0 of multiple concurrent instruments in polyphonic recordings (multipitch estimation). For a detailed review, the reader is referred to [20]. As seen in the section "Introduction," melody extraction differs from both monophonic and multipitch estimation in two important ways. Unlike monophonic pitch estimation, here we are dealing with polyphonic material and the challenges it entails. Unlike multipitch estimation, melody extraction requires the identification of the specific voice that carries the melody within the polyphony, but does not involve estimating the pitch values of the remaining sources.

It is instructive to consider melody extraction systems as elaborations of monophonic pitch trackers. Monophonic pitch trackers usually take the audio signal $x(t)$ and calculate a function $S_x(f_\tau, \tau)$ evaluated across a range of candidate pitch frequencies f that indicates the relative score or likelihood of the pitch candidates at each time frame τ. The function can be calculated either in the time domain (e.g., the autocorrelation evaluated over a range of lags) or the frequency domain (e.g., some function of the magnitude spectrum evaluated over a range of frequencies). The local estimates of period are then typically subject to sequential constraints, for instance, via dynamic programming. Thus, the estimated sequence of pitch values $\hat{\mathbf{f}}$, represented as a vector with one value for each time frame, is derived as

$$\hat{\mathbf{f}}_{\text{mon}} = \arg\max_{\mathbf{f}} \sum_{\tau} S_x(f_\tau, \tau) + C(\mathbf{f}), \tag{1}$$

where f_τ is the τth element of $\mathbf{f}$, and $C(\mathbf{f})$ accounts for the temporal constraints. For example, a common choice for $S_x(f, \tau)$ is an autocorrelation function such as

$$S_x(f, \tau) = r_{xx}\left(\frac{1}{f}; \tau\right) = \frac{1}{W}\int_{\tau - W/2}^{\tau + W/2} x(t)\, x\left(t + \frac{1}{f}\right) dt, \tag{2}$$

where W is the length of the autocorrelation analysis window. In melody extraction, the observed signal $y(t)$ consists of a target monophonic melody signal $x(t)$ with added accompaniment "noise"

$$y(t) = x(t) + n(t). \tag{3}$$

There are two paths to extending monophonic trackers to succeed in such conditions: we could improve the robustness of the underlying pitch candidate scoring function, so it continues to reflect the desired pitch even in the presence of other periodicities; we call this *salience-based* melody extraction

$$\hat{\mathbf{f}}_{\text{sal}} = \arg\max_{\mathbf{f}} \sum_{\tau} S'_y(f_\tau, \tau) + C'(\mathbf{f}), \tag{4}$$

where S'_y is the modified pitch salience function calculated over the mixed signal y. There are many different approaches for calculating the salience function (cf. the section "Salience Function"). For instance, some functions compute the salience of a candidate frequency f as the weighted sum of its harmonics

$$S'_y(f_\tau, \tau) = \sum_{h=1}^{N_h} g(f_\tau, h)\, |Y(h \cdot f, \tau)|, \tag{5}$$

where N_h is the number of harmonics in the summation, $g(f_\tau, h)$ is a harmonic weighting function [5], and $Y(f, \tau)$ is the short-time Fourier transform (STFT),

$$Y(f, \tau) = \int_{-W/2}^{W/2} w(t)\, y(\tau + t)\, e^{-j2\pi f t}\, dt, \tag{6}$$

where $w(t)$ is a windowing function.

Note that in (4) we now use $C'(\mathbf{f})$ to represent the temporal constraints instead of $C(\mathbf{f})$, since for the polyphonic case this is a far more complex problem: even with a modified salience function there is no guarantee that the frequency of the melody will always be found at the maximum of the function. As shall be seen in the section "Tracking," this is addressed by employing tracking techniques such as Viterbi decoding, tracking agents, clustering, etc.

Alternatively, we could attempt to decompose the mixed signal into separate sources, at least one of which, $\hat{x}(t)$, is dominated by the melody signal to a degree that makes it suitable for a largely unmodified pitch tracker; we call this *source separation* melody extraction

$$\hat{\mathbf{f}}_{\text{sep}} = \arg\max_{\mathbf{f}} \sum_{\tau} S_{\hat{x}}(f_\tau, \tau) + C'(\mathbf{f}), \tag{7}$$

where $\hat{x}(t)$ is estimated using decomposition or matrix factorization techniques (cf. the section "Source Separation-Based Approaches").

THE MIREX MELODY EXTRACTION EVALUATIONS

Since its initiation in 2005, over 50 melody extraction algorithms have been submitted to MIREX [3]. In this annual campaign, different algorithms are evaluated against the same set of music collections to obtain a quantitative comparison between methods and assess the accuracy of the current state of the art in melody extraction. We believe MIREX is a good point of reference for this review, given that the large majority of melody extraction algorithms that have had an impact on the research community have participated in MIREX at some point. Due to space limitations, approaches predating 2005 (e.g., [1]) are not discussed in this article, and we refer the reader to [20] for further information on earlier work.

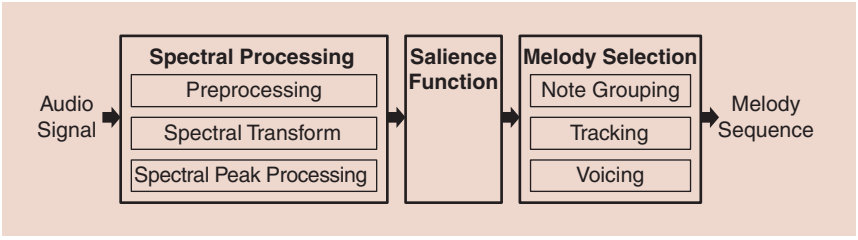
In Table 1, we provide a summary of the characteristics of a selection of 16 representative algorithms out of all the

[TABLE 1] ALGORITHMIC ARCHITECTURE OF 16 MELODY EXTRACTION ALGORITHMS FROM MIREX FROM 2005 TO 2012.

FIRST AUTHOR/ MIREX YEAR	PREPROCESSING	SPECTRAL TRANSFORM AND PROCESSING	MULTIPITCH REP. (SALIENCE FUNCTION)	TRACKING	VOICING	APPROACH TYPE
PAIVA 2005 [25]	—	AUDITORY MODEL + AUTOCORRELATION PEAKS	SUMMARY CORRELOGRAM	MULTIPITCH TRAJECTORIES + NOTE DELETION	SALIENCE VALLEYS	SALIENCE BASED
MAROLT 2005 [26]	—	STFT + SMS HARMONICS PLUS NOISE	EM FIT TO TONE MODELS	FRAGMENTS + FRAGMENT CLUSTERING	LOUDNESS FILTER	SALIENCE BASED
GOTO 2005 [27]	BANDPASS FILTER	MULTIRATE FILTERBANK + IF-BASED PEAK SELECTION	EM FIT TO TONE MODELS	TRACKING AGENTS	—	SALIENCE BASED
CANCELA 2008 [28]	—	CONSTANT-Q + HIGH PASS FILTER + LOG POWER NORM.	HARMONICITY MAP	CONTOUR TRACKING + WEIGHTING + SMOOTHING	ADAPTIVE THRESHOLD	SALIENCE BASED
RYYNÄNEN 2008 [5]	—	STFT + SPECTRAL WHITENING	HARMONIC SUMMATION	NOTE EVENT HMM + GLOBAL HMM	SILENCE MODEL	SALIENCE BASED
DRESSLER 2009 [29]	—	MRFFT + IF PEAK CORRECTION + MAGNITUDE THRESH.	PAIRWISE COMPARISON OF SPECTRAL PEAKS	STREAMING RULES	DYNAMIC THRESHOLD	SALIENCE BASED
RAO 2009 [30]	—	HIGH RESOLUTION FFT + MAIN-LOBE MAG. MATCHING	SMS + TWM	DYNAMIC PROGRAMMING	NHC THRESHOLD	SALIENCE BASED
SALAMON 2011 [16]	EQUAL LOUDNESS FILTER	STFT + IF PEAK CORRECTION	HARMONIC SUMMATION	CONTOUR TRACKING + CONTOUR FILTERING	SALIENCE DISTRIBUTION	SALIENCE BASED
JO 2011 [31]	—	STFT WITH VARYING WINDOW LENGTH	HARMONIC SUMMATION	STABLE CANDIDATES + RULE-BASED SELECTION	IMPLICIT	SALIENCE BASED
ARORA 2012 [32]	—	STFT + LOG SPECTRUM + PEAK SELECTION	IFT OF LOG SPECTRUM	HARMONIC CLUSTER TRACKING + CLUSTER SCORE	HARM. SUM. THRESHOLD	SALIENCE BASED
HSU 2010 [33]	HARM/PERC SOUND SEP.	MRFFT + VOCAL PARTIAL DISCRIMINATION	NORMALIZED SUBHARMONIC SUMMATION	GLOBAL TREND + DYNAMIC PROGRAMMING	CLASSIFICATION	SALIENCE BASED + SOURCE SEP. PREPROCESSING
YEH 2012 [34]	HARM/PERC SOUND SEP.	MRFFT + VOCAL PARTIAL DISCRIMINATION	NORMALIZED SUBHARMONIC SUMMATION	TREND ESTIMATION + HMM	—	SALIENCE BASED + SOURCE SEP. PREPROCESSING
DURRIEU 2009 [22]	SOURCE/FILTER MODEL FOR MELODY SOURCE SEPARATION			VITERBI SMOOTHING	ENERGY THRESHOLD	SOURCE SEPARATION
TACHIBANA 2011 [23]	TWO-STAGE HARMONIC/PERCUSSIVE SOUND SEPARATION			DYNAMIC PROGRAMMING	SIGNAL/NOISE RATIO THRESHOLD	SOURCE SEPARATION
POLINER 2006 [21]	DOWNSAMPLE TO 8 kHz	STFT + LIMIT TO 2 kHz + NORMALIZE MAGNITUDE	N/A	SUPPORT VECTOR MACHINE CLASSIFIER	ENERGY THRESHOLD	DATA DRIVEN
SUTTON 2006 [35]	SEMITONE ATT. + BANDPASS	N/A	N/A	HMM COMBINATION OF MONOPHONIC PITCH TRACKERS	CONFIDENCE HMM	MONOPHONIC

submissions to MIREX since 2005. To do so, we have attempted to break down the extraction process into a series of steps that are common to most algorithms. Since some authors submitted several algorithms over the years, we have opted to include only their most recent (published) contribution, as in most cases it represents the latest version in the evolution of a single algorithm. If a certain step is not included in an algorithm (or otherwise not mentioned by the authors) a "—" is placed in the table. "N/A" means a step is not relevant to the method (e.g., Poliner and Ellis [21] determine the melody directly from the power spectrum and hence a multipitch representation of the audio signal is not relevant for this approach). Finally, we note that some algorithms (those by Durrieu [22] and Tachibana [23]) cannot be broken down into the same steps as the rest of the approaches. This is indicated by fusing the columns of some steps in the table for these algorithms.

The last column of the table, "Approach Type," attempts to classify the algorithms based on their underlying approach, with most falling into the categories of salience based and source separation introduced above. Some approaches, however, do not fit into either category, including the data-driven approach in which the power spectrum is fed directly into a machine-learning algorithm that attempts to classify the melody frequency based on the observed spectrum at each frame.



[FIG3] **A block diagram of salience-based melody extraction algorithms.**

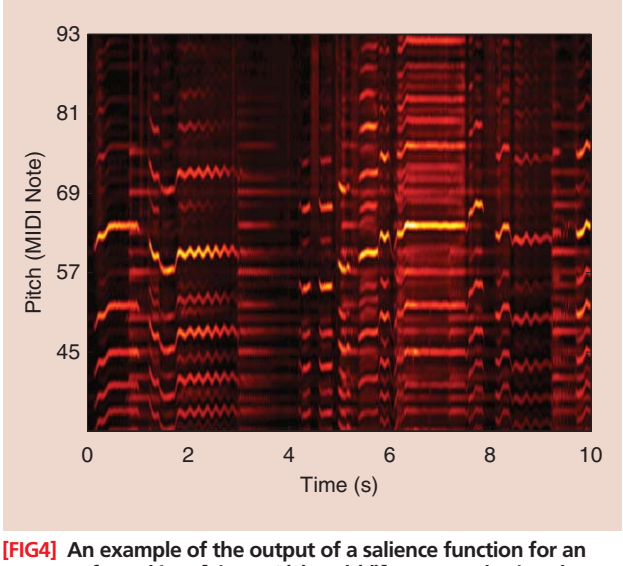
Note that while melody extraction includes detecting both sung melodies and melodies played by lead instruments, many algorithms are developed particularly for singing voice extraction. The reason for this is twofold: first, there is a large body of popular music with sung melodies, which makes vocal melody extraction commercially attractive. Second, the singing voice has unique characteristics that are different from most instruments [24], and algorithms can exploit these unique features to identify the melody more accurately.

SALIENCE-BASED APPROACHES

As evident in Table 1, the largest set of approaches are those based on time-frequency representations of pitch salience (a salience function). The general architecture of these approaches, with possible substeps, is depicted in Figure 3.

PREPROCESSING

As a first step, some approaches apply some type of preprocessing, normally a filter to enhance the frequency content where we expect to find the melody: Goto [27] applies a bandpass filter between 261.6 Hz and approximately 4 kHz, while Salamon and Gómez [16] apply a perceptually motivated equal loudness filter [7]. Some approaches use source separation to enhance the melody signal before it is further processed: Hsu [33] and Yeh [34] use a technique originally designed for harmonic-percussive sound separation (HPSS) adapted to perform melody-accompaniment separation (cf. the section "Source Separation-Based Approaches").



[FIG4] **An example of the output of a salience function for an excerpt of vocal jazz [Figure 2(a) and (d)] computed using the algorithm proposed in [16].**

SPECTRAL TRANSFORM AND PROCESSING

Next, the signal is chopped into time frames and a transform function is applied to obtain a spectral representation of each frame. The most straightforward approach is to apply the STFT, with a window length typically between 50 and 100 ms [5], [16], [26], [30], [32]. Such a window length usually provides sufficient frequency resolution to distinguish different notes while maintaining adequate time resolution to track pitch changes in the melody over short time periods. Still, some approaches attempt to overcome the time-frequency resolution limitation inherent to the Fourier transform by applying a multiresolution transform such as a multirate filterbank [27], the constant-Q transform [28], or the multiresolution FFT (MRFFT) [33], [34], [36]. In general, these transforms use larger windows at low frequencies (where we require greater frequency resolution to resolve close notes) and small windows at higher frequencies (where we need high-temporal resolution to track rapidly changing harmonics). In [16], a comparison between the STFT and MRFFT showed there was no statistically significant difference between using one transform over another for melody extraction. Nonetheless, since each step in a melody extraction system is highly sensitive to the output of the preceding step, it is possible that some algorithms do benefit from using multiresolution transforms. Finally, we note that some methods use transforms designed to emulate the human auditory system [7] such as the model used by Paiva [25].

After applying the transform, most approaches only use the spectral peaks for further processing. Apart from detecting the peaks themselves, different peak processing techniques may be applied: some methods filter peaks based on magnitude or sinusoidality criteria in an attempt to filter out peaks that do not represent harmonic content or the lead voice [26], [27], [30], [33], [34]. Other approaches apply spectral magnitude normalization in an attempt to reduce the influence of timbre on the analysis—Cancela [28] and Arora [32] take the log spectrum and Ryynänen and Klapuri (who use the whole spectrum, not just the peaks) apply spectral whitening [5]. Finally, Dressler [36] and Salamon and Gómez [16] obtain more accurate frequency and amplitude estimates for each spectral peak by computing its instantaneous frequency from the phase spectrum.

SALIENCE FUNCTION

At the core of salience-based algorithms lies the multipitch representation, i.e., the salience function. This function provides an estimate of the salience of each possible pitch value (within the range where we expect to find the melody) over time. An example of the output of a salience function (used by

Salamon and Gómez [16]) is depicted in Figure 4. The peaks of this function are taken as possible candidates for the melody, which are further processed in the next stages. Different methods can be used to obtain a salience function: most approaches use some form of harmonic summation, by which the salience of a certain pitch is calculated as the weighted sum of the amplitude of its harmonic frequencies [5], [16], [28], [31], [33], [34]. Goto [27] and Marolt [26] use expectation maximisation to fit a set of tone models to the observed spectrum. The estimated maximum a posteriori probability (MAP) of the tone model whose f_0 corresponds to a certain pitch is considered to be the salience of that pitch. Other approaches include two-way mismatch computed by Rao [30], summary autocorrelation used by Paiva [25], and pairwise analysis of spectral peaks as done by Dressler [37].

As evident in Figure 4, the salience function approach has one main undesirable effect—the appearance of the "ghost" pitch values whose f_0 is an exact multiple (or submultiple) of the f_0 of the actual pitched sound. This effect can lead to what is commonly referred to as *octave errors*, in which an algorithm selects a pitch value that is exactly one octave above or below the correct pitch of the melody. [This type of error can be observed in Figure 2(f).] Different algorithms adopt different strategies to reduce the number of octave errors they commit. Some algorithms, such as the ones by Cancela [28] and Dressler [29], attempt to directly reduce the number of ghost pitch values present in the salience function. Dressler does this by examining pairs of spectral peaks that potentially belong to the same harmonic series and attenuating the result of their summation if there are many high amplitude spectral peaks whose frequencies lie between the pair being considered. Cancela attenuates the harmonic summation supporting a certain f_0 if the mean amplitude of spectral components at frequencies $2k \cdot f_0, 3k \cdot f_0/2$ and $3k \cdot f_0$ is above the mean of the components at frequencies $k \cdot f_0$ (this will attenuate ghost pitch values whose f_0 is 1/2, 2/3, or 1/3 of the real f_0). In [20], Klapuri proposes a method for reducing octave errors based on spectral smoothness. The amplitude of each peak in the salience function is recalculated after smoothing the spectral envelope of its corresponding harmonic frequencies. Peaks representing octave errors will have an irregular envelope (compared to a smoother envelope for real notes) and thus will be attenuated by this process. An alternative approach for coping with octave errors is proposed by Paiva [25] and Salamon [16], who first group the peaks of the salience function into pitch contours and then determine which contours are actually ghost contours and remove them. The underlying idea is that once salience peaks are grouped into contours, detecting duplicate contours becomes easier since they have identical trajectories one octave apart. Determining which of the two is the ghost contour is done using criteria based on contour salience and the overall pitch continuity of the melody. Finally, we note that practically all methods reduce octave errors nonexplicitly by penalizing large jumps in pitch during the tracking stage of the algorithm.

TRACKING

Given the peaks of the salience function, the remaining task is to determine which peaks (i.e., pitch values) belong to the melody. This is one of the most crucial stages of each algorithm and, interestingly, it is also perhaps the most varied step where practically every algorithm uses a different approach. Most approaches attempt to directly track the melody from the salience peaks, though some (Paiva, Marolt, Cancela, and Salamon) include a preliminary grouping stage where peaks are grouped into continuous pitch contours (also referred to as *fragments* or *trajectories*) out of which the melody is later selected [16], [25], [26], [28]. This grouping is usually performed by tracking sequential peaks based on time, pitch, and salience continuity constraints. Given the pitch contours (or salience peaks if no grouping is applied), a variety of tracking techniques have been proposed to obtain the final melody sequence: Marolt [26] uses clustering, while Goto [27] and Dressler [29] use heuristic-based tracking agents. Ryynänen [5] and Yeh [34] use HMMs, while Rao [30] and Hsu [33] use dynamic programming. Finally, Paiva [25] and Salamon [16] take a different approach—rather than tracking the melody, they attempt to delete all pitch contours (or notes) that do not belong to the melody.

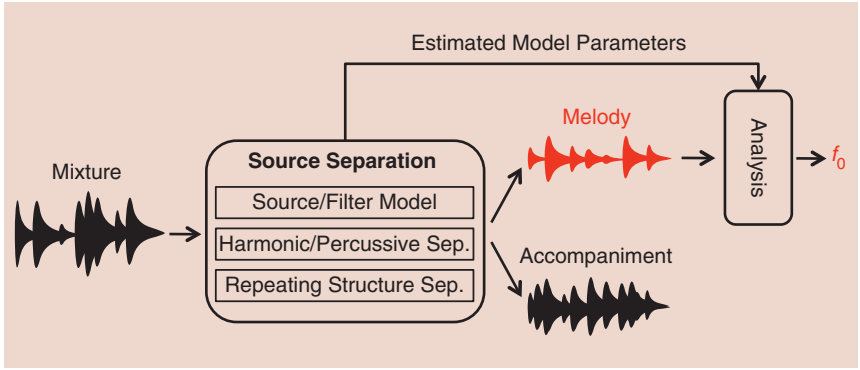
VOICING

An important part of melody extraction that is sometimes overlooked is voicing detection, i.e., determining when the melody is present and when it is not. The voicing detection step of an algorithm is usually applied at the very end, though exceptions do exist (e.g., Salamon uses a threshold based on the salience distribution of pitch contours in the entire piece to remove nonsalient contours before proceeding to filter out other nonmelody contours). A common approach is to use a fixed or dynamic per-frame salience-based threshold, as done by Paiva, Marolt, Cancela, Dressler, Rao, and Arora. Alternative strategies include Ryynänen's algorithm, which incorporates a silence model into the HMM tracking part of the algorithm, and Hsu's algorithm, which uses timbre-based classification to determine the presence (or absence) of human voice.

SOURCE SEPARATION-BASED APPROACHES

An alternative strategy to salience-based melody extraction is to use source separation algorithms to isolate the melody source from the mixture. A block diagram illustrating some of the strategies for melody extraction using source separation is provided in Figure 5. This type of approach is the most recent of the ones mentioned in Table 1 and has gained popularity in recent years following the advances in audio source separation research. While there is a large body of research on melody and lead voice source separation (cf. [22] and [38]–[43] and references therein), such algorithms are usually evaluated using measures based on signal to noise ratios, and only few have been evaluated in terms of estimating the frequency sequence of the melody, as is our goal here.

Two methods in Table 1 are source separation based—those of Durrieu et al. [22] and Tachibana et al. [23]. Durrieu models the power spectrogram of the signal as the instantaneous sum of two



[FIG5] **A block diagram of source separation-based melody extraction algorithms.**

contributions: the lead voice and the accompaniment. The contribution of the lead voice is represented with a source/filter model, and the contribution of the accompaniment as the sum of an arbitrary number of sources with distinct spectral shapes. Two different representations are proposed for the source/filter model: a smooth instantaneous mixture model (SIMM) and a smooth Gaussian scaled mixture model (SGSMM). The former represents the lead instrument (or voice) as the instantaneous mixture of all possible notes, while the latter is more realistic in that it only allows one source/filter couple to be active at any moment, albeit computationally heavier. In both cases, the model parameters are estimated using an expectation maximization framework. Once the model parameters are estimated, the final melody sequence is obtained using the Viterbi algorithm to find a smooth trajectory through the model parameters (which include the f_0 of the source). Voicing detection is done by first using Wiener filtering to separate the melody signal based on the estimated model parameters, and then computing the energy of this signal at every frame to determine an energy threshold for frames where the melody is present.

The approach proposed by Tachibana et al. is quite distinct. It is based on exploiting the temporal variability of the melody compared to more sustained chord notes. To do so, they make use of the HPSS algorithm [44]. The algorithm was originally designed to separate harmonic from percussive elements in a sound mixture by separating sources that are smooth in time (harmonic content) and sources smooth in frequency (percussive content). By changing the window length used for the analysis, the algorithm can be used to separate "sustained" (i.e., chord) sounds from "temporally variable" (melody plus percussive) sounds. Once the accompaniment is removed, the algorithm is run again, this time in its original form to remove percussive elements. After these two passes, the melody in the resulting signal should be significantly enhanced. The melody frequency sequence is obtained directly from the spectrogram of the enhanced signal using dynamic programming by finding the path which maximizes the MAP of the frequency sequence, where the probability of a frequency given the spectrum is proportional to the weighted sum of the energy at its harmonic multiples, and transition probabilities are a function of the distance between two subsequent frequency values. Voicing detection is done by setting a threshold on the (Mahalanobis) distance between the two signals produced by the second run of the HPSS algorithm (the melody signal and the percussive signal).

Finally, in Table 1 we see that some authors attempt to combine salience-based and source separation approaches. Here, source separation is used as a preprocessing step to attenuate the accompaniment signal, and then a salience function is computed from the processed signal. Both Hsu [33] and Yeh [34] use the HPSS method proposed by Tachibana, but rather than attempt to estimate the melody directly from the spectrum of the resulting signal, they continue to compute a salience function and further steps similar to other salience-based approaches.

For completeness, we briefly describe some singing voice source separation algorithms here. As mentioned earlier, while these methods have not been evaluated in terms of melody extraction, they could be used to build melody extraction systems by combining them with a monophonic pitch tracking algorithm that estimates the melody f_0 sequence from the separated voice signal, or by using them as a preprocessing step similar to the aforementioned approaches by Hsu and Yeh. We have already seen the source/filter model proposed by Durrieu et al. [22] and the HPSS method employed by Tachibana et al. [23]. A different strategy for separating the lead voice is to exploit the fact that the music accompaniment often has a repetitive structure, while the voice contains more variation. Huang et al. [41] exploit this by assuming that the spectrogram of the accompaniment can be modeled by a low-rank matrix, and the spectrogram of the voice by a sparse matrix. They use robust principal component analysis (RPCA) to factorize the spectrogram of the signal into the desired voice and accompaniment matrices. A different way of exploiting repetition is proposed by Rafii and Pardo [42]—they first compute the repetition period of the accompaniment using autocorrelation applied to the spectrogram of the mixture. By computing the median of the spectrograms of consecutive repetitions, they obtain a spectrogram that contains only the repeating signal (the accompaniment). This spectrogram is used to derive a time-frequency mask used to separate the voice from the accompaniment. This approach was extended by Liutkus et al. [43] to work on full songs (where the repetition period can change between verse and chorus) by searching for local periodicities in a song, and again by Rafii and Pardo by applying the algorithm to local windows of the signal and by computing a self-similarity matrix to better identify repeating segments in a song. In [42], the authors also present some experiments on combining their approach with existing pitch trackers to perform melody extraction, and we expect to see an increase in the number of source separation-based melody extraction algorithms participating in MIREX in the future.

ALTERNATIVE APPROACHES

While most melody extraction approaches are either salience or source separation based, some very different strategies have been

proposed as well. The first to appear in Table 1 is the data-driven approach by Poliner and Ellis [21]. Rather than handcraft knowledge about musical acoustics into the system (e.g., in the form of a salience function based on harmonic summation), they propose to use machine learning to train a classifier to estimate the melody note directly from the power spectrum. As a preprocessing step they downsample the audio to 8 kHz, and use the STFT to obtain a spectral representation. Bins corresponding to frequencies above 2 kHz are discarded and the magnitude of the remaining bins is normalized over a short time period to reduce the influence of different instrument timbres. The resulting 256 feature vector is used to train a support vector machine classifier using training data where each frame is labeled with one of 60 possible output classes corresponding to 60 MIDI notes spanning five octaves. Voicing detection is done by means of a global threshold based on the magnitude squared energy found between 200 and 1,800 Hz.

Another completely different strategy is the one proposed by Sutton et al. [35]. Rather than design an algorithm to handle polyphonic audio signals, they compute the pitch sequences returned by two different monophonic pitch estimators and then combine them using an HMM. The underlying assumption is that while monophonic pitch estimators are not designed to handle audio where there is more than one pitch present at a time (normally leading to a large degree of estimation errors), by combining the output of different estimators a more reliable result could be obtained.

EVALUATION: MEASURES AND MUSIC COLLECTIONS

As explained earlier, melody extraction algorithms are expected to accomplish two goals: estimate the correct pitch of the melody (pitch estimation), and estimate when the melody is present and when it is not (voicing detection). The output of a melody extraction algorithm typically includes two columns, the first with timestamps at a fixed interval (e.g., for MIREX a 10-ms interval is used), and the second with f_0 values representing the algorithm's pitch estimate for the melody at each timestamp (i.e., at each analysis frame). Algorithms can report a pitch even for frames where they estimate the melody to be absent (nonmelody frames), in this way allowing us to evaluate pitch estimation and voicing detection independently.

To evaluate the performance of an algorithm for a given audio excerpt, we compare the algorithm's output with the excerpt's ground truth. The ground truth file has the same format as the output file, and contains the correct series of f_0 values representing the melody of the excerpt. The ground truth is produced by running a monophonic pitch tracker on the solo melody track of the excerpt (meaning we require access to the multitrack recording of every song we use for evaluation). Using a graphical user interface such as SMSTools or WaveSurfer (see "For More Information"), the output of the monophonic pitch tracker is manually inspected and corrected if necessary. Given the ground truth file, an algorithm is evaluated by comparing its output on a per-frame basis to the ground truth. For nonmelody frames in the ground truth, the algorithm is expected to indicate that it has detected the absence of melody. For melody frames, the algorithm is expected to return a frequency value matching the one in the ground truth. An algorithm's frequency estimate is considered correct if it is within 50 cents (i.e., half a semitone) of the ground truth.

MELODY EXTRACTION ALGORITHMS ARE EXPECTED TO ACCOMPLISH TWO GOALS: ESTIMATE THE CORRECT PITCH OF THE MELODY (PITCH ESTIMATION), AND ESTIMATE WHEN THE MELODY IS PRESENT AND WHEN IT IS NOT (VOICING DETECTION).

MEASURES

Based on this per-frame comparison, we compute five global measures that assess different aspects of the algorithm's performance for the audio excerpt in question. These measures were first used in MIREX 2005 [2], and have since become the de facto set of measures for evaluating melody extraction algorithms. If the system's estimated melody pitch frequency vector is $\mathbf{f}$ and the true sequence is $\mathbf{f}^*$, let us also define a voicing indicator vector $\mathbf{v}$, whose τth element $v_\tau = 1$ when a melody pitch is detected, with corresponding ground truth $\mathbf{v}^*$. We also define an "unvoicing" indicator $\bar{v}_\tau = 1 - v_\tau$. Recall that an algorithm may report an estimated melody pitch ($f_\tau > 0$) even for times where it reports no voicing ($v_\tau = 0$). Then the measures are as follows:

- **Voicing recall rate:** The proportion of frames labeled as melody frames in the ground truth that are estimated as melody frames by the algorithm

$$\text{Rec}_{\text{vx}} = \frac{\sum_\tau v_\tau v_\tau^*}{\sum_\tau v_\tau^*}. \tag{8}$$

- **Voicing false alarm rate:** The proportion of frames labeled as nonmelody in the ground truth that are mistakenly estimated as melody frames by the algorithm

$$\text{FA}_{\text{vx}} = \frac{\sum_\tau v_\tau \bar{v}_\tau^*}{\sum_\tau \bar{v}_\tau^*}. \tag{9}$$

- **Raw pitch accuracy:** The proportion of melody frames in the ground truth for which f_τ is considered correct (i.e., within half a semitone of the ground truth f_τ^*)

$$\text{Acc}_{\text{pitch}} = \frac{\sum_\tau v_\tau^* \mathcal{T}[\mathcal{M}(f_\tau) - \mathcal{M}(f_\tau^*)]}{\sum_\tau v_\tau^*}, \tag{10}$$

where $\mathcal{T}$ is a threshold function defined by

$$\mathcal{T}[a] = \begin{cases} 1 & \text{if } |a| < 0.5 \\ 0 & \text{if } |a| \geq 0.5 \end{cases} \tag{11}$$

and $\mathcal{M}$ maps a frequency in Hertz to a melodic axis as a real-valued number of semitones above an arbitrary reference frequency f_{ref} (55 Hz, or note pitch A1, in this work):

[TABLE 2] TEST COLLECTIONS FOR MELODY EXTRACTION EVALUATION IN MIREX.

COLLECTION	DESCRIPTION
ADC2004	20 EXCERPTS OF ROUGHLY 20 s IN THE GENRES OF POP, JAZZ, AND OPERA. INCLUDES REAL RECORDINGS, SYNTHESIZED SINGING, AND AUDIO GENERATED FROM MIDI FILES. TOTAL PLAY TIME: 369 s.
MIREX05	25 EXCERPTS OF 10–40 s DURATION IN THE GENRES OF ROCK, R&B, POP, JAZZ, AND SOLO CLASSICAL PIANO. INCLUDES REAL RECORDINGS AND AUDIO GENERATED FROM MIDI FILES. TOTAL PLAY TIME: 686 s.
INDIAN08	FOUR 1-MIN-LONG EXCERPTS FROM NORTH INDIAN CLASSICAL VOCAL PERFORMANCES. THERE ARE TWO MIXES PER EXCERPT WITH DIFFERENT AMOUNTS OF ACCOMPANIMENT, RESULTING IN A TOTAL OF EIGHT AUDIO CLIPS. TOTAL PLAY TIME: 501 s.
MIREX09 (0 dB)	374 KARAOKE RECORDINGS OF CHINESE SONGS (i.e., RECORDED SINGING WITH KARAOKE ACCOMPANIMENT). THE MELODY AND ACCOMPANIMENT ARE MIXED AT A 0-dB SIGNAL-TO-ACCOMPANIMENT RATIO. TOTAL PLAY TIME: 10,020 s.
MIREX09 (–5 dB)	SAME 374 EXCERPTS AS MIREX09 (0 dB), BUT HERE THE MELODY AND ACCOMPANIMENT ARE MIXED AT A –5-dB SIGNAL-TO-ACCOMPANIMENT RATIO. TOTAL PLAY TIME: 10,020 s.
MIREX09 (+5 dB)	SAME 374 EXCERPTS AS MIREX09 (0 dB), BUT HERE THE MELODY AND ACCOMPANIMENT ARE MIXED AT A +5-dB SIGNAL-TO-ACCOMPANIMENT RATIO. TOTAL PLAY TIME: 10,020 s.

$$\mathcal{M}(f) = 12\log_2\left(\frac{f}{f_{\text{ref}}}\right). \tag{12}$$

■ **Raw chroma accuracy:** As raw pitch accuracy, except that both the estimated and ground truth f_0 sequences are mapped onto a single octave. This gives a measure of pitch accuracy that ignores octave errors, a common error made by melody extraction systems

$$\text{Acc}_{\text{chroma}} = \frac{\sum_\tau v_\tau^* \mathcal{T}[\langle \mathcal{M}(f_\tau) - \mathcal{M}(f_\tau^*)\rangle_{12}]}{\sum_\tau v_\tau^*}. \tag{13}$$

Octave equivalence is achieved by taking the difference between the semitone-scale pitch values modulo 12 (one octave), where

$$\langle a\rangle_{12} = a - 12\left\lfloor \frac{a}{12} + 0.5\right\rfloor. \tag{14}$$

■ **Overall accuracy:** This measure combines the performance of the pitch estimation and voicing detection tasks to give an overall performance score for the system. It is defined as the proportion of all frames correctly estimated by the algorithm, where for nonmelody frames this means the algorithm labeled them as nonmelody, and for melody frames the algorithm both labeled them as melody frames and provided a correct f_0 estimate for the melody (i.e., within half a semitone of the ground truth)

$$\text{Acc}_{\text{ov}} = \frac{1}{L}\sum_\tau v_\tau^* \mathcal{T}[\mathcal{M}(f_\tau) - \mathcal{M}(f_\tau^*)] + \bar{v}_\tau^* \bar{v}_\tau, \tag{15}$$

where L is the total number of frames.

The performance of an algorithm on an entire music collection for a given measure is obtained by averaging the per-excerpt scores for that measure over all excerpts in the collection.

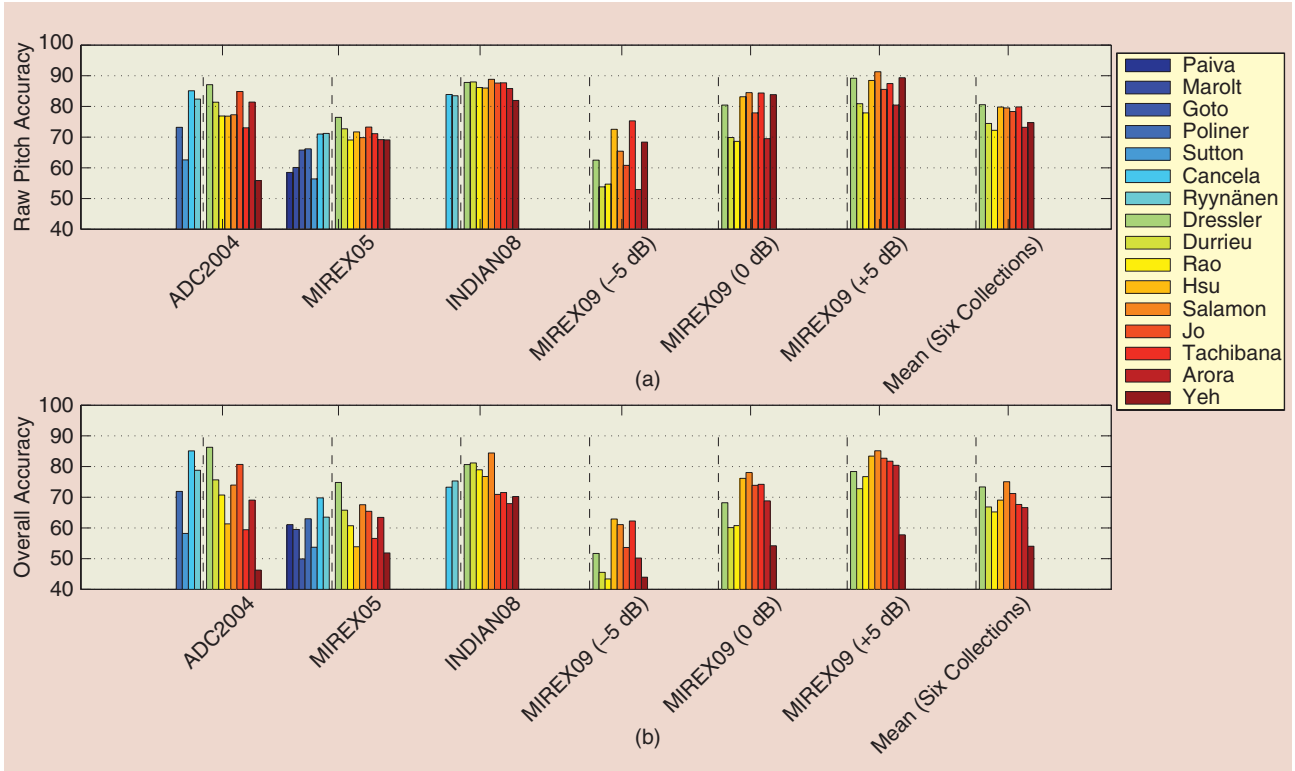
MUSIC COLLECTIONS

Over the years, different research groups have contributed annotated music collections for evaluating melody extraction in MIREX. The limited amount of multitrack recordings freely available, and the time-consuming annotation process, mean most of these collections are relatively small compared to those used in other MIR tasks. The collections currently used for evaluation in MIREX, which have remained fixed since 2009, are described in Table 2.

PERFORMANCE: 2005 TO DATE

EXTRACTION ACCURACY

In Figure 6, we present the results obtained by the 16 algorithms in Table 1 for the MIREX evaluation collections. Note that some algorithms only participated in MIREX before all the collections were added, meaning we only have partial results for these algorithms. This is indicated in the graph with vertical dashed lines that separate the algorithms that were only evaluated on some of the collections (to the left of the line) from those evaluated on all collections (to the right of the line). We only compute the mean for algorithms evaluated on all six collections. To get a general idea of the performance of the algorithms, it is sufficient to focus on two evaluation measures—the raw pitch accuracy [Figure 6(a)] and the overall accuracy [Figure 6(b)]. The former tells us how well the algorithm tracks the pitch of the melody, and the latter combines this measure with the efficiency of the algorithm's voicing detection, meaning the voicing-related measures are (to an extent) also reflected in this measure. Starting with the raw pitch, the first thing we note is that the accuracy of all algorithms varies depending on the collection being analyzed. While some collections are generally harder for all approaches (e.g., MIREX09 (−5 dB) where the accompaniment is louder and masks the melody), in general the variability in performance is not homogeneous. This highlights the advantages and disadvantages of different approaches with respect to the music material being analyzed. For instance, we see that Dressler's method outperforms all others for the ADC2004 and MIREX05 collections, which contain a mixture of vocal and instrumental pieces, but does not for the other collections where the melody is always vocal. On the one hand this means that her approach is generalizable to a wider range of musical material, but on the other hand we see that approaches that take advantage of specific features of the human voice (e.g., Tachibana or Salamon) can do better on vocal melodies. We also see that the HPSS melody enhancement applied by Hsu, Tachibana, and Yeh is particularly advantageous when the melody source is relatively weak compared to the accompaniment [MIREX09 (−6 dB)]. Finally, examining the raw pitch accuracy results for the MIREX05 collection, we see that results have improved gradually from 2005 to 2009, after which raw



[FIG6] (a) Raw pitch accuracy and (b) overall accuracy obtained in MIREX by the 16 melody extraction algorithms in Table 1. The vertical dashed line separates the algorithms that were only evaluated on some of the collections (left of the line) from those evaluated on all six collections (right of the line).

pitch accuracies have remained relatively unchanged (more on the evolution of performance in the section "Are We Improving?"). Overall, we see that the average pitch accuracy over all collections lies between 70 and 80%.

Turning over to the overall accuracy, we see that performance goes down compared to the raw pitch accuracy for all algorithms, since voicing detection is now factored into the results. Note that the results for Goto and Yeh are artificially low since these methods do not include a voicing detection step. The importance of this step depends on the intended use of the algorithm. For example, if we intend to use it as a first step in a transcription system, it is very important that we do not include notes that do not belong to the melody in our output. On the other hand, similarity-based applications which rely on matching algorithms that can handle gaps in the alignment of melodic sequences may be less sensitive to voicing mistakes. If we look at the average results over all six collections, we see that the algorithms obtaining the best overall accuracy are those that obtain good raw pitch accuracy combined with an effective voicing detection method. Generally, we see that overall accuracy results lie between 65 and 75% for the best performing algorithms. While this clearly indicates that there are still many challenges remaining (see the section "Challenges"), this degree of accuracy is in fact good enough for new applications to be built on top of melody extraction algorithms (cf. the section "Software and Applications").

Finally, we note that one important aspect of performance that is not reflected in Figure 6 is the computational cost of each approach. Depending on the intended application, we may have limited resources (e.g., time, computing power) and this can influence our decision when choosing which algorithm to use. While deriving O-notation complexity estimates is too complicated for some of the algorithms, generally we observe that algorithms involving source separation techniques (which are often implemented as iterative matrix operations) tend to be significantly more computationally complex than salience-based approaches. In this respect Dressler's algorithm is of particular interest, obtaining both the lowest runtime and the highest mean overall accuracy among the algorithms participating in 2009 (only Salamon and Gómez obtain a higher mean accuracy, but there is no runtime information for 2011).

ARE WE IMPROVING?

In the previous section we noted that, for some collections, performance has not improved much over the last three to four years. In Figure 7, we present the evolution of the overall accuracy obtained for the six MIREX collections over the years. For each collection, we plot the best overall accuracy result obtained up to a given year (e.g., for 2008 we plot the best result obtained up to 2008, for 2009 the best result obtained up to 2009, etc.). Indeed, our previous observation seems to be confirmed—for the

two earliest collections (ADC2004 and MIREX05), we observe a steady improvement in results from 2005 to 2009, after which performance does not improve. For the more recent collections (INDIAN08 and the three MIREX09 collections), we see a gradual improvement up to 2011; in 2012 no algorithm outperformed its predecessors for any of the collections. This highlights an important limitation of the MIREX evaluation campaign—since the collections are kept secret, it is very hard for researchers to learn from the results to improve their algorithms. This limitation is discussed further in the section "Challenges."

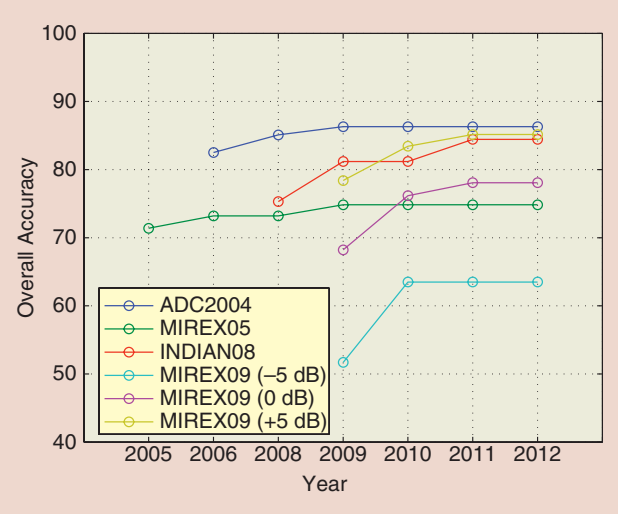
SOFTWARE AND APPLICATIONS

SOFTWARE

While various melody extraction algorithms have been proposed, relatively few implementations are freely available for people to download and use. Such tools are important for facilitating comparative evaluations, increasing the reproducibility of research and facilitating the development of new applications that make use of melody extraction technology (cf. the section "Applications Based on Melody Extraction"). Below we provide a list of known melody extraction related tools that are freely available (for links to all tools mentioned below see "For More Information")

- LabROSAmelodyextract2005 includes the code for the melody extraction system submitted by Poliner and Ellis to MIREX 2005 [21]. Runs on Linux and OSX systems and requires both MATLAB and Java.
- FChT is an open source MATLAB/C++ implementation of the Fan Chirp Transform (FChT) and f_0 gram (salience function) proposed by Cancela et al. in [45].
- separateLeadStereo is an open-source python implementation of the algorithm by Durrieu et al. reported in [40]. The code includes functionality for melody extraction, as well as lead instrument/accompaniment source separation.
- IMMF0salience is an open-source vamp plug-in for visualizing a salience function derived from the intermediate steps of the algorithm by Durrieu et al. [22], [40].
- MELODIA is a vamp plug-in available as a compiled library for Windows, OSX, and Linux. The plug-in implements the melody extraction algorithm by Salamon and Gómez [16], and in addition to its final output (i.e., the f_0 sequence of the melody) it provides visualizations of intermediate steps of the algorithm such as the salience function and pitch contours computed before selecting the final melody.

For completeness, we also briefly mention some commercially available software: Dressler's algorithm is incorporated in Fraunhofer's "Melody Extraction for Music Games" library, and certain melody extraction functionality is also incorporated in Adobe Audition and Melodyne, though the details of the algorithms used in these products are not published.



[FIG7] The evolution of the best overall accuracy result over the years for the six MIREX collections.

> THE ADVANCES IN ALGORITHMIC PERFORMANCE OF MELODY EXTRACTION ALGORITHMS OVER THE PAST DECADE MEAN THEY NOW PROVIDE SUFFICIENTLY GOOD RESULTS FOR MORE COMPLEX APPLICATIONS TO BE BUILT ON TOP OF THEM.

APPLICATIONS BASED ON MELODY EXTRACTION

The advances in algorithmic performance of melody extraction algorithms over the past decade mean they now provide sufficiently good results for more complex applications to be built on top of them. Below we provide a summary of some of these applications, whose wide range evidences the importance of melody extraction algorithms for MIR and computational music analysis.

RETRIEVAL

One of the most commercially attractive applications for melody extraction is music retrieval. That is, helping users find the music they are interested in or discover new music by means of automatically analyzing and comparing songs. Within this large application area we highlight two different yet related retrieval applications: version identification (version ID) and QBH. Version ID (also known as cover song ID) is the task of automatically retrieving different versions of a musical recording provided to the system by the user. Use cases range from the detection of copyright violations on Web sites such as YouTube, to automating the analysis of how musicians influence each other's compositions. Since the melody is often one of the few musical facets that remain unchanged across different renditions, various studies have explored the use of melody extraction for version ID, either by attempting to fully transcribe it [46], by using it as a midlevel representation for computing similarity [47], or by combining it with other tonal features (e.g., harmony, bass line, or the accompaniment as a whole) [8], [9].

The second retrieval task, QBH, is designed to help in the scenario where the user remembers the melody of a song but does not have any of its editorial information (e.g., title, album, or artist). QBH systems help the user retrieve this information by allowing them to sing or hum the melody as a search query. One important problem in the creation of QBH systems is the generation of a melody database (song index) against which the sung queries are to be compared. While it is possible to find MIDI versions of many songs on the Internet, such an approach will always be limited since it is not feasible to generate (i.e., transcribe) MIDI files manually for the very large music collections in existence today. Another solution is to match queries against other queries (i.e., user-recorded melodies), as performed by SoundHound (see "For More Information"). While this avoids the need for manual transcription, the approach still suffers from the same "cold start" problem—a song "does not exist" until a user records it. This problem can be alleviated by using melody extraction to automatically create a melody index for QBH systems. While no commercial QBH system based on melody extraction has been launched yet, research prototypes have shown promising results [9], [48], [49].

CLASSIFICATION

Automatic music classification attempts to help individual users as well as managers of large music corpora to organize their collections by automatically assigning descriptive labels to the songs in these collections. One of the most commonly used labels for organizing music is musical genre. The characteristics of the melody are often related to the musical genre (e.g., use of vibrato, pitch range), and could help in its identification. In [10], the authors present a genre classification system based on melody-related features obtained using melody extraction and demonstrate how combining these features with more commonly used timbre-related features such as Mel-frequency cepstral coefficients (MFCCs) can help to improve classification accuracy.

DE-SOLOING

Music de-soloing involves "removing" the lead instrument from a polyphonic music mixture. Doing this automatically is a highly attractive application for karaoke bars and fans—any song could automatically be converted into a karaoke accompaniment. Melody extraction can be used as a first step for de-soloing by providing a "score" of the melody that can be used to guide source separation algorithms in eliminating the melody from the audio mix [11].

TRANSCRIPTION

As we have already shown, a midlevel frequency-based representation of the melody is already very useful for various applications. However, sometimes transcribing all the way to symbolic notation (e.g., Western score notation) is desirable. For starters, music transcription is an attractive end goal in its own right, helping users learn music from automatically generated scores [5]. Automatic transcription can also help formalize the symbolic representation of orally transmitted music traditions, such as Flamenco [13]. Finally, by obtaining a symbolic representation of the melody we can apply the wide range of techniques that have been developed for symbolic melodic similarity and retrieval [4]. In all cases, the first step for obtaining a symbolic transcription of the melody from a polyphonic recording is by applying a melody extraction algorithm, whose output is then quantized in time and pitch to produce musical notes.

COMPUTATIONAL MUSIC ANALYSIS

As a final application, we discuss a couple of examples where melody extraction is useful for computational music analysis. Unlike the previous applications, whose goal was to enhance the way we find, represent, and interact with music, here our goal is to learn about the musical content itself by means of automated analysis. In [15], the authors combine melody extraction with a pattern recognition algorithm to detect the presence (or absence) of musical patterns that were predefined by musicologists. This type of analysis allows musicologists to study important aspects of the given musical style, e.g., to confirm existing musical hypotheses.

In [14], melody extraction is used for a different type of analysis. Here, melodies are extracted from excerpts of Indian classical music and summarized as pitch histograms with a high-frequency resolution. The resulting histograms are used for intonation analysis—an important aspect in Carnatic music (a type of Indian classical music). The intonation of a singer can be used to identify the raga of the piece, as well as characterize the musical expression of the performer.

CHALLENGES

While melody extraction algorithms have improved considerably since 2005, many challenges still remain. In the following sections we discuss some of the important issues, in terms of both algorithmic design and evaluation, that future research on melody extraction will have to address.

INSTRUMENTAL MUSIC AND HIGH DEGREES OF POLYPHONY

Earlier in our review, we mentioned that while most approaches can process instrumental music, many of them are particularly tailored for vocal music. We noted that this stems both from the popularity of vocal music, and from the uniqueness of the human voice which can be exploited by algorithms. However, if we wish to develop algorithms which generalize to a broader range of music material, melody extraction for instrumental music must be properly addressed. This presents two challenges compared with vocal melody extraction: first, instrumental music is not as constrained as vocal music. Instruments have a wider pitch range, can produce rapidly changing pitch sequences and include large jumps in pitch. Second, an instrument playing the melody may be closer, both in timbre and in the pitch contour of individual notes, to other accompanying instruments, which makes the task of distinguishing the melody from the accompaniment more complicated.

Regardless of the instrument playing the melody, the task becomes harder as we increase the number of instruments in the

mixture. This causes greater overlap of spectral content, making it harder to determine individual pitched sources correctly. Even when we manage to correctly distinguish the pitch values of different notes, determining which of these belong to the melody is now harder. Currently, algorithms are designed to handle material that is primarily homophonic, i.e., a single dominant lead instrument (or voice) with some harmonic accompaniment (strictly speaking, homophonic implies that the accompaniment shares the same rhythm as the melody, here we use the term more generally to refer to all music which has a lead melody with some form of harmonic accompaniment). Accurately extracting a specific melody from (for example) a fugue with a high degree of polyphony and several competing melodic lines is something current melody extraction algorithms can not do yet. Even in the simpler homophonic case we can think of challenging examples for melody extraction, for instance, songs that have backing vocals or even just a second voice. A second voice will usually move very similarly to the melody, reside in the same pitch range, and often be equally loud. This makes the task of determining which of the two voices is the actual melody highly challenging.

MELODY IS WITHOUT DOUBT A VERY IMPORTANT AND DISTINCT ASPECT OF MUSIC INFORMATION, AND SYSTEMS FOR AUTOMATICALLY EXTRACTING IT FROM MUSIC AUDIO ARE SURE TO BE CENTRAL TO FUTURE MUSIC INFORMATION TECHNOLOGIES.

VOICING DETECTION

When considering algorithmic performance, we saw that the key to obtaining high overall accuracy is the combination of high raw pitch accuracy with a good voicing detection method. To date, most approaches focus primarily on the former aspect of melody extraction, and less so on the latter (in Table 1 we see that some algorithms do not even include a voicing detection step). Often, voicing detection is only considered at the very end of the processing chain by applying a simple global energy threshold. Currently, even the algorithms with the most effective voicing detection methods obtain an average voicing false alarm rate (i.e., detecting melody where there isn't any) of more than 20%. In [16], the authors note that the most significant potential improvement in the performance of their algorithm would come from reducing the voicing false alarm rate, even though it is already one of the lowest in MIREX.

DEVELOPMENT CYCLE AND EVALUATION

In the section "Are We Improving?" we saw that for some MIREX collections performance has not improved significantly in recent years, and it was noted that this highlights a problem in the research and development cycle of melody extraction algorithms. Since the MIREX collections (with the exception of ADC2004) are kept secret for use in future evaluations, researchers have no way of analyzing the data to understand *why* their algorithms fail. Without listening to the audio content and examining the output of intermediate steps of the algorithm, the final results obtained, even if broken into several metrics, only tell you where and how you fail, but not why.

For algorithmic research and development, researchers use open data sets that are freely available. Since preparing a data set usually requires access to multitrack recordings and a considerable amount of manual annotation, there are very few such collections: the ADC2004 data set, the MIREX05 train data set, the MIR-1K data set, and the RWC pop data set (see "For More Information"). But the problem does not end here—the former two collections, while varied in terms of music material, are very small in size (20 and 13 excerpts, respectively), and the latter two, which are larger, are limited to a single musical genre (Chinese and Japanese pop, respectively). This means the collections are either too small to give statistically stable results, or too homogeneous to represent the universe of musical styles on which we would like our algorithms to work.

The current challenges in melody extraction evaluation are studied in detail in [50]. The authors focus on three aspects of evaluation in the MIREX campaign: ground truth generation, the duration of the excerpts used in test collections, and the size and content of the collections themselves. They first show how the lack of a common protocol for generating ground truth annotations could potentially lead to systematic errors in evaluation. By comparing algorithms' performance on excerpts with their performance on shorter subclips taken from the same excerpts, they also show that often short excerpts are not representative of the full song, implying that test collections should use complete songs rather than excerpts. Finally, they discuss the stability and representativeness of the results based on the size of the data sets, as we have already commented above. As the authors note, these findings do not invalidate the MIREX results, but rather emphasize the fact that we can not generalize them with confidence to significantly larger data sets of full songs. In an attempt to answer these problems, the Audio Melody Extraction Annotation Initiative (AMEAI) was launched in late 2012 (see "For More Information"). The goal of the initiative is to establish a common annotation protocol and compile a new, open data set for evaluation. The data set is planned to comprise full songs, large enough to provide statistically stable results and varied enough to represent a larger set of musical genres than those currently represented by existing evaluation collections.

SUMMARY AND CONCLUSIONS

In this article, we provided a review of melody extraction algorithms, considering not only aspects of algorithmic design and performance, but also the very definition of the task, its potential applications, and the challenges that still need to be solved. We started by considering the definition of melody and noted that to develop and evaluate melody extraction algorithms, we require a simplified and pragmatic definition. This was achieved by limiting the task to "single source predominant fundamental frequency estimation from musical content with a lead voice or

instrument." We described the challenges melody extraction entails from a signal processing point of view, and noted the differences between melody extraction, monophonic pitch estimation and multipitch estimation. By means of a case study, we highlighted some of the most common errors made by melody extraction algorithms and identified their possible causes. Next, we provided a comprehensive review of algorithmic design by considering 16 of the most relevant algorithms submitted to the MIREX evaluation campaign since 2005. We noted the great diversity of approaches and signal processing techniques applied, and identified two main algorithmic categories: salience-based methods and source separation-based methods. The evaluation measures most commonly used to assess melody extraction algorithms were described, and algorithmic performance was considered in terms of these measures. We saw that the best performing algorithms obtain a raw pitch accuracy between 70 and 80% and an overall accuracy of between 65 and 75%. We also saw that while performance has not improved much for some of the earlier collections, overall performance has improved gradually over the years.

Next, we provided a list of freely available melody extraction software, and considered some of the applications that have already been built on top of melody extraction algorithms, including: retrieval (version ID and QBH), genre classification, automatic de-soloing, music transcription, and computational music analysis. Finally, we considered some of the challenges that still need to be addressed by the research community. We noted that current algorithms are primarily designed to handle homophonic vocal music, and that in the future they will have to be extended to handle instrumental and highly polyphonic material. We highlighted the importance of voicing detection and noted the problem in the development cycle caused by the lack of open evaluation collections. We finally considered the evaluation process itself and noted that to be able to generalize the results obtained by melody extraction algorithms to larger music collections, we require new, larger and more heterogeneous test collections.

After nearly a decade of formal evaluations and many dozens of complete systems, it is fair to ask what we have learned about the best approaches to this problem. In our distinction between salience-based and source separation approaches, we find representatives of both among the best-performing systems according to the evaluations. One might argue that further progress in source separation (and full polyphonic transcription) will ultimately subsume this problem, but even despite the issue of greater computational expense, it remains an open question how best to model the perception and cognitive processing of the full music signal that goes on in the heads of listeners, who are not, we assume, performing a full analysis of the sound into individual sources when they listen to music. Notwithstanding the difficulties in obtaining a precise definition, melody is without doubt a very important and distinct aspect of music information, and systems for automatically extracting it from music audio are sure to be central to future music information technologies.

AUTHORS

Justin Salamon (justin.salamon@upf.edu) obtained a B.A. degree (Hons.) in computer science from the University of Cambridge, United Kingdom, in 2007. In 2008 he obtained his M.S. degree in cognitive systems and interactive media from Universitat Pompeu Fabra (UPF), Barcelona, Spain. Currently he is a researcher and Ph.D. student at the Music Technology Group (MTG), UPF. As part of his doctoral studies, he was a visiting researcher at the Sound Analysis-Synthesis research team of the Institut de Recherche et Coordination Acoustique/Musique (IRCAM), Paris, France. His main field of interest is MIR, with a focus on content-based MIR and audio and music processing, including musical stream estimation, melody and bass line extraction and characterization, QBH/example, classification, music, and melodic similarity and indexing. He is a Student Member of the IEEE.

Emilia Gómez (emilia.gomez@upf.edu) is a postdoctoral researcher and assistant professor (professor lector) at the MTG, Department of Information and Communication Technologies, UPF. She graduated as a telecommunication engineer specialized in signal processing at the Universidad de Sevilla. She received a DEA degree in acoustics, signal processing, and computer science applied to music at IRCAM, Paris. In 2006, she completed her Ph.D. degree in computer science and digital communication at UPF on the topic of tonal description of music audio signals. Her main research interests are related to melodic and tonal description of music audio signals, computer-assisted music analysis, and computational ethnomusicology. She is a Member of the IEEE.

Daniel P.W. Ellis (dpwe@ee.columbia.edu) received the Ph.D. degree in electrical engineering from the Massachusetts Institute of Technology (MIT). He is an associate professor in the Electrical Engineering Department, Columbia University, New York. His Laboratory for Recognition and Organization of Speech and Audio is concerned with all aspects of extracting high-level information from audio, including speech recognition, music description, and environmental sound processing. At MIT, he was a research assistant at the Media Lab, and he spent several years as a research scientist at the International Computer Science Institute, Berkeley, California, where he remains an external fellow. He is a Senior Member of the IEEE.

Gaël Richard (gael.richard@telecom-paristech.fr) received the State Engineering degree from Telecom ParisTech, France, (formerly ENST) in 1990, the Ph.D. degree from LIMSI-CNRS, University of Paris-XI, in 1994 in speech synthesis, and the Habilitation à Diriger des Recherches degree from the University of Paris XI in 2001. He spent two years at the Center for Advanced Information Processing, Rutgers University, Piscataway, New Jersey, and from 1997 to 2001, he successively worked for Matra, Bois d'Arcy, France, and Philips, Montrouge, France. He joined the Department of Signal and Image Processing, Telecom ParisTech, in 2001, where he is now a full professor in audio signal processing and head of the Audio, Acoustics, and Waves research group. He is a coauthor of over 100 papers and patents and is also one of the experts of the European commission in the field of audio signal processing and man/machine interfaces. He is a Senior Member of the IEEE.

REFERENCES

[1] M. Goto and S. Hayamizu, "A real-time music scene description system: Detecting melody and bass lines in audio signals," in *Working Notes of the IJCAI-99 Workshop on Computational Auditory Scene Analysis*, 1999, pp. 31–40.

[2] G. E. Poliner, D. P. W. Ellis, F. Ehmann, E. Gómez, S. Streich, and B. Ong, "Melody transcription from music audio: Approaches and evaluation," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 15, no. 4, pp. 1247–1256, 2007.

[3] J. S. Downie, "The music information retrieval evaluation exchange (2005–2007): A window into music information retrieval research," *Acoust. Sci. Technol.*, vol. 29, no. 4, pp. 247–255, 2008.

[4] R. Typke, "Music retrieval based on melodic similarity," Ph.D. dissertation, Dept. Inform. Computing Sci., Utrecht University, The Netherlands, 2007.

[5] M. Ryynänen and A. Klapuri, "Automatic transcription of melody, bass line, and chords in polyphonic music," *Comput. Music J.*, vol. 32, no. 3, pp. 72–86, 2008.

[6] A. L. Ringer. (2013, Jan.). Melody. Grove Music Online, Oxford Music Online. [Online]. Available: www.oxfordmusiconline.com/subscriber/article/grove/music/18357

[7] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, 5th ed. San Diego, CA: Academic Press, 2003.

[8] R. Foucard, J.-L. Durrieu, M. Lagrange, and G. Richard, "Multimodal similarity between musical streams for cover version detection," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 2010, pp. 5514–5517.

[9] J. Salamon, J. Serrà, and E. Gómez, "Tonal representations for music retrieval: From version identification to query-by-humming," *Int. J. Multimedia Inform. Retrieval* (*Special Issue on Hybrid Music Information Retrieval*), vol. 2, no. 1, pp. 45–58, Mar. 2013.

[10] J. Salamon, B. Rocha, and E. Gómez, "Musical genre classification using melody features extracted from polyphonic music signals," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 81–84.

[11] J.-L. Durrieu, G. Richard, and B. David, "An iterative approach to monaural musical mixture de-soloing," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 2009, pp. 105–108.

[12] A. Mesaros, T. Virtanen, and A. Klapuri, "Singer identification in polyphonic music using vocal separation and pattern recognition methods," in *Proc. 8th Int. Conf. Music Information Retrieval*, 2007, pp. 375–378.

[13] E. Gómez, F. Cañadas, J. Salamon, J. Bonada, P. Vera, and P. Cabañas, "Predominant fundamental frequency estimation vs. singing voice separation for the automatic transcription of accompanied flamenco singing," in *Proc. 13th Int. Society for Music Information Retrieval Conf.*, Porto, Portugal, Oct. 2012, pp. 601–606.

[14] G. K. Koduri, J. Serrà, and X. Serra, "Characterization of intonation in carnatic music by parametrizing pitch histograms," in *Proc. 13th Int. Society for Music Information Retrieval Conf.*, Porto, Portugal, Oct. 2012, pp. 199–204.

[15] A. Pikrakis, F. Gómez, S. Oramas, J. M. D. Báñez, J. Mora, F. Escobar, E. Gómez, and J. Salamon, "Tracking melodic patterns in flamenco singing by analyzing polyphonic music recordings," in *Proc. 13th Int. Society for Music Information Retrieval Conf.*, Porto, Portugal, Oct. 2012, pp. 421–426.

[16] J. Salamon and E. Gómez, "Melody extraction from polyphonic music signals using pitch contour characteristics," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 20, no. 6, pp. 1759–1770, Aug. 2012.

[17] J. Kreiman and D. V. L. Sidtis, *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Hoboken, NJ: Wiley-Blackwell, 2011, Ch. 10.

[18] W. Hess, *Pitch Determination of Speech Signals: Algorithms and Devices*. Berlin, Germany: Springer-Verlag, 1983.

[19] E. Gómez, A. Klapuri, and B. Meudic, "Melody description and extraction in the context of music content processing," *J. New Music Res.*, vol. 32, no. 1, pp. 23–40, 2003.

[20] A. Klapuri, "Signal processing methods for the automatic transcription of music," Ph.D. dissertation, Tampere University of Technology, Finland, Apr. 2004.

[21] G. Poliner and D. Ellis, "A classification approach to melody transcription," in *Proc. 6th Int. Conf. Music Information Retrieval*, London, Sept. 2005, pp. 161–166.

[22] J.-L. Durrieu, G. Richard, B. David, and C. Févotte, "Source/filter model for unsupervised main melody extraction from polyphonic audio signals," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 18, no. 3, pp. 564–575, Mar. 2010.

[23] H. Tachibana, T. Ono, N. Ono, and S. Sagayama, "Melody line estimation in homophonic music audio signals based on temporal-variability of melodic source," in *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2010, pp. 425–428.

[24] J. Sundberg, *The Science of the Singing Voice*. DeKalb, IL: Northern Illinois Univ. Press, 1987.

[25] R. P. Paiva, T. Mendes, and A. Cardoso, "Melody detection in polyphonic musical signals: Exploiting perceptual rules, note salience, and melodic smoothness," *Comput. Music J.*, vol. 30, pp. 80–98, Dec. 2006.

[26] M. Marolt, "On finding melodic lines in audio recordings," in *Proc. 7th Int. Conf. Digial Audio Effects (DAFx'04)*, Naples, Italy, Oct. 2004, pp. 217–221.

[27] M. Goto, "A real-time music-scene-description system: predominant-f0 estimation for detecting melody and bass lines in real-world audio signals," *Speech Commun.*, vol. 43, no. 4, pp. 311–329, Sept. 2004.

[28] P. Cancela, "Tracking melody in polyphonic audio," in *Proc. 4th Music Information Retrieval Evaluation eXchange (MIREX)*, 2008.

[29] K. Dressler, "An auditory streaming approach for melody extraction from polyphonic music," in *Proc. 12th Int. Society for Music Information Retrieval Conf.*, Miami, FL, Oct. 2011, pp. 19–24.

[30] V. Rao and P. Rao, "Vocal melody extraction in the presence of pitched accompaniment in polyphonic music," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 18, no. 8, pp. 2145–2154, Nov. 2010.

[31] S. Jo, S. Joo, and C. D. Yoo, "Melody pitch estimation based on range estimation and candidate extraction using harmonic structure model," in *Proc. InterSpeech*, Makuhari, Japan, Sept. 2010, pp. 2902–2905.

[32] V. Arora and L. Behera, "On-line melody extraction from polyphonic audio using harmonic cluster tracking," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 21, no. 3, pp. 520–530, Mar. 2013.

[33] C. Hsu and J. R. Jang, "Singing pitch extraction by voice vibrato/tremolo estimation and instrument partial deletion," in *Proc. 11th Int. Society for Music Information Retrieval Conf.*, Utrecht, The Netherlands, Aug. 2010, pp. 525–530.

[34] T.-C. Yeh, M.-J. Wu, J.-S. Jang, W.-L. Chang, and I.-B. Liao, "A hybrid approach to singing pitch extraction based on trend estimation and hidden Markov models," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 457–460.

[35] C. Sutton, "Transcription of vocal melodies in popular music," Master's thesis, Queen Mary, School of Electron. Eng. Comput. Sci., University of London, United Kingdom, 2006.

[36] K. Dressler, "Sinusoidal extraction using an efficient implementation of a multi-resolution FFT," in *Proc. 9th Int. Conf. Digital Audio Effects (DAFx-06)*, Montreal, Canada, Sept. 2006, pp. 247–252.

[37] K. Dressler, "Pitch estimation by the pair-wise evaluation of spectral peaks," in *Proc. AES 42nd Int. Conf.*, Ilmenau, Germany, July 2011, pp. 278–290.

[38] A. Ozerov, P. Philippe, F. Bimbot, and R. Gribonval, "Adaptation of Bayesian models for single-channel source separation and its application to voice/music separation in popular songs," *IEEE Trans. Audio, Speech, and Lang. Processing*, vol. 15, no. 5, pp. 1564–1578, July 2007.

[39] Y. Li and D. Wang, "Separation of singing voice from music accompaniment for monaural recordings," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 15, no. 4, pp. 1475–1487, May 2007.

[40] J.-L. Durrieu, B. David, and G. Richard, "A musically motivated mid-level representation for pitch estimation and musical audio source separation," *IEEE J. Select. Topics Signal Processing*, vol. 5, no. 6, pp. 1180–1191, Oct. 2011.

[41] P.-S. Huang, S. D. Chen, P. Smaragdis, and M. Hasegawa-Johnson, "Singing voice separation from monaural recordings using robust principal component analysis," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 57–60.

[42] Z. Rafii and B. Pardo, "Repeating pattern extraction technique (REPET): A simple method for music/voice separation," *IEEE Trans. Audio, Speech, Lang. Processing*, vol. 21, no. 1, pp. 71–82, Jan. 2013.

[43] A. Liutkus, Z. Rafii, R. Badeau, B. Pardo, and G. Richard, "Adaptive filtering for music/voice separation exploiting the repeating musical structure," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 53–56.

[44] N. Ono, K. Miyamoto, H. Kameoka, J. Le Roux, Y. Uchiyama, E. Tsunoo, T. Nishimoto, and S. Sagayama, "Harmonic and percussive sound separation and its application to MIR-related tasks," in *Advances in Music Information Retrieval* (Studies in Computational Intelligence, vol. 274), Z. Ras and A. Wieczorkowska, Eds. Berlin, Germany: Springer, 2010, pp. 213–236.

[45] P. Cancela, E. López, and M. Rocamora, "Fan chirp transform for music representation," in *Proc. 13th Int. Conf. Digital Audio Effects (DAFx)*, Graz, Austria, Sept. 2010, pp. 54–61.

[46] W.-H. Tsai, H.-M. Yu, and H.-M. Wang, "Using the similarity of main melodies to identify cover versions of popular songs for music document retrieval," *J. Inform. Sci. Eng.*, vol. 24, no. 6, pp. 1669–1687, 2008.

[47] M. Marolt, "A mid-level representation for melody-based retrieval in audio collections," *IEEE Trans. Multimedia*, vol. 10, no. 8, pp. 1617–1625, Dec. 2008.

[48] J. Song, S. Y. Bae, and K. Yoon, "Mid-level music melody representation of polyphonic audio for query-by-humming system," in *Proc. 3rd Int. Conf. Music Information Retrieval*, Paris, France, Oct. 2002, pp. 133–139.

[49] M. Ryynänen and A. Klapuri, "Query by humming of MIDI and audio using locality sensitive hashing," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 2008, pp. 2249–2252.

[50] J. Salamon and J. Urbano, "Current challenges in the evaluation of predominant melody extraction algorithms," in *Proc. 13th Int. Society for Music Information Retrieval Conf.*, Porto, Portugal, Oct. 2012, pp. 289–294.

[SP]

[sp TIPS&TRICKS]

Chun-Su Park and Sung-Jea Ko

The Hopping Discrete Fourier Transform

The discrete Fourier transform (DFT) produces a Fourier representation for finite-duration data sequences. In addition to its theoretical importance, the DFT plays a key role in the implementation of a variety of digital signal-processing algorithms. Several algorithms including the fast Fourier transform (FFT) and the Goertzel algorithm have been introduced for the fast implementation of the DFT [1], [2].

Recently, there has been a growing interest in a sliding DFT process where the transform window is shifted one sample at a time and the transform process is repeated. Basically, a sliding transform algorithm computes the DFT bins on a sample-by-sample basis. This means that, even if the DFT output needs to be calculated only every L $(L > 0)$ samples, the algorithm has to perform the DFT at all time indices.

This article introduces a new hopping DFT (HDFT) algorithm that can adjust the time hop L between successive DFT outputs. The new algorithm computes the DFT output at time index n using the precalculated one at $(n-L)$. We will show that the computational requirement of the proposed HDFT algorithm is the lowest among existing fast sliding DFT (SDFT) algorithms.

EXISTING SLIDING DFT ALGORITHMS

In the sliding transform, the transform is computed on a fixed-length window of the signal, which is continuously updated with new samples as the oldest ones are discarded [3]. Consider a complex input signal $x(n)$, $n = 0, 1, 2, \ldots$, which will be divided into overlapping windows of size M. In [4], the sliding FFT (SFFT) algorithm was introduced, which requires M complex multiplications and $2M$ complex additions for computing all bins of the shifted window. The SFFT algorithm calculates the bins of the shifted window by exploiting delayed intermediate calculations of the previous window. In [5], the generalized SFFT (GSFFT) algorithm was proposed for an efficient implementation of the hopping FFT. When the DFT output needs to be calculated only every L samples, the GSFFT algorithm directly computes the bins at time index n by exploiting the delayed intermediate calculations at time index $(n-L)$. The GSFFT algorithm requires $(M(\log_2 L + 2)/2 - L)$ complex multiplications and $(M(\log_2 L + 2) - 2L)$ complex additions for computing all bins of the shifted window. The SFFT and GSFFT algorithms require additional memory to store the intermediate calculations of the previous window.

Recently, the SDFT algorithm exploiting the circular shift property was proposed in [6] and [7]. Let k be a frequency-domain index in the range $0 \le k < M$. Then, at time index n, the k th bin of an M-point DFT is computed as

$$X_n(k) = \sum_{m=0}^{M-1} x(\hat{n}+m) W_M^{-\mathrm{km}}, \qquad (1)$$

where $\hat{n} = n - M + 1$ and $W_M = e^{j2\pi/M}$. Accordingly, as shown in [7], the SDFT is formulated as follows:

$$\begin{aligned} X_n(k) &= \sum_{m=0}^{M-1} x(\hat{n}+m) W_M^{-\mathrm{km}} \\ &= \sum_{m=0}^{M-1} x(\hat{n}+m-1) W_M^{-k(m-1)} \\ &\quad + x(\hat{n}+M-1) W_M^{-k(M-1)} \\ &\quad - x(\hat{n}-1) W_M^{k} \\ &= W_M^k \sum_{m=0}^{M-1} x(\hat{n}+m-1) W_M^{-\mathrm{km}} \\ &\quad + x(\hat{n}+M-1) W_M^k \\ &\quad - x(\hat{n}-1) W_M^k \\ &= W_M^k (X_{n-1}(k) + x(n) \\ &\quad - x(n-M)), \end{aligned} \qquad (2)$$

where the periodicity property of the complex twiddle factor is exploited $(W_M^{k+M} = W_M^k)$. The SDFT shown in (2) requires one complex multiplication and two complex additions for computing an output bin of the shifted window.

As mentioned in [6], even when the DFT outputs need to be computed only every L input samples, the SDFT should compute the output for each input sample due to its recursive nature. The SDFT is computationally superior to the traditional butterfly-based FFT only when L is less than $(\log_2 M)/2$ [1]. Now, let us explain the proposed HDFT.

PROPOSED HOPPING DFT ALGORITHM FOR TIME HOP 2 (L = 2)

We first illustrate the proposed HDFT algorithm for $L = 2$. In the next section, the proposed algorithm for arbitrary time hops is introduced. The relationship between $X_n(k)$ and $X_{n-2}(k)$ can be easily derived by substituting $X_n(k)$ into $X_{n-1}(k)$ in (2). The resultant formula relating $X_n(k)$ and $X_{n-2}(k)$ is given by

$$\begin{aligned} X_n(k) &= W_M^k \{W_M^k [X_{n-2}(k) \\ &\quad + x(n-1) - x(n-M-1)] \\ &\quad + x(n) - x(n-M)\} \\ &= W_M^{2k} \{X_{n-2}(k) + [d(n-1) \\ &\quad + W_M^k d(n)]\}, \end{aligned} \qquad (3)$$

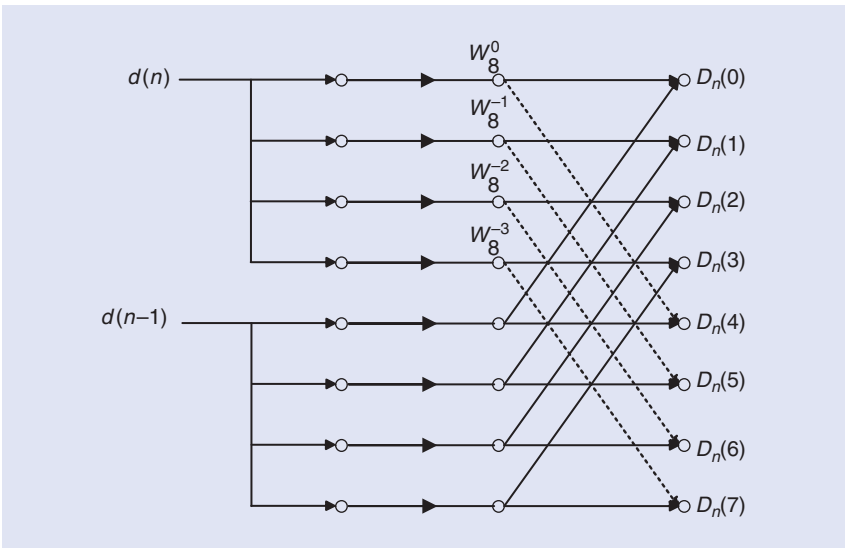
where $d(n) = x(n) - x(n-M)$. Let $D_n(k) = d(n-1) + W_M^k d(n)$ and, then, (3) can be rewritten as

$$X_n(k) = W_M^{2k} (X_{n-2}(k) + D_n(k)). \qquad (4)$$

The above equation indicates that the DFT outputs at time index n can be directly computed from those at time index $(n-2)$ by exploiting $D_n(k)$. Based on these observations, we introduce an algorithm to compute $D_n(k)$ with less computational load. Figure 1 describes how to efficiently calculate

Digital Object Identifier 10.1109/MSP.2013.2292891
Date of publication: 12 February 2014

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

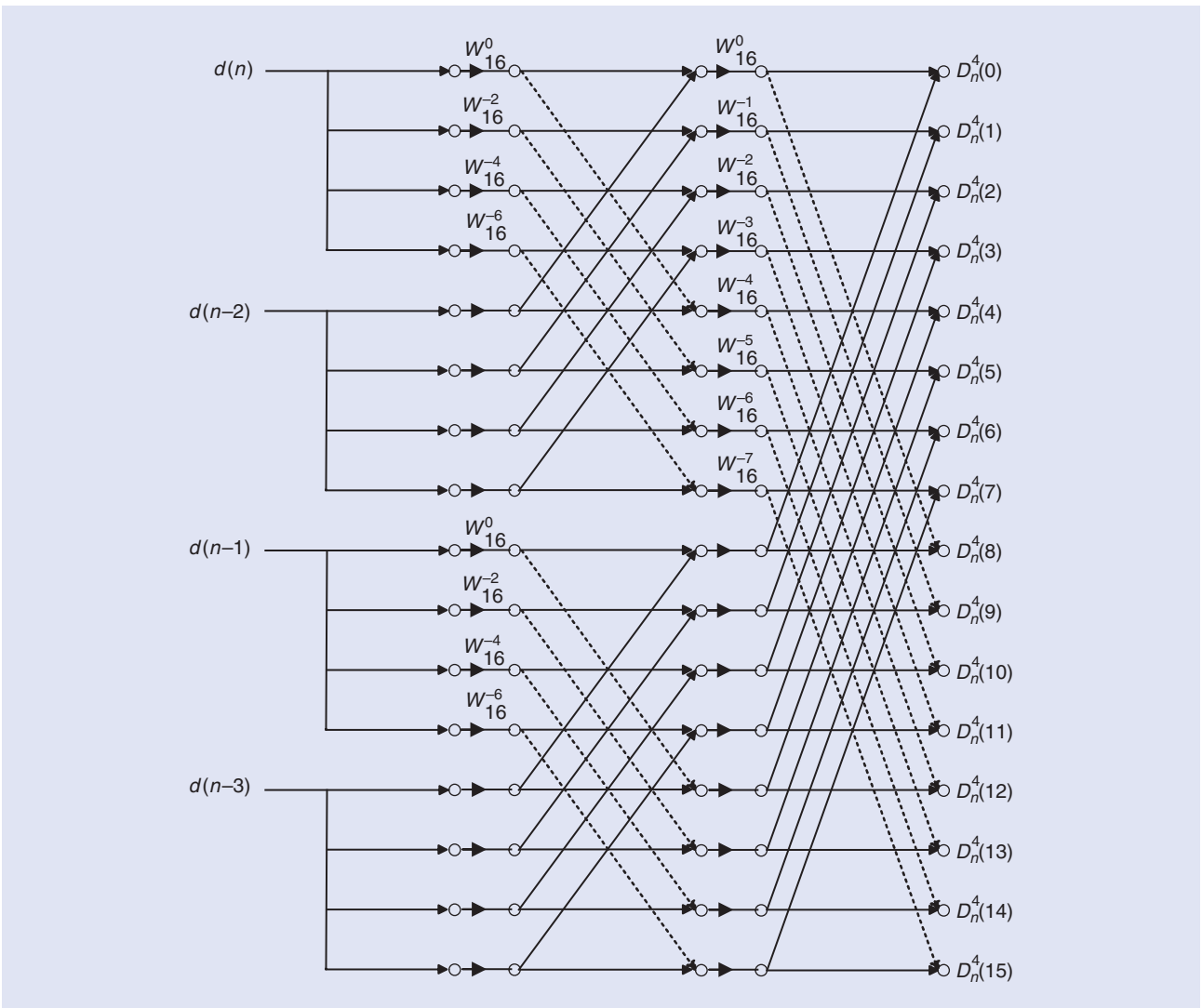


[FIG 1] The calculation of $D_n(k)$, for $M = 8$ and $L = 2$, where solid and dotted lines indicate the plus and minus signs, respectively.

$D_n(k)$ using $d(n)$ and $d(n-1)$, where the symmetry property of the twiddle factor is exploited ($W_M^{k+M/2} = -W_M^k$).

The computational requirement of the HDFT algorithm for $L = 2$ can be summarized as follows. At first, two additions are required for computing $d(n)$ and $d(n-1)$. Further, as shown in Figure 1, $M/2$ complex multiplications and M complex additions are required for the calculation of $D_n(k)$. It can be seen from (4) that, after obtaining $D_n(k)$, only one complex multiplication and one complex addition are needed per output bin.

In total, using the precalculated DFT outputs at time index $(n-2)$, the proposed algorithm computes all outputs at time index n with $3M/2$ complex multiplications and $(2M+2)$ complex



[FIG 2] The UVT calculation for $M = 16$ and $L = 4$, where solid and dotted lines indicate the plus and minus signs, respectively.

additions, whereas the SDFT algorithm in [7] requires $2M$ complex multiplications and $4M$ complex additions. Thus, the proposed algorithm reduces the numbers of complex multiplications and additions by about 25% and 50%, respectively.

PROPOSED HOPPING DFT ALGORITHM FOR TIME HOPS OF ARBITRARY LENGTH

This section describes the proposed HDFT algorithm for arbitrary time hops. For simplicity of explanation, we only consider the case $L = 2^a,\ a \geq 0$. The other cases can be easily generalized. Similar to the case $L = 2$, the relationship between $X_n(k)$ and $X_{n-L}(k)$ is derived by recursively substituting $X_n(k)$ into $X_{n-1}(k)$ L times in (2). The resultant formula is given by

$$X_n(k) = W_M^{Lk}\{X_{n-L}(k) + [d(n-L+1) + W_M^{-k} d(n-L+2) + \cdots + W_M^{-(L-1)k} d(n)]\}. \quad (5)$$

Let us define $D_n^L(k)$ as the kth bin of the L-point updating vector transform (UVT), which is represented by

$$D_n^L(k) = \sum_{m=0}^{L-1} d(n-m)\, W_M^{(m-L+1)k}, \quad (6)$$

where $0 \leq k < M$. Then, (5) is simplified as

$$X_n(k) = W_M^{Lk}(X_{n-L}(k) + D_n^L(k)). \quad (7)$$

This leads to the result that the DFT outputs at time index n can be directly computed from those at time index $(n-L)$ by exploiting $D_n^L(k)$.

We now focus our attention on the fast implementation of the UVT shown in (6). There are several FFT algorithms for computing the DFT efficiently. Since the computation of the UVT is similar to that of the DFT, we can apply the traditional FFT algorithms to the computation of $D_n^L(k)$. A radix-2 decimation-in-time (DIT) FFT is the simplest and most common form of the FFT algorithms, which divides a DFT of size M into two interleaved DFTs of size $M/2$ at each recursive stage. Based on the DIT approach, we express $D_n^L(k)$ using decimated sequences as follows:

$$\begin{aligned} D_n^L(k) &= \sum_{m=0}^{L-1} d(n-m)\, W_M^{(m-L+1)k} \\ &= \sum_{p=0}^{L/2-1} d(n-2p)\, W_M^{(2p-L+1)k} \\ &\quad + \sum_{p=0}^{L/2-1} d(n-2p-1)\, W_M^{(2p-L+2)k}. \end{aligned} \quad (8)$$

By substituting W_M^2 with $W_{M/2}$, the equation can be rewritten as

$$\begin{aligned} D_n^L(k) &= W_M^{-k} \sum_{p=0}^{L/2-1} d(n-2p)\, W_{M/2}^{(p-L/2+1)k} \\ &\quad + \sum_{p=0}^{L/2-1} d(n-2p-1)\, W_{M/2}^{(p-L/2+1)k} \\ &= W_M^{-k} \dot{D}_n^{L/2}(k) + \ddot{D}_n^{L/2}(k), \end{aligned} \quad (9)$$

where

$$\dot{D}_n^{L/2}(k) = \sum_{p=0}^{L/2-1} d(n-2p)\, W_{M/2}^{(p-L/2+1)k}$$

and

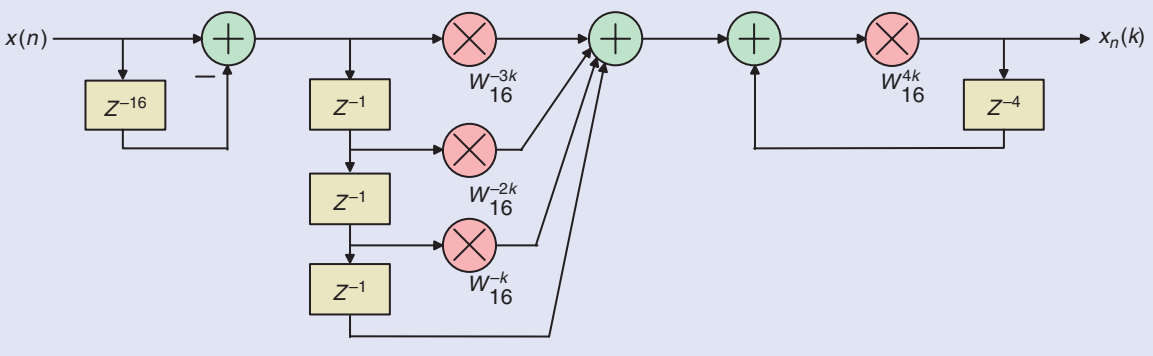
$$\ddot{D}_n^{L/2}(k) = \sum_{p=0}^{L/2-1} d(n-2p-1)\, W_{M/2}^{(p-L/2+1)k}.$$

This equation implies that, instead of directly computing $D_n^L(k)$, we can obtain $D_n^L(k)$ using two UVT bins of the decimated sequences, $\{d(n), d(n-2), \ldots, d(n-L+2)\}$ and $\{d(n-1), d(n-3), \ldots, d(n-L+1)\}$. The decimation process is repeated until the resulting sequences are reduced to one-point sequences. For example, if $L = 2^a$, the decimation is performed $a = \log_2 L$ times. Figure 2 shows an example that computes $D_n^L(k)$ using the DIT approach. Note that, since the twiddle factors used in the UVT are identical to those used in the traditional butterfly-based FFT algorithm, the proposed HDFT algorithm has the same precision as the traditional FFT [1].

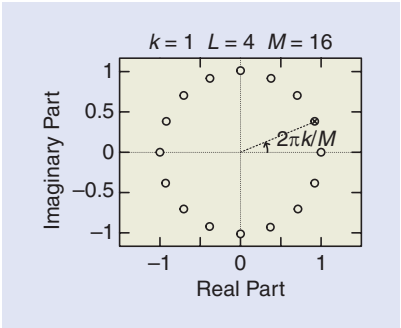
The UVT is performed using only the input signal samples, i.e., $D_n^L(k)$ can be precalculated independent of $X_{n-L}(k)$ in (7). This may have a significant effect on run times in favor of the HDFT algorithm. For example, the separable feature of the HDFT is highly beneficial for

[TABLE 1] THE PSEUDOCODE OF THE PROPOSED HDFT ALGORITHM.

```
INPUT: INPUT SAMPLES x(n) AND DFT OUTPUT X_{n-L} AT TIME INDEX (n − L)
OUTPUT: DFT OUTPUT X_n AT TIME INDEX N

/* STEP A): C_M = 0, C_A = L * /
  FOR i = 0, 1, ..., L − 1 DO
    d(n − i) = x(n − i) − x(n − i − M)
  END FOR

/* STEP B): C_M = (M/2) log2 L, C_A = M · log2 L * /
  FOR i = 1, 2, ..., log2 L DO
    COMPUTE THE ORDER-2^i UVT D_n^{2^i}(k) [SEE (9)]
  END FOR

/* STEP C): C_M = 0 OR M, C_A = M * /
  FOR k = 0, 1, ..., M − 1 DO
    COMPUTE X_n(k) USING X_{n−L}(k) AND D_n^L(k) [SEE (7)]
  END FOR
RETURN X_n
```



[FIG 3] The single-bin HDFT structure for $M = 16$ and $L = 4$.



[FIG 4] The z-plain pole/zero locations for $k = 1$, $L = 4$, and $M = 16$.

hardware implementation and multicore processors which are widely used in many signal processing applications.

At time index n, the overall algorithm for window size M proceeds as follows:

a) Compute $\{d(n), d(n-1), \ldots, d(n-L+1)\}$ using the input samples. Memory of size L is required to store the outputs. Only L complex additions are needed in this step.

b) Compute the order-L UVT $D_n^L(k)$ using the DIT approach in (9). The resultant $D_n^L(k)$ will be an input of the following step. Note that, since the decimation is performed $\log_2 L$ times, the computation of $D_n^L(k)$ requires $(M/2)\log_2 L$ complex multiplications and $M \log_2 L$ complex additions. This step requires size M memory for storing the UVT outputs.

c) According to (7), calculate $X_n(k)$ using $X_{n-L}(k)$ and $D_n^L(k)$. Since this step requires the DFT outputs of the previous window to be maintained in memory, this step requires additional memory of size M. The computational requirement of this step can be minimized for the special twiddle factors $(W_M^0 = -W_M^{M/2} = 1;\ W_M^{M/4} = -W_M^{3M/4} = j)$. For example, if $L = M/4$ or $L = M/2$, this step needs only M complex additions. Otherwise, M complex multiplications and M complex additions are needed in this step.

A pseudocode of the proposed HDFT algorithm is given in Table 1 where C_M and C_A indicate the numbers of complex multiplications and additions, respectively. Then, C_M and C_A of the overall HDFT process are summarized as

$$C_M = \begin{cases} (M/2)\log_2 L, & \text{if } L = M/4 \text{ or } M/2, \\ M(1 + (\log_2 L)/2), & \text{otherwise,} \end{cases} \tag{10}$$

and

$$C_A = M(1 + \log_2 L) + L. \tag{11}$$

A computational comparison with varying M and L is presented in Table 2. In particular, when L is equal to $M/4$, the HDFT algorithm reduces the number of complex multiplications by as much as 81.25% as compared to the SFFT and SDFT.

Figure 3 shows the single-bin HDFT structure for $M = 16$ and $L = 4$. It can be observed from Figure 3 that the z-domain transfer function for the kth bin of the single-bin HDFT filter is

$$\begin{aligned} H_{\text{HDFT}}(z) &= \frac{W_M^k \sum_{m=0}^{L-1} W_M^{mk} (z^{-m} - z^{-(m+M)})}{1 - W_M^{Lk} z^{-L}} \\ &= \frac{W_M^k (1 - z^{-M}) \sum_{m=0}^{L-1} W_M^k z^{-1})^m}{1 - W_M^{Lk} z^{-L}}. \end{aligned} \tag{12}$$

Figure 4 shows an example for the pole/zero locations of the single-bin HDFT filter. In Figure 4, a single pole is located on the unit circle and the number of zeroes varies depending on M. The single-bin structure can be used for applications which do not require to compute all DFT bins. If all DFT bins should be computed, we would suggest the use of the UVT-based implementation shown in Table 1.

In summary, the proposed HDFT algorithm has several advantages:

- The computational requirement of the proposed HDFT is the lowest among existing sliding transform algorithms.
- The UVT can be precalculated independent of the recursive calculation. This separable feature is highly beneficial for real-world applications such as hardware implementation and multicore processors.
- Due to the reduced number of complex multiplications using the imprecise twiddle factors, the accumulated error in the sliding transform process can be reduced drastically. For

[TABLE 2] COMPUTATIONAL WORKLOAD FOR A COMPLEX INPUT SIGNAL.

		FFT		SFFT		GSFFT		SDFT		HDFT	
M	L	C_M	C_A	C_M	C_A	C_M	C_A	C_M	C_A	C_M	C_A
16	1	32	64	16	32	16	32	16	32	16	17
	2	32	64	32	64	22	44	32	64	24	34
	4	32	64	64	128	28	56	64	128	16	52
	8	32	64	128	256	32	64	128	256	24	72
32	1	80	160	32	64	32	64	32	64	32	33
	2	80	160	64	128	46	92	64	128	48	66
	4	80	160	128	256	60	120	128	256	64	100
	8	80	160	256	512	72	144	256	512	48	136
	16	80	160	512	1,024	80	160	512	1,024	64	176

[TABLE 3] MEASURED PROCESSING TIME WITH VARYING L.

		FFT	SFFT		GSFFT		SDFT		HDFT	
M	L	(ms)	(ms)	(%)	(ms)	(%)	(ms)	(%)	(ms)	(%)
16	1	73.17	37.79	51.64	37.79	51.64	30.02	41.03	27.43	37.49
	2	36.13	37.79	104.58	26.42	73.12	30.02	83.09	24.18	66.93
	4	18.09	37.79	208.89	17.14	94.76	30.02	165.96	13.14	72.66
	8	9.03	37.79	418.29	9.03	100.00	30.02	332.33	8.74	96.71
32	1	192.63	74.35	38.60	74.35	38.60	64.54	33.50	52.55	27.28
	2	96.08	74.35	77.38	52.91	55.07	64.54	67.17	47.86	49.82
	4	48.80	74.35	152.34	36.98	75.77	64.54	132.24	28.80	59.01
	8	24.22	74.35	306.97	22.28	91.99	64.54	266.46	17.83	73.61
	16	12.17	74.35	611.08	12.17	100.00	64.54	530.44	11.77	96.72

example, in the SDFT, L complex multiplications are sequentially performed for computing $X_n(k)$ from $X_{n-L}(k)$. However, the proposed HDFT algorithm needs only a single complex multiplication regardless of time hop L.

SIMULATION RESULTS

We investigated the computational efficiency of the proposed HDFT algorithm using a complex test signal which was zero-mean Gaussian noise with a standard deviation equal to one. The length of the test signal was 10^6 and M was set to 16 and 32. The SFFT [4], GSFFT [5], SDFT [7], and HDFT algorithms were compared to the FFT [1] with varying L. All algorithms were implemented using a highly efficient ANSI-C code and the performance was evaluated on an Intel i5 3.4 GHz CPU with 8 GB RAM. The simulation was performed in 64-bit double-precision arithmetic.

Table 3 shows the measured processing time of all algorithms. When $L = 1$, the SFFT, GSFFT, SDFT, and HDFT significantly reduce the processing time as compared to the FFT. As mentioned, the SFFT and GSFFT algorithms require additional memory to store intermediate calculations of the previous window, which introduces additional memory accesses as compared to the SDFT and HDFT. In the case of $L = 1$, although the computational requirements of the SFFT and GSFFT are the same as that of the SDFT, their processing times are longer than that of the SDFT.

As L increases, the performance gains of all algorithms decrease. When $L = M/4$, only the GSFFT and HDFT algorithms can reduce the processing time as compare to the FFT. For example, when $M = 32$ and $L = 8$, the time savings of the GSFFT and HDFT are 8.01% and 26.39%, respectively. Further, in the case of $L = M/2$, the HDFT is the only algorithm that can reduce the complexity as compare to the FFT. Table 3 shows that the HDFT consistently outperforms the other algorithms for all values of L.

SUMMARY

A new hopping DFT algorithm with reduced computational complexity was presented in this article. Theoretical analysis showed that the proposed algorithm provides computational advantages over conventional algorithms including the FFT, SFFT, GSFFT, and SDFT. The experimental results clearly demonstrated that the proposed method consistently achieves a substantially higher acceleration than the other algorithms. The proposed HDFT algorithm can be utilized for a variety of signal processing applications requiring successive DFT output calculations.

ACKNOWLEDGMENTS

The author would like to thank Rick Lyons for insightful suggestions for improving the manuscript and the anonymous reviewers for their valuable comments. This research was supported by a research grant from Sangmyung University.

AUTHORS

Chun-Su Park (cspark@smu.ac.kr) is an assistant professor in the Department of Electrical Engineering at Sangmyung University in Cheonan, South Korea.

Sung-Jea Ko (sjko@korea.ac.kr) is a professor in the Department of Electrical Engineering at Korea University in Seoul, South Korea.

REFERENCES

[1] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Englewood-Cliffs, NJ: Prentice Hall, 1999, p. 635.

[2] K. Banks, "The Goertzel algorithm," *Embedded Syst. Programming Mag.*, vol. 15, no. 9, pp. 34–42, Sept. 2002.

[3] J. Wu, H. Shu, L. Wang, and L. Senhadji, "Fast algorithms for the computation of sliding sequency-ordered complex Hadamard transform," *IEEE Trans. Signal Processing*, vol. 58, no. 11, pp. 5901–5909, Nov. 2010.

[4] B. Farhang-Boroujeny and Y. Lim, "A comment on the computational complexity of sliding FFT," *IEEE Trans. Circuits Syst. II*, vol. 39, no. 12, pp. 875–876, Dec. 1992.

[5] B. Farhang-Boroujeny and S. Gazor, "Generalized sliding FFT and its application to implementation of block LMS adaptive filters," *IEEE Trans. Signal Processing*, vol. 42, no. 3, pp. 532–538, Mar. 1994.

[6] E. Jacobsen and R. Lyons, "The sliding DFT," *IEEE Signal Processing Mag.*, vol. 20, no. 2, pp. 74–80, Mar. 2003.

[7] E. Jacobsen and R. Lyons, "An update to the sliding DFT," *IEEE Signal Processing Mag.*, vol. 21, no. 1, pp. 110–111, Jan. 2004.

[8] K. Duda, "Accurate, guaranteed stable, sliding discrete Fourier transform," *IEEE Signal Processing Mag.*, vol. 27, no. 6, pp. 124–127, Nov. 2010.

[SP]

moving?

You don't want to miss any issue of this magazine!

change your address

BY E-MAIL: address-change@ieee.org

BY PHONE: +1 800 678 IEEE (4333) in the U.S.A.
or +1 732 981 0060 outside the U.S.A.

ONLINE: www.ieee.org, click on quick links, change contact info

BY FAX: +1 732 562 5445

Be sure to have your member number available.

[lecture NOTES]

Manuel D. Ortigueira

On the Recursive Solution to the Normal Equations: Some Other Results

The recursive solution of the autocorrelation normal equations was recovered in a recent article by Baran and Oppenheim [1]. The algorithm they obtained is slightly different from the one currently described in the literature [9]. However, they did not reallize the importance of the difference. I discovered it almost 30 years ago [4], [6] and used it to propose a modified Burg method as alternative to the well-known Burg method [2] that emerged directly from the classic Durbin-Levinson algorithm. In fact, the solution of the normal equations Baran, Oppenheim, and I obtained leads to predictors that are different from those acquired with the classic Durbin-Levinson algorithm. We will see the difference by looking into the recursions verified by the corresponding prediction error filters.

RELEVANCE

The autocorrelation normal equations constitute the most important step in many signal processing algorithms, mainly in autoregressive (AR) and autoregressive-moving average (ARMA) modeling. In general, the autocorrelation matrix is a Toeplitz matrix, but in the AR case it is also a symmetric (Hermitian if it is complex) matrix. An efficient solution to the normal equations is very important [2], [9], [10] for reducing the computational burden, the numerical errors, and the time spent.

PREREQUISITES

This lecture note assumes the basic knowledge of the discrete-time linear systems theory and Z-transform. In particular the concepts of the moving average, AR, and ARMA filters and consequently the relations with the finite impulse response and infinite impulse response filters.

Digital Object Identifier 10.1109/MSP.2013.2292431
Date of publication: 12 February 2014

PROBLEM STATEMENT AND CLASSIC RESULTS

The AR modeling and the linear prediction are closely related. It is well known that the prediction error filter coefficients are the solution of the normal equations that we can write as

$$\sum_{i=0}^{N} a_i R(j-i) = P_N \delta_j,\ j = 0, \cdots, N, \quad (1)$$

where $R(n)$ is the autocorrelation function of the process x_n, P_N is the Nth-order prediction error power, and δ_n is the Kronecker symbol. The above relation can be written in a matrix format and normally solved recursively by means of the Levinson-Durbin recursion [2].

In the classic formulation the Nth-order forward and backward prediction error filters needed to construct the solution are defined by

$$A^N(z) = \sum_{0}^{N} a_i z^{-i} \text{ and } B^N(z) = \sum_{0}^{N} a_{N-i} z^{-i}, \quad (2)$$

where $a_i, i = 0, 1, \ldots, N$ are the AR (predictor) parameters and verify the following recursions:

$$A'^N(z) = A'^{N-1}(z) + C_N z^{-1} B'^{N-1}(z) \quad (3)$$
$$B'^N(z) = z^{-1} B'^{N-1}(z) + C_N A'^{N-1}(z), \quad (4)$$

where C_N is the reflection coefficient (PARCOR).

Let $x_n, n \in \mathbb{Z}$ be a stationary stochastic process. The forward, f'^N_n, and backward, b'^N_n, prediction errors obtained with the above filters $(f'^0_n = b'^0_n = x_n)$ verify the recursions

$$f^N_n = f^{N-1}_n + C_N b^{N-1}_{n-1} \quad (5)$$
$$b^N = b^{N-1}_{n-1} + C_N f^{N-1}_n \quad (6)$$

that are the basis for the Burg method and several lattice algorithms introduced in the 1970s and 1980s [9].

If one represents each filter by an arrow, we can illustrate the action of each filter by the following scheme:

$$\begin{array}{c} A^N(z) \\ \xrightarrow{\qquad} \\ x_{n-N} x_{n-N+1} \cdots x_{n-2} x_{n-1} x_n \\ \xleftarrow{\qquad} \\ \mathrm{B}^N(z) \end{array}.$$

As can be seen, both filters run over the same datum points and share them simultaneously. This does not happen in the algorithm described next.

ALTERNATIVE SOLUTION

A slightly different solution for the above system of normal equations (1) was presented in [6] where the bordering algorithm was used. The same result was found recently by Baran and Oppenheim [1]. This procedure has an interesting interpretation in terms of the action of the prediction error filter over the autocorrelation function and the involved entropy [10] leading to an interesting formulation. It will be briefly described next.

Let us denote the function defined by the left-hand side in (1) for the (N–1)th order by $u^{N-1}(n)$

$$u^{N-1}(n) = \sum_{i=0}^{N-1} a_i R(n-i). \quad (7)$$

This function is defined for all $n \in \mathbb{Z}$ and has $N-1$ gaps (points where the function is zero) for $n = 1, \ldots, N-1$. The function $u^{N-1}(-n)$ has also $N-1$ gaps for $n = -1, \ldots, -N+1$. It can be shifted to the right to make its gaps coincide with those of $u^{N-1}(n)$. We call this new function $v^{N-1}(n) = u^{N-1}(N-n)$. This means that the function $v^{N-1}(n)$ has also $N-1$ gaps for $n = 1, \ldots, N-1$. So any linear combination of both functions

1053-5888/14/$31.00©2014IEEE

$$u^N(n) = u^{N-1}(n) + C_N v^{N-1}(n) \quad (8)$$

will have also $N-1$ gaps but for $n = 1, \ldots, N-1$. To compute the coefficient C_N, we force the Nth-order function $u^N(n)$ to have a gap at $n = N$. This leads immediately to

$$C_N = -\frac{u^{N-1}(N)}{v^{N-1}(N)} = -\frac{u^{N-1}(N)}{u^{N-1}(0)} \quad (9)$$

$$v^N(n) = v^{N-1}(n-1) + C_N u^{N-1}(n-1) \quad (10)$$

$$u^N(0) = u^{N-1}(0)(1 - C_n^2) \quad (11)$$

that, together with (8), allow us to obtain recursively the set of reflection coefficients by starting with $u^0(n) = v^0(n) = R(n)$. We can summarize the algorithm in the following steps:

1) Let $u^0(n) = v^0(n) = R(n)$
2) For $N = 1, \ldots, N_0$ compute
3) $C_N = -(u^{N-1}(N))/(u^{N-1}(0))$
   $u^N(n) = u^{N-1}(n) + C_N v^{N-1}(n)$
   $v^N(n) = v^{N-1}(n-1) + C_N u^{N-1}(n-1)$. If $P_x = R(0)$ is the power of the signal, the prediction error power is given by

$$P_{N_0} = u^{N_0}(0) = P_x \prod_1^{N_0}(1 - C_n^2).$$

We can conclude that, if $R(n)$ is available for all $n \in \mathbb{Z}$, the right-hand side of $u^N(n)$ becomes identically null, while $u^N(n), n \le 0$ converges to $h(-n)$, $h(n)$ being the impulse response of the system. This procedure can be related to the Schur algorithm [9].

Rewriting (7) for the order N and attending to the relation $v^N(n) = u^N(N+1-n)$, we arrive to the forward and backward prediction error filters defined by [4] and [6]

$$A^N(z) = \sum_0^N a_i z^{-i} \text{ and}$$
$$B^N(z) = \sum_1^{N+1} a_{N+1-i} z^{-i} \quad (12)$$

and verifying the recursions

$$A^N(z) = A^{N-1}(z) + C_N B^{N-1}(z) \quad (13)$$
$$B^N(z) = z^{-1} B^{N-1}(z) + C_N z^{-1} A^{N-1}(z) \quad (14)$$

that were also obtained by Baran and Oppenheim. The corresponding forward, f_n^N, and backward, b_n^N, prediction errors verify the recursions [6]

$$f_n^N = f_n^{N-1} + C_N b_n^{N-1} \quad (15)$$
$$b_n^N = b_{n-1}^{N-1} + C_N f_{n-1}^{N-1} \quad (16)$$

under initial conditions $f_n^0 = x_n, b_n^0 = x_{n-1}$. With these recursions, I devised a modified Burg method that gives better results than the classic Burg method, mainly with short data records. In fact, it exhibits less phase dependence and smaller coefficient bias and prediction error power [6]. As seen above, we can illustrate the action of forward and backward prediction error filters over the signal x_n by

$$\xrightarrow{A^N(z)} \quad x_{n-N-1}\ x_{n-N}\ x_{n-N+1} \cdots x_{n-2}\ x_{n-1}\ x_n \quad \xleftarrow[B^N(z)]{}$$

The action of these filters is somehow different from those described above. In fact, the filters only share N–1 points. The past of the forward filter is the future of the backward one.

We must emphsize two important facts:

- In the AR(p) case, the number of reflection coefficients is equal to p, and we need only the first $p+1$ values of the autocorrelation function. At the end of the pth recursion $u^N(n)$ has infinite gaps for $n > 0$.
- In the MA(q) case, the number of reflection coefficients is infinite, but the $q+1$ available autocorrelation values are enough to compute them [3], [5]. The recursion is self-maintained.

SOLUTION IN THE ARMA CASE

The algorithm just described is suitable for AR or even MA modeling, but not for ARMA. We are going to examine how to devise an algorithm for solving the normal equations in the ARMA case. The algorithm we are going to describe is a reformulation of the one presented in [5] by using the above approach, and it was described for the first time in [8] where it was applied to modeling impulse responses.

The normal equations in the ARMA (p, q) case have a nonsymmetric Toeplitz matrix and can be written in the format [5]

$$\sum_{i=0}^N a_i R[M + (j-i)] = 0, j = 1, \ldots, N \quad (17)$$

with $N \le p$ and M an assumed MA order. To solve the ARMA normal equations, we introduce an adjoint system defined by

$$\sum_{i=0}^N \gamma_i R[M - (j-i)] = 0, j = 1, \ldots, N. \quad (18)$$

This system of equations has a matrix that is the transposed of the one in (9). As it is evident, in the AR case they are the same.

As in the above, we introduce two functions $u_M^{N-1}(n)$ and $\phi_M^{N-1}(n)$ defined by

$$u_M^{N-1}(n) = \sum_{i=0}^N a_i R[M + (n-i)] \quad (19)$$

and

$$\phi_M^{N-1}(n) = \sum_{i=0}^N \gamma_i R[M - (n-i)]. \quad (20)$$

We also introduce two corresponding time-reversed and shifted functions with coincident gaps: $v_M^{N-1}(n) = u_M^{N-1}(N-n)$ and $\psi_M^{N-1}(n) = \phi_M^{N-1}(N-n)$. Construct the recursions

$$u_M^N(n) = u_M^{N-1}(n) + C_{N,M} \psi_M^{N-1}(n) \quad (21)$$
$$\phi_M^N(n) = \phi_M^{N-1}(n) + K_{N,M} v_M^{N-1}(n) \quad (22)$$

with

$$\phi_M^0(n) = R(M+n) \quad (23)$$
$$\phi_M^0(n) = R(M-n) \quad (24)$$

leading to

$$v_M^N(n) = v_M^{N-1}(n-1) + C_{N,M} \phi_M^{N-1}(n-1) \quad (25)$$

$$\psi_M^N(n) = \psi_M^{N-1}(n-1) + K_{N,M} u_M^{N-1}(n-1). \quad (26)$$

$C_{N,M}$ and $K_{N,M}$ are two reflection coefficients that are obtained by forcing $u_M^N(n)$ and $\phi_M^N(n)$ to have a gap at $n = N$. This leads to

$$C_{N,M} = -(u_M^{N-1}(N))/(\psi_M^{N-1}(N)) = -(u_M^{N-1}(N))/(\phi_M^{N-1}(0)) \quad (27)$$

$$K_{N,M} = -(\phi_M^{N-1}(N))/(v_M^{N-1}(N)) = -(\phi_M^{N-1}(N))/(u_M^{N-1}(0)) \quad (28)$$

$$v_M^N(0) = u_M^N(0)(1 - C_{N,M} K_{N,M}). \quad (29)$$

With these two sets of reflection coefficients we can define two AR polynomials $A_M^N(z) = \sum_{i=0}^N a_M^N(i) z^{-i}$ and $\Gamma_M^N(z) = \sum_{i=0}^N \gamma_M^N(i) z^{-i}$ with coefficients

[lecture NOTES] continued

verifying the so-called double Levinson recursion [7]

$$a_M^N(i) = a_M^{N-1}(i) + C_{N,M}\gamma_M^{N-1}(N-i) \quad (30)$$
$$\gamma_M^N(i) = \gamma_M^{N-1}(i) + K_{N,M}a_M^{N-1}(N-i) \quad (31)$$

with $i = 0, 1, \ldots, N$. These recursions reproduce twice (8), because both polynomials are forward. We can define also two backward polynomials verifying two recursions identical to (9). When $M = 0$ these polynomials are equal and the recursion recovers the one stated in (15). The sets of $a_M^N(i), \gamma_M^N(i) i = 0, 1, \ldots, N$ are the solutions of the normal equations, (17) and (18), respectively.

It can be shown that $A_M^N(z)$ and $\Gamma_{M+1}^N(z)$ are aside a constant reverse polynomials, which leads to

$$C_{N,M}K_{N,M+1} = 1, \quad (32)$$

an important relation that is useful in the determination of the orders.

From these polynomials we can

- construct a recursive algorithm for obtaining the MA parameters
- obtain an ARMA Burg-like algorithm, but this goes beyond the objectives of this column [5], [8].

CONCLUSIONS

There is an alternative solution for the normal equations that lead to different use of the predictor coefficients when defining the forward and backward prediction error filters. A simple algorithm for the computation of the predictor coefficients was presented. Its generalization for the computation of the AR parameters of an ARMA model was also described.

ACKNOWLEDGMENT

This work was partially funded by national funds through the Foundation for Science and Technology of Portugal.

AUTHOR

Manuel D. Ortigueira (mdo@fct.unl.pt) is an associate professor of the Department of Electrical Engineering of Faculty of Sciences and Technology of Universidade Nova de Lisboa and senior investigator in the signal processing group of UNINOVA.

REFERENCES

[1] T. A. Baran and A. V. Oppenheim, "A derivation of the recursive solution to the autocorrelation normal equations," *IEEE Signal Processing Mag.*, vol. 30, no. 1, pp. 142–144, Jan. 2013.

[2] J. P. Burg, "Maximum entropy spectral analysis," Ph.D. dissertation, Dept. Geophys., Stanford Univ., Stanford, CA, May 1975.

[3] J. Le Roux and Y. Grenier, "An iterative procedure for moving average models estimation," in *Proc. IEEE ICASSP*, 1980.

[4] M. D. Ortigueira, "Methodology, implementation and evaluation in spectral analysis (in Portuguese)," Ph.D. dissertation, Instituto Superior Tecnico, Lisbon, 1984.

[5] M. D. Ortigueira, "ARMA realization from the reflection coefficient sequence," *Signal Processing*, vol. 32, no. 3, June 1993.

[6] M. D. Ortigueira and J. M. Tribolet, "Global versus local minimization in least-squares AR spectral estimation," *Signal Processing*, vol. 7, pp. 267–281, 1984.

[7] M. D. Ortigueira and J. M. Tribolet, "On the double Levinson recursion formulation of ARMA spectral estimation," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Boston, MA, 1983.

[8] M. D. Ortigueira and A. J. Serralheiro, "Pseudo-fractional ARMA modelling using a double Levinson recursion," *IET Control Theory Applicat.*, vol. 1, no. 1, pp. 173–178, Jan. 2007.

[9],J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*. Englewood Cliffs, NJ: Prentice Hall, 2007.

[10] E. A. Robinson and S. Treitel, "Maximum entropy and the relationship of the partial autocorrelation to the reflection coefficients of a layered system," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-28, no. 2, Apr. 1980.

[SP]

[from the GUEST EDITORS] (continued from page 15)

the receiver signal processing task to multiple slower processors, performed in the frequency domain (FD) rather than in the time domain. They show that subband-based FD parallelization is especially suited to long-haul optical fiber communication systems.

Lau et al. describe the advances in DSP techniques that enable Tb/s transmission as well as software-defined flexible transponders supporting adaptive modulation formats and elastic optical networks. Among the covered DSP topics are carrier phase estimation for high spectral efficiency and adaptive signal processing techniques. Progress towards a universal DSP platform for arbitrary QAM transmission is also covered.

Beygi et al. review the joint design of forward error correction and modulation for fiber-optic communications. The authors use an information-theoretic design framework to investigate coded modulation (CM) techniques for fiber-optic channels. The authors further discuss two- and four-dimensional CM schemes. They also address the computational complexity and hardware constraints of CM schemes. Finally, the authors treat signal shaping and rate-adaptation capabilities to accommodate different signal qualities.

Djordjevic et al. describe how to jointly address the limited bandwidth of the transport infrastructure, high energy consumption, and network heterogeneity issues. They present an adaptive software-defined low-density parity check-coded multiband approach involving spatial-MIMO and all-optical-OFDM, enabling energy-efficient high-bandwidth delivery with fine granularity and elastic bandwidth utilization. The modulation is based on multidimensional signaling to improve the tolerance to fiber nonlinearities and imperfect compensation of channel impairments, based on a hybrid nature employing both electrical and optical degrees of freedom.

The guest editors would like to express their gratitude to colleagues and individuals who directly or indirectly contributed to this special issue. In particular, we would like to acknowledge all the authors who have submitted their manuscripts as well as all the reviewers whose review comments have greatly improved the overall quality of the issue. Special thanks are extended to Editor-in-Chief Abdelhak Zoubir and Special Issues Area Editor Fulvio Gini for their tremendous support during the solicitation and review processes, as well as to Rebecca Wollman for her professional administrative assistance in the logistics of the special issue and its promotion. All the aforementioned supports that we received are indispensable for making this special issue happen.

We hope that this issue will provide the readers with insights into a broad scope of advanced DSP and coding technologies for future multi-Tb/s optical transport.

[SP]

dates AHEAD

Please send calendar submissions to:
Dates Ahead, c/o Jessica Barragué
IEEE Signal Processing Magazine
445 Hoes Lane
Piscataway, NJ 08855 USA
e-mail: j.barrague@ieee.org
(Colored conference title indicates SP-sponsored conference.)

2014

MARCH

IEEE World Forum on Internet of Things (IOT)
6–8 March, Seoul, South Korea.
General Chair: Roberto Minerva
General Cochairs: Jaiyong Lee, Edward Cole, Stephen Dukes, Paul Hartmann, Abbas Jamalipour, Jean Philippe Vasseur, John Vig, Yen-Kuang Chen, and Mengchu Zhou
URL: http://sites.ieee.org/wf-iot//

Data Compression Conference (DCC)
26–28 March, Salt Lake City, Utah.
URL: http://cs.brandeis.edu/~dcc/index.html

APRIL

IEEE International Symposium on Biomedical Imaging (ISBI)
28 April–2 May, Beijing, China.
General Chairs: Ge Wang and Bin He
URL: http://biomedicalimaging.org/2014/

Digital Object Identifier 10.1109/MSP.2013.2296120
Date of publication: 12 February 2014

MAY

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
4–9 May, Florence, Italy.
General Cochairs: Fulvio Gini and Marco Luise
URL: http://www.icassp2014.org/

IEEE Radar Conference (RadarCon)
19–23 May, Cincinnati, Ohio.
General Chair: Brian Rigling
URL: http://www.radarcon2014.org/

6th International Symposium on Communications, Control, and Signal Processing (ISCCSP)
21–23 May, Athens, Greece.
Honorary Chair: Anastasios Venetsanopoulos
General Cochairs: Sanjit K. Mitra and Thanos Stouraitis
URL: http://isccsp2014.upatras.gr/

JUNE

IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)
22–25 June, A Caruña, Spain.
URL: http://www.gtec.udc.es/sam2014/

IEEE Statistical Signal Processing Workshop (SSP)
29 June–2 July, Gold Coast, Australia.
General Cochairs: Rob Evans and Abd-Krim Seghouane
URL: http://www.ee.unimelb.edu.au/SSP2014/index.html

JULY

2nd IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)
9–13 July, Xi'an, China.
General Chairs: Mingyi He and Kung Yao
URL: http://www.chinasip2014.org/CfP.htm

IEEE International Conference on Multimedia and Expo (ICME)
14–18 July, Chengdu, China.
General Cochairs: Touradj Ebrahimi, Shipeng Li, Houjun Wang, and Jie Yang
URL: http://www.icme2014.org/

AUGUST

11th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)
26–29 August, Seoul, South Korea.
General Chair: Hanseok Ko
General Cochair: Jin Young Choi
URL: http://www.avss2014.org/

OCTOBER

IEEE Workshop on Signal Processing Systems (SIPS)
20–23 October, Belfast, Ireland.

DECEMBER

IEEE Global Conference on Signal and Information Processing (GlobalSIP)
3–5 December, Atlanta, Georgia.
General Chairs: Geoffrey Li and Fred Juang
URL: http://renyi.ece.iastate.edu/globalsip2014/

IEEE Spoken Language Technology Workshop (SLT)
6–9 December, South Lake Tahoe, California.
General Chairs: Murat Akbacak and John Hansen

2015

APRIL

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
19–24 April, Brisbane, Australia.
General Cochairs: Vaughan Clarkson and Jonathan Manton
URL: http://icassp2015.org/

advertisers INDEX

The Advertisers Index contained in this issue is compiled as a service to our readers and advertisers: the publisher is not liable for errors or omissions although every effort is made to ensure its accuracy. Be sure to let our advertisers know you found them through *IEEE Signal Processing Magazine*.

ADVERTISER	PAGE	URL	PHONE
Asilomar Conference	7	www.asilomarssc.org	
IEEE MDL/Marketing	3	www.ieee.org/go/trymdl	
Mathworks	CVR 4	www.mathworks.com/accelerate	+1 508 647 7040
Mini-Circuits	CVR 2, 5, CVR 3	www.minicircuits.com	+1 718 934 4500

advertising SALES OFFICES

James A. Vick
Sr. Director, Advertising
Phone: +1 212 419 7767;
Fax: +1 212 419 7589
jv.ieeemedia@ieee.org

Marion Delaney
Advertising Sales Director
Phone: +1 415 863 4717;
Fax: +1 415 863 4717
md.ieeemedia@ieee.org

Susan E. Schneiderman
Business Development Manager
Phone: +1 732 562 3946;
Fax: +1 732 981 1855
ss.ieeemedia@ieee.org

Product Advertising

MIDATLANTIC
Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, PA, DE, MD, DC, KY, WV

NEW ENGLAND/SOUTH CENTRAL/ EASTERN CANADA
Jody Estabrook
Phone: +1 774 283 4528;
Fax: +1 774 283 4527
je.ieeemedia@ieee.org
ME, VT, NH, MA, RI, CT, AR, LA, OK, TX
Canada: Quebec, Nova Scotia, Newfoundland, Prince Edward Island, New Brunswick

SOUTHEAST
Thomas Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
tf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

MIDWEST/CENTRAL CANADA
Dave Jones
Phone: +1 708 442 5633;
Fax: +1 708 442 7620
dj.ieeemedia@ieee.org
IL, IA, KS, MN, MO, NE, ND, SD, WI, OH
Canada: Manitoba, Saskatchewan, Alberta

MIDWEST/ ONTARIO, CANADA
Will Hamilton
Phone: +1 269 381 2156;
Fax: +1 269 381 2556
wh.ieeemedia@ieee.org
IN, MI. Canada: Ontario

WEST COAST/MOUNTAIN STATES/ WESTERN CANADA
Marshall Rubin
Phone: +1 818 888 2407;
Fax: +1 818 888 4907
mr.ieeemedia@ieee.org
AZ, CO, HI, NM, NV, UT, AK, ID, MT, WY, OR, WA, CA. Canada: British Columbia

EUROPE/AFRICA/MIDDLE EAST ASIA/FAR EAST/PACIFIC RIM
Louise Smith
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
les.ieeemedia@ieee.org
Europe, Africa, Middle East
Asia, Far East, Pacific Rim, Australia, New Zealand

Recruitment Advertising

MIDATLANTIC
Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, CT, PA, DE, MD, DC, KY, WV

NEW ENGLAND/EASTERN CANADA
Liza Reich
Phone: +1 212 419 7578;
Fax: +1 212 419 7589
e.reich@ieee.org
ME, VT, NH, MA, RI. Canada: Quebec, Nova Scotia, Prince Edward Island, Newfoundland, New Brunswick

SOUTHEAST
Cathy Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
cf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

MIDWEST/SOUTH CENTRAL/ CENTRAL CANADA
Darcy Giovingo
Phone: +224 616 3034;
Fax: +1 847 729 4269
dg.ieeemedia@ieee.org;
AR, IL, IN, IA, KS, LA, MI, MN, MO, NE, ND, SD, OH, OK, TX, WI. Canada: Ontario, Manitoba, Saskatchewan, Alberta

WEST COAST/SOUTHWEST/ MOUNTAIN STATES/ASIA
Tim Matteson
Phone: +1 310 836 4064;
Fax: +1 310 836 4067
tm.ieeemedia@ieee.org
AZ, CO, HI, NV, NM, UT, CA, AK, ID, MT, WY, OR, WA. Canada: British Columbia

EUROPE/AFRICA/MIDDLE EAST
Louise Smith
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
les.ieeemedia@ieee.org
Europe, Africa, Middle East

Digital Object Identifier 10.1109/MSP.2013.2290960

CERAMIC FILTERS

LOW PASS BANDPASS HIGH PASS

45 MHz to 15 GHz from **99¢** ea. qty.3000

Over ~~167~~ 185 models...only 0.12 x 0.06" These tiny hermetically sealed filters utilize our advanced Low Temperature Co-fired Ceramic **(LTCC)** technology to offer superior thermal stability, high reliability, and very low cost, making them a must for your system requirements. Visit our website to choose and view comprehensive performance curves, data sheets, pcb layouts, and everything you need to make your choice. You can even order direct from our web store and have units in your hands as early as tomorrow!

***Wild Card** KWC-LHP LTCC Filter Kits* only $98

Choose any 8 LFCN or HFCN models
Receive 5 of ea. model, for a total of 40 filters
Order your KWC-LHP FILTER KIT TODAY!

RoHS compliant U.S. Patents 7,760,485 and 6,943,646

Now available in small-quantity reels at *no extra charge:*
Standard counts of 20, 50, 100, 200, 500, 1000 or 2000. Save time, money, and inventory space!

Mini-Circuits®

www.minicircuits.com P.O. Box 35166, Brooklyn, NY 11235-0003 (718) 934-4500 sales@minicircuits.com

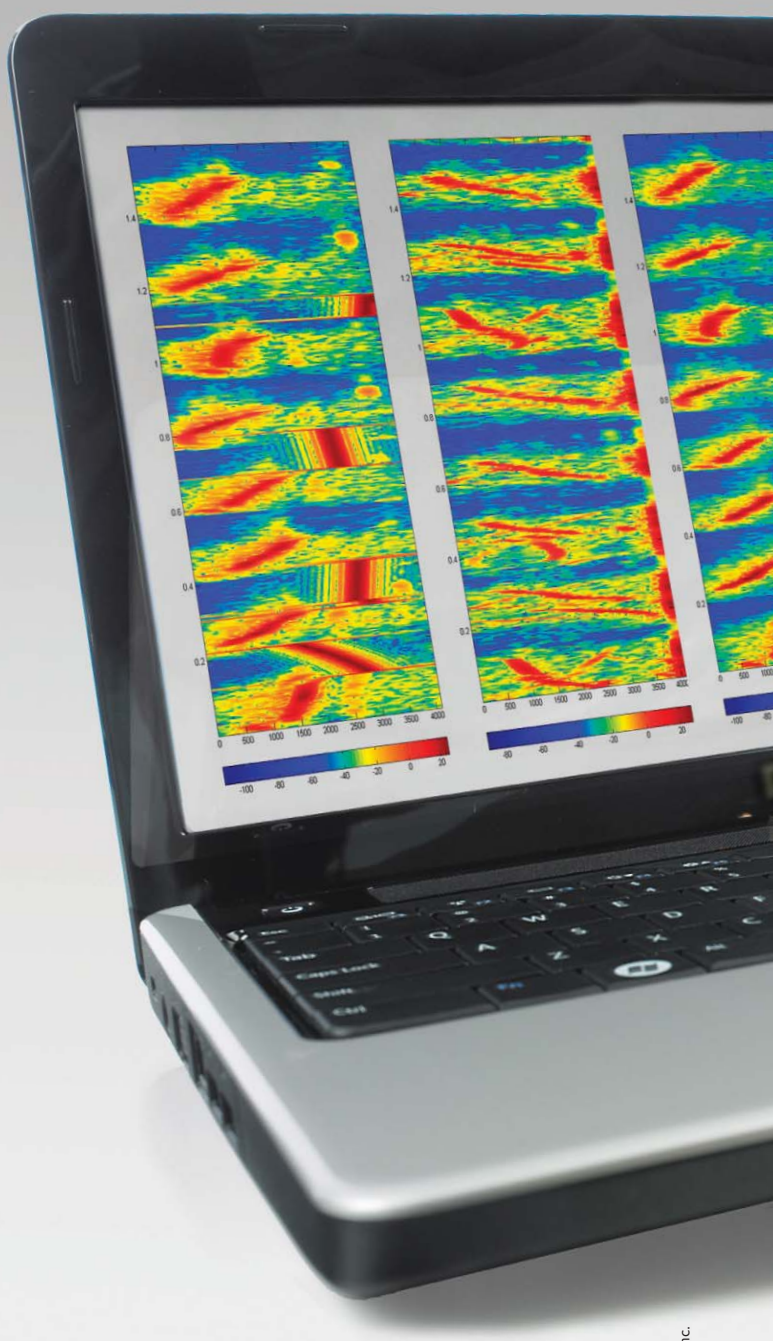
504 Rev C

MathWorks®
Accelerating the pace of engineering and science

Find it at
mathworks.com/accelerate
datasheet
video example
trial request

RUN MATLAB PROGRAMS IN PARALLEL

with
Parallel Computing Toolbox™



If you can write a FOR-loop, you can write a MATLAB parallel program for high-performance computing. Parallel Computing Toolbox lets you quickly adapt MATLAB programs to run on your multicore computer, GPU or cluster, with features for task-parallel and data-parallel programs.

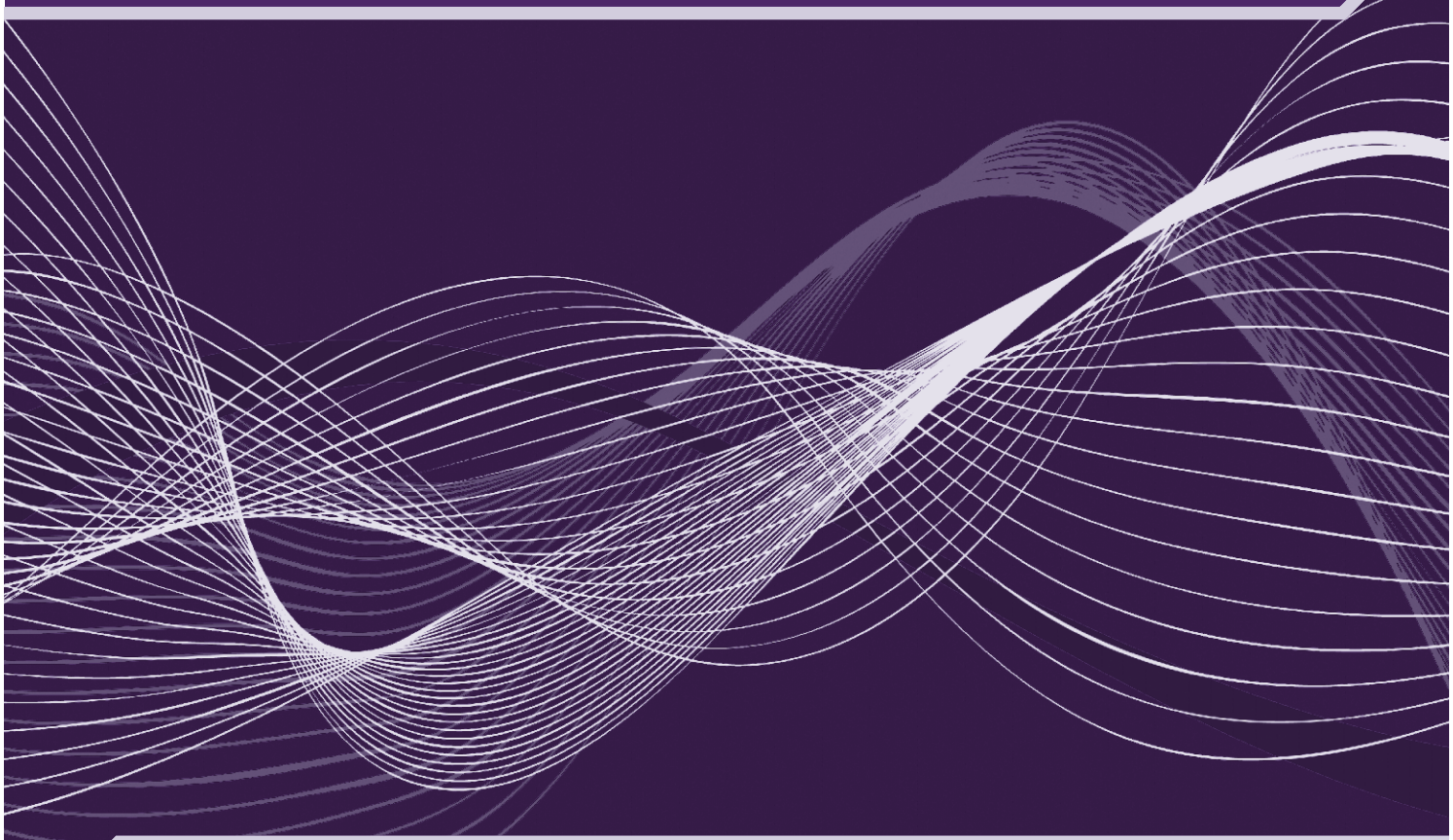
MATLAB® & SIMULINK®

©2010 The MathWorks, Inc.

IEEE SIGNAL PROCESSING SOCIETY

CONTENT GAZETTE

[ISSN 2167-5023]



MARCH 2014

IEEE Signal Processing Society

IEEE

The IEEE Global Conference on Signal and Information Processing (GlobalSIP) is a recently launched flagship conference of the IEEE Signal Processing Society. GlobalSIP14 will focus broadly on signal and information processing with an emphasis on up-and-coming signal processing themes. The conference will feature world-class speakers, tutorials, exhibits, and technical sessions consisting of poster or oral presentations. GlobalSIP14 will be comprised of co-located symposia.

Topics of interests include (but are not limited to):

- Workshop on Information Forensics and Security
- Signal Processing for Cognitive Radios and Networks
- Network Theory
- Energy Efficiency and Energy Harvesting Related Signal Processing and Communications
- Game Theory for Signal Processing and Communications
- Large-Scale MIMO Communications
- Optimizing Future Energy Systems
- Active Authentication
- Dynamic Sparse and/or Low-Rank Recovery
- Subspace Methods for High-Dimensional Data
- Internet of Things
- Data Flow Algorithms and Architecture for Signal Processing Systems
- Mixed-Signal and Optical Sensing: Hardware to Algorithms
- Signal Processing Challenges and Architectures for High Throughput Satellite Communications
- Advanced Topics in Signal and Information Processing
- Machine Learning for Acoustic Signal Processing
- Signal Processing for Next Generation Semiconductor Integrated Circuits

Organizers

- **General Chairs**
  - Geoffrey Li (Georgia Inst. of Tech.)
  - Fred Juang (Georgia Inst. of Tech.)
- **Technical Chairs**
  - Douglas Williams (Georgia Inst. of Tech.)
  - Timothy Davidson (McMaster University)
  - Ghassan AlRegib (Georgia Inst. of Tech.)
- **Finance Chair (Treasurer)**
  David V. Anderson (Georgia Inst. of Tech.)
- **Local Arrangement Chair**
  Xiaoli Ma (Georgia Inst. of Tech.)
- **Publications and Publicity Chair**
  Zhengdao Wang (Iowa State University)
- **Industrial Liaison Chair**
  Ahsan Aziz (National Instruments)

Important Dates

- May 16, 2014: Paper submission deadline
- June 27, 2014: Review results announced
- September 5, 2014: Camera-ready papers due

Website

http://renyi.ece.iastate.edu/globalsip2014

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

JANUARY 1, 2014 VOLUME 62 NUMBER 1 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

A Dual-Phase Power Allocation Scheme for Multicarrier Relay System With Direct Link http://dx.doi.org/10.1109/TSP.2013.2283455 ..... *Y. Ma, A. Liu, and Y. Hua* 5

An EM Approach for Time-Variant Poisson-Gaussian Model Parameter Estimation http://dx.doi.org/10.1109/TSP.2013.2283839 ..... *A. Jezierska, C. Chaux, J.-C. Pesquet, H. Talbot, and G. Engler* 17

Resilient Detection in the Presence of Integrity Attacks http://dx.doi.org/10.1109/TSP.2013.2284145 ..... *Y. Mo, J. P. Hespanha, and B. Sinopoli* 31

Nonuniform Sampling Walls in Wideband Signal Detection http://dx.doi.org/10.1109/TSP.2013.2284765 ..... *J. Font-Segura, G. Vazquez, and J. Riba* 44

The Stationary Phase Approximation, Time-Frequency Decomposition and Auditory Processing http://dx.doi.org/10.1109/TSP.2013.2284479 ..... *B. Mulgrew* 56

Matching Demodulation Transform and SynchroSqueezing in Time-Frequency Analysis http://dx.doi.org/10.1109/TSP.2013.2276393 ..... *S. Wang, X. Chen, G. Cai, B. Chen, X. Li, and Z. He* 69

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Joint User Grouping and Transceiver Design in a MIMO Interfering Broadcast Channel http://dx.doi.org/10.1109/TSP.2013.2284753 ..... M. Razaviyayn, M. Baligh, A. Callard, and Z.-Q. Luo 85

Cramér-Rao Bounds for UMTS-Based Passive Multistatic Radar http://dx.doi.org/10.1109/TSP.2013.2284758 ..... S. Gogineni, M. Rangaswamy, B. D. Rigling, and A. Nehorai 95

Robust Training Sequence Design for Correlated MIMO Channel Estimation http://dx.doi.org/10.1109/TSP.2013.2284763 ..... N. Shariati, J. Wang, and M. Bengtsson 107

Exact LTP Representation of the Generalized Periodic-Reference FxLMS Algorithm http://dx.doi.org/10.1109/TSP.2013.2285513 ..... T. Haarnoja, K. Tammi, and K. Zenger 121

A New Deterministic Identification Approach to Hammerstein Systems http://dx.doi.org/10.1109/TSP.2013.2286103 ..... C. Yu, C. Zhang, and L. Xie 131

Learning Nonlinear Generative Models of Time Series With a Kalman Filter in RKHS http://dx.doi.org/10.1109/TSP.2013.2283842 ..... P. Zhu, B. Chen, and J. C. Prncipe 141

Uniformly Improving Maximum-Likelihood SNR Estimation of Known Signals in Gaussian Channels http://dx.doi.org/10.1109/TSP.2013.2274638 ..... E. Stathakis, J. Jaldén, L. K. Rasmussen, and M. Skoglund 156

EM- and JMAP-ML Based Joint Estimation Algorithms for Robust Wireless Geolocation in Mixed LOS/NLOS Environments http://dx.doi.org/10.1109/TSP.2013.2286779 ..... F. Yin, C. Fritsche, F. Gustafsson, and A. M. Zoubir 168

Convergence and Stability of Iteratively Re-weighted Least Squares Algorithms http://dx.doi.org/10.1109/TSP.2013.2287685 ..... D. Ba, B. Babadi, P. L. Purdon, and E. N. Brown 183

Permutation Meets Parallel Compressed Sensing: How to Relax Restricted Isometry Property for 2D Sparse Signals http://dx.doi.org/10.1109/TSP.2013.2284762 ..... H. Fang, S. A. Vorobyov, H. Jiang, and O. Taheri 196

Non-Negative Matrix Factorization Revisited: Uniqueness and Algorithm for Symmetric Decomposition http://dx.doi.org/10.1109/TSP.2013.2285514 ..... K. Huang, N. D. Sidiropoulos, and A. Swami 211

Bayesian Model Comparison With the g-Prior http://dx.doi.org/10.1109/TSP.2013.2286776 ..... J. K. Nielsen, M. G. Christensen, A. T. Cemgil, and S. H. Jensen 225

Biosensor Arrays for Estimating Molecular Concentration in Fluid Flows http://dx.doi.org/10.1109/TSP.2013.2287680 ..... M. Abolfath-Beygi and V. Krishnamurthy 239

Statistical Analysis of a Jointly Optimized Beamformer-Assisted Acoustic Echo Canceler http://dx.doi.org/10.1109/TSP.2013.2284138 ..... M. H. Maruo, J. C. M. Bermudez, and L. S. Resende 252

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

JANUARY 1, 2014 VOLUME 62 NUMBER 1 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Statistical Signal Processing

Resilient Detection in the Presence of Integrity Attacks http://dx.doi.org/10.1109/TSP.2013.2284145 ..... *Y. Mo, J. P. Hespanha, and B. Sinopoli* 31

Nonuniform Sampling Walls in Wideband Signal Detection http://dx.doi.org/10.1109/TSP.2013.2284765 ..... *J. Font-Segura, G. Vazquez, and J. Riba* 44

Uniformly Improving Maximum-Likelihood SNR Estimation of Known Signals in Gaussian Channels http://dx.doi.org/10.1109/TSP.2013.2274638 ..... *E. Stathakis, J. Jaldén, L. K. Rasmussen, and M. Skoglund* 156

Non-Negative Matrix Factorization Revisited: Uniqueness and Algorithm for Symmetric Decomposition http://dx.doi.org/10.1109/TSP.2013.2285514 ..... *K. Huang, N. D. Sidiropoulos, and A. Swami* 211

Bayesian Model Comparison With the g-Prior http://dx.doi.org/10.1109/TSP.2013.2286776 ..... *J. K. Nielsen, M. G. Christensen, A. T. Cemgil, and S. H. Jensen* 225

An EM Approach for Time-Variant Poisson-Gaussian Model Parameter Estimation http://dx.doi.org/10.1109/TSP.2013.2283839 ..... *A. Jezierska, C. Chaux, J.-C. Pesquet, H. Talbot, and G. Engler* 17

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Adaptive Signal Processing

Exact LTP Representation of the Generalized Periodic-Reference FxLMS Algorithm http://dx.doi.org/10.1109/TSP.2013.2285513 ..... *T. Haarnoja, K. Tammi, and K. Zenger* 121

Learning Nonlinear Generative Models of Time Series With a Kalman Filter in RKHS http://dx.doi.org/10.1109/TSP.2013.2283842 ..... *P. Zhu, B. Chen, and J. C. Prncipe* 141

Statistical Analysis of a Jointly Optimized Beamformer-Assisted Acoustic Echo Canceler http://dx.doi.org/10.1109/TSP.2013.2284138 ..... *M. H. Maruo, J. C. M. Bermudez, and L. S. Resende* 252

Digital and Multirate Signal Processing

Matching Demodulation Transform and SynchroSqueezing in Time-Frequency Analysis http://dx.doi.org/10.1109/TSP.2013.2276393 ..... *S. Wang, X. Chen, G. Cai, B. Chen, X. Li, and Z. He* 69

The Stationary Phase Approximation, Time-Frequency Decomposition and Auditory Processing http://dx.doi.org/10.1109/TSP.2013.2284479 ..... *B. Mulgrew* 56

Permutation Meets Parallel Compressed Sensing: How to Relax Restricted Isometry Property for 2D Sparse Signals http://dx.doi.org/10.1109/TSP.2013.2284762 ..... *H. Fang, S. A. Vorobyov, H. Jiang, and O. Taheri* 196

Convergence and Stability of Iteratively Re-weighted Least Squares Algorithms http://dx.doi.org/10.1109/TSP.2013.2287685 ..... *D. Ba, B. Babadi, P. L. Purdon, and E. N. Brown* 183

Multidimensional Signal Processing

A New Deterministic Identification Approach to Hammerstein Systems http://dx.doi.org/10.1109/TSP.2013.2286103 ..... *C. Yu, C. Zhang, and L. Xie* 131

Sensor Array and Multichannel Processing

Cramér-Rao Bounds for UMTS-Based Passive Multistatic Radar http://dx.doi.org/10.1109/TSP.2013.2284758 ..... *S. Gogineni, M. Rangaswamy, B. D. Rigling, and A. Nehorai* 95

Signal Processing for Communications

A Dual-Phase Power Allocation Scheme for Multicarrier Relay System With Direct Link http://dx.doi.org/10.1109/TSP.2013.2283455 ..... *Y. Ma, A. Liu, and Y. Hua* 5

MIMO Communications & Signal Processing

Robust Training Sequence Design for Correlated MIMO Channel Estimation http://dx.doi.org/10.1109/TSP.2013.2284763 ..... *N. Shariati, J. Wang, and M. Bengtsson* 107

Signal Processing for Sensor Networks

EM- and JMAP-ML Based Joint Estimation Algorithms for Robust Wireless Geolocation in Mixed LOS/NLOS Environments http://dx.doi.org/10.1109/TSP.2013.2286779 ..... *F. Yin, C. Fritsche, F. Gustafsson, and A. M. Zoubir* 168

Signal Processing for Wireless Networks

Joint User Grouping and Transceiver Design in a MIMO Interfering Broadcast Channel http://dx.doi.org/10.1109/TSP.2013.2284753 ..... *M. Razaviyayn, M. Baligh, A. Callard, and Z.-Q. Luo* 85

Biomedical Signal Processing

Biosensor Arrays for Estimating Molecular Concentration in Fluid Flows http://dx.doi.org/10.1109/TSP.2013.2287680 ..... *M. Abolfath-Beygi and V. Krishnamurthy* 239

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed | MEDLINE U.S. National Library of Medicine

JANUARY 15, 2014 VOLUME 62 NUMBER 2 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

An MMSE Strategy at Relays With Partial CSI for a Multi-Layer Relay Network http://dx.doi.org/10.1109/TSP.2013.2284471 ..... *P. S. Elamvazhuthi, B. K. Dey, and S. Bhashyam* 271

Fully Distributed Optimal Channel Assignment for Open Spectrum Access http://dx.doi.org/10.1109/TSP.2013.2285512 ..... *O. Naparstek and A. Leshem* 283

Matrix Completion in Colocated MIMO Radar: Recoverability, Bounds & Theoretical Guarantees http://dx.doi.org/10.1109/TSP.2013.2287673 ..... *D. S. Kalogerias and A. P. Petropulu* 309

Time Domain Channel Estimation for OQAM-OFDM Systems: Algorithms and Performance Bounds http://dx.doi.org/10.1109/TSP.2013.2290498 ..... *D. Kong, D. Qu, and T. Jiang* 322

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Persymmetric Adaptive Detectors in Homogeneous and Partially Homogeneous Environments http://dx.doi.org/10.1109/TSP.2013.2288087 ..... *Y. Gao, G. Liao, S. Zhu, X. Zhang, and D. Yang* 331

Multi-Cell Beamforming With Decentralized Coordination in Cognitive and Cellular Networks http://dx.doi.org/10.1109/TSP.2013.2287681 ..... *H. Pennanen, A. Tölli, and M. Latva-aho* 295

MIMO Radar Waveform Design With Constant Modulus and Similarity Constraints http://dx.doi.org/10.1109/TSP.2013.2288086 ..... *G. Cui, H. Li, and M. Rangaswamy* 343

Adaptive Identification and Recovery of Jointly Sparse Vectors http://dx.doi.org/10.1109/TSP.2013.2288679 ..... *R. Amel and A. Feuer* 354

Sensor Selection Based on Generalized Information Gain for Target Tracking in Large Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2289881 ..... *X. Shen and P. K. Varshney* 363

Non-Negative Blind Source Separation Algorithm Based on Minimum Aperture Simplicial Cone http://dx.doi.org/10.1109/TSP.2013.2287683 ..... *W. S. B. Ouedraogo, A. Souloumiac, M. Jaidane, and C. Jutten* 376

Cache-Enabled Opportunistic Cooperative MIMO for Video Streaming in Wireless Systems http://dx.doi.org/10.1109/TSP.2013.2291211 ..... *A. Liu and V. K. N. Lau* 390

Estimation of Space-Time Varying Parameters Using a Diffusion LMS Algorithm http://dx.doi.org/10.1109/TSP.2013.2289888 ..... *R. Abdolee, B. Champagne, and A. H. Sayed* 403

Spatial Compressive Sensing for MIMO Radar http://dx.doi.org/10.1109/TSP.2013.2289875 .... *M. Rossi, A. M. Haimovich, and Y. C. Eldar* 419

Asymptotically Optimal Truncated Multivariate Gaussian Hypothesis Testing With Application to Consensus Algorithms http://dx.doi.org/10.1109/TSP.2013.2288673 ..... *J. Zhang and R. S. Blum* 431

Energy Efficiency Optimization in Relay-Assisted MIMO Systems With Perfect and Statistical CSI http://dx.doi.org/10.1109/TSP.2013.2292031 ..... *A. Zappone, P. Cao, and E. A. Jorswieck* 443

Sampling Curves With Finite Rate of Innovation http://dx.doi.org/10.1109/TSP.2013.2292033 ..... *H. Pan, T. Blu, and P. L. Dragotti* 458

Distributed Least Mean-Square Estimation With Partial Diffusion http://dx.doi.org/10.1109/TSP.2013.2292035 ..... *R. Arablouei, S. Werner, Y.-F. Huang, and K. Doğançay* 472

MIMO Systems With Quantized Covariance Feedback http://dx.doi.org/10.1109/TSP.2013.2292036 ..... *R. T. Krishnamachari, M. K. Varanasi, and K. Mohanty* 485

Energy Efficient Collaborative Beamforming in Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2288080 ..... *B. Béjar Haro, S. Zazo, and D. P. Palomar* 496

MIMO Detection by Lagrangian Dual Maximum-Likelihood Relaxation: Reinterpreting Regularized Lattice Decoding http://dx.doi.org/10.1109/TSP.2013.2292040 ..... *J. Pan, W.-K. Ma, and J. Jaldén* 511

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

JANUARY 15, 2014 VOLUME 62 NUMBER 2 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Statistical Signal Processing

Persymmetric Adaptive Detectors in Homogeneous and Partially Homogeneous Environments http://dx.doi.org/10.1109/TSP.2013.2288087 ..... *Y. Gao, G. Liao, S. Zhu, X. Zhang, and D. Yang* 331

Adaptive Signal Processing

Estimation of Space-Time Varying Parameters Using a Diffusion LMS Algorithm http://dx.doi.org/10.1109/TSP.2013.2289888 ..... *R. Abdolee, B. Champagne, and A. H. Sayed* 403

Distributed Least Mean-Square Estimation With Partial Diffusion http://dx.doi.org/10.1109/TSP.2013.2292035 ..... *R. Arablouei, S. Werner, Y.-F. Huang, and K. Doğançay* 472

Digital and Multirate Signal Processing

Adaptive Identification and Recovery of Jointly Sparse Vectors http://dx.doi.org/10.1109/TSP.2013.2288679 ..... *R. Amel and A. Feuer* 354

Sampling Curves With Finite Rate of Innovation http://dx.doi.org/10.1109/TSP.2013.2292033 ..... *H. Pan, T. Blu, and P. L. Dragotti* 458

Machine Learning

Non-Negative Blind Source Separation Algorithm Based on Minimum Aperture Simplicial Cone http://dx.doi.org/10.1109/TSP.2013.2287683 ..... *W. S. B. Ouedraogo, A. Souloumiac, M. Jaidane, and C. Jutten* 376

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Sensor Array and Multichannel Processing

MIMO Radar Waveform Design With Constant Modulus and Similarity Constraints http://dx.doi.org/10.1109/TSP.2013.2288086 ..... *G. Cui, H. Li, and M. Rangaswamy* 343

Spatial Compressive Sensing for MIMO Radar http://dx.doi.org/10.1109/TSP.2013.2289875 .... *M. Rossi, A. M. Haimovich, and Y. C. Eldar* 419

Radar and Sonar Signal Processing

Matrix Completion in Colocated MIMO Radar: Recoverability, Bounds & Theoretical Guarantees http://dx.doi.org/10.1109/TSP.2013.2287673 ..... *D. S. Kalogerias and A. P. Petropulu* 309

Signal Processing for Communications

Time Domain Channel Estimation for OQAM-OFDM Systems: Algorithms and Performance Bounds http://dx.doi.org/10.1109/TSP.2013.2290498 ..... *D. Kong, D. Qu, and T. Jiang* 322

Fully Distributed Optimal Channel Assignment for Open Spectrum Access http://dx.doi.org/10.1109/TSP.2013.2285512 ..... *O. Naparstek and A. Leshem* 283

Energy Efficiency Optimization in Relay-Assisted MIMO Systems With Perfect and Statistical CSI http://dx.doi.org/10.1109/TSP.2013.2292031 ..... *A. Zappone, P. Cao, and E. A. Jorswieck* 443

MIMO Communications & Signal Processing

MIMO Systems With Quantized Covariance Feedback http://dx.doi.org/10.1109/TSP.2013.2292036 ..... *R. T. Krishnamachari, M. K. Varanasi, and K. Mohanty* 485

MIMO Detection by Lagrangian Dual Maximum-Likelihood Relaxation: Reinterpreting Regularized Lattice Decoding http://dx.doi.org/10.1109/TSP.2013.2292040 ..... *J. Pan, W.-K. Ma, and J. Jaldén* 511

Multi-Cell Beamforming With Decentralized Coordination in Cognitive and Cellular Networks http://dx.doi.org/10.1109/TSP.2013.2287681 ..... *H. Pennanen, A. Tölli, and M. Latva-aho* 295

Signal Processing for Sensor Networks

Asymptotically Optimal Truncated Multivariate Gaussian Hypothesis Testing With Application to Consensus Algorithms http://dx.doi.org/10.1109/TSP.2013.2288673 ..... *J. Zhang and R. S. Blum* 431

Energy Efficient Collaborative Beamforming in Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2288080 ..... *B. Béjar Haro, S. Zazo, and D. P. Palomar* 496

Signal Processing for Wireless Networks

Sensor Selection Based on Generalized Information Gain for Target Tracking in Large Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2289881 ..... *X. Shen and P. K. Varshney* 363

An MMSE Strategy at Relays With Partial CSI for a Multi-Layer Relay Network http://dx.doi.org/10.1109/TSP.2013.2284471 ..... *P. S. Elamvazhuthi, B. K. Dey, and S. Bhashyam* 271

Cache-Enabled Opportunistic Cooperative MIMO for Video Streaming in Wireless Systems http://dx.doi.org/10.1109/TSP.2013.2291211 .. *A. Liu and V. K. N. Lau* 390

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

FEBRUARY 1, 2014 VOLUME 62 NUMBER 3 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Variational Mode Decomposition http://dx.doi.org/10.1109/TSP.2013.2288675 ..... *K. Dragomiretskiy and D. Zosso* 531

Magnetometer Modeling and Validation for Tracking Metallic Targets http://dx.doi.org/10.1109/TSP.2013.2274639 ..... *N. Wahlström and F. Gustafsson* 545

Vector Perturbation Based on Symbol Scaling for Limited Feedback MISO Downlinks http://dx.doi.org/10.1109/TSP.2013.2291223 ..... *C. Masouros, M. Sellathurai, and T. Ratnarajah* 562

Channel Idle Time Statistics Based Spectrum Accessing Strategies With CSMA Based Primary Networks http://dx.doi.org/10.1109/TSP.2013.2291216 ..... *Y. Liu, N. Kundargi, and A. Tewfik* 572

On the Complexity of Joint Subcarrier and Power Allocation for Multi-User OFDMA Systems http://dx.doi.org/10.1109/TSP.2013.2293130 ..... *Y.-F. Liu and Y.-H. Dai* 583

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Ground Clutter Detection Using the Statistical Properties of Signals Received With a Polarimetric Radar http://dx.doi.org/10.1109/TSP.2013.2293118 ..... *Y. Li, G. Zhang, and R. J. Doviak* 597

Bi-Directional Training for Adaptive Beamforming and Power Control in Interference Networks http://dx.doi.org/10.1109/TSP.2013.2293122 ..... *C. Shi, R. A. Berry, and M. L. Honig* 607

High-Throughput Energy-Efficient LDPC Decoders Using Differential Binary Message Passing http://dx.doi.org/10.1109/TSP.2013.2293116 ..... *K. Cushon, S. Hemati, C. Leroux, S. Mannor, and W. J. Gross* 619

Novel Blind Identification of LDPC Codes Using Average LLR of Syndrome *a Posteriori* Probability http://dx.doi.org/10.1109/TSP.2013.2293975 ..... *T. Xia and H.-C. Wu* 632

Decomposition by Partial Linearization: Parallel Optimization of Multi-Agent Systems http://dx.doi.org/10.1109/TSP.2013.2293126 ..... *G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang* 641

Intrinsically Optimal Bayesian Robust Filtering http://dx.doi.org/10.1109/TSP.2013.2291213 ..... *L. A. Dalton and E. R. Dougherty* 657

Distributed Detection in Sensor Networks Over Fading Channels With Multiple Antennas at the Fusion Centre http://dx.doi.org/10.1109/TSP.2013.2293970 ..... *I. Nevat, G. W. Peters, and I. B. Collings* 671

Backhaul Limited Asymmetric Cooperation for MIMO Cellular Networks via Semidefinite Relaxation http://dx.doi.org/10.1109/TSP.2013.2293972 ..... *F. Zhuang and V. K. N. Lau* 684

High-Dimensional Screening Using Multiple Grouping of Variables http://dx.doi.org/10.1109/TSP.2013.2294591 ..... *D. Vats* 694

Locality Statistics for Anomaly Detection in Time Series of Graphs http://dx.doi.org/10.1109/TSP.2013.2294594 ..... *H. Wang, M. Tang, Y. Park, and C. E. Priebe* 703

Reducing Artifacts in JPEG Decompression Via a Learned Dictionary http://dx.doi.org/10.1109/TSP.2013.2290508 ..... *H. Chang, M. K. Ng, and T. Zeng* 718

Stochastic Ordering of Interference in Large-Scale Wireless Networks http://dx.doi.org/10.1109/TSP.2013.2293977 ..... *J. Lee and C. Tepedelenlioğlu* 729

Coordinated Multicell Multiuser Precoding for Maximizing Weighted Sum Energy Efficiency http://dx.doi.org/10.1109/TSP.2013.2294595 ..... *S. He, Y. Huang, L. Yang, and B. Ottersten* 741

CORRESPONDENCE

Optimal Memory for Discrete-Time FIR Filters in State-Space http://dx.doi.org/10.1109/TSP.2013.2290504 ..... *F. Ramirez-Echeverria, A. Sarr, and Y. S. Shmaliy* 557

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

FEBRUARY 1, 2014 VOLUME 62 NUMBER 3 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Statistical Signal Processing

Locality Statistics for Anomaly Detection in Time Series of Graphs http://dx.doi.org/10.1109/TSP.2013.2294594 ..... *H. Wang, M. Tang, Y. Park, and C. E. Priebe* 703

Magnetometer Modeling and Validation for Tracking Metallic Targets http://dx.doi.org/10.1109/TSP.2013.2274639 ..... *N. Wahlström and F. Gustafsson* 545

Adaptive Signal Processing

Variational Mode Decomposition http://dx.doi.org/10.1109/TSP.2013.2288675 ..... *K. Dragomiretskiy and D. Zosso* 531

Nonlinear Signal Processing

Reducing Artifacts in JPEG Decompression Via a Learned Dictionary http://dx.doi.org/10.1109/TSP.2013.2290508 ..... *H. Chang, M. K. Ng, and T. Zeng* 718

Multidimensional Signal Processing

Intrinsically Optimal Bayesian Robust Filtering http://dx.doi.org/10.1109/TSP.2013.2291213 ..... *L. A. Dalton and E. R. Dougherty* 657

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Machine Learning

High-Dimensional Screening Using Multiple Grouping of Variables http://dx.doi.org/10.1109/TSP.2013.2294591 ..... *D. Vats* 694

Signal Processing for Networks

Distributed Detection in Sensor Networks Over Fading Channels With Multiple Antennas at the Fusion Centre http://dx.doi.org/10.1109/TSP.2013.2293970 ..... *I. Nevat, G. W. Peters, and I. B. Collings* 671

Signal Processing for Communications

Stochastic Ordering of Interference in Large-Scale Wireless Networks http://dx.doi.org/10.1109/TSP.2013.2293977 ..... *J. Lee and C. Tepedelenlioğlu* 729

Novel Blind Identification of LDPC Codes Using Average LLR of Syndrome *a Posteriori* Probability http://dx.doi.org/10.1109/TSP.2013.2293975 ..... *T. Xia and H.-C. Wu* 632

Ground Clutter Detection Using the Statistical Properties of Signals Received With a Polarimetric Radar http://dx.doi.org/10.1109/TSP.2013.2293118 ..... *Y. Li, G. Zhang, and R. J. Doviak* 597

On the Complexity of Joint Subcarrier and Power Allocation for Multi-User OFDMA Systems http://dx.doi.org/10.1109/TSP.2013.2293130 ..... *Y.-F. Liu and Y.-H. Dai* 583

Channel Idle Time Statistics Based Spectrum Accessing Strategies With CSMA Based Primary Networks http://dx.doi.org/10.1109/TSP.2013.2291216 ..... *Y. Liu, N. Kundargi, and A. Tewfik* 572

Decomposition by Partial Linearization: Parallel Optimization of Multi-Agent Systems http://dx.doi.org/10.1109/TSP.2013.2293126 ..... *G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang* 641

Vector Perturbation Based on Symbol Scaling for Limited Feedback MISO Downlinks http://dx.doi.org/10.1109/TSP.2013.2291223 ..... *C. Masouros, M. Sellathurai, and T. Ratnarajah* 562

MIMO Communications & Signal Processing

Bi-Directional Training for Adaptive Beamforming and Power Control in Interference Networks http://dx.doi.org/10.1109/TSP.2013.2293122 ..... *C. Shi, R. A. Berry, and M. L. Honig* 607

Coordinated Multicell Multiuser Precoding for Maximizing Weighted Sum Energy Efficiency http://dx.doi.org/10.1109/TSP.2013.2294595 ..... *S. He, Y. Huang, L. Yang, and B. Ottersten* 741

Backhaul Limited Asymmetric Cooperation for MIMO Cellular Networks via Semidefinite Relaxation http://dx.doi.org/10.1109/TSP.2013.2293972 ..... *F. Zhuang and V. K. N. Lau* 684

Implementation of Signal Processing Systems

High-Throughput Energy-Efficient LDPC Decoders Using Differential Binary Message Passing http://dx.doi.org/10.1109/TSP.2013.2293116 ..... *K. Cushon, S. Hemati, C. Leroux, S. Mannor, and W. J. Gross* 619

CORRESPONDENCE

Digital and Multirate Signal Processing

Optimal Memory for Discrete-Time FIR Filters in State-Space http://dx.doi.org/10.1109/TSP.2013.2290504 ..... *F. Ramirez-Echeverria, A. Sarr, and Y. S. Shmaliy* 557

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

FEBRUARY 15, 2014 VOLUME 62 NUMBER 4 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Coalitional Games for Transmitter Cooperation in MIMO Multiple Access Channels http://dx.doi.org/10.1109/TSP.2013.2290496 ..... S. Yerramalli, R. Jain, and U. Mitra 757

A Doppler Robust Design of Transmit Sequence and Receive Filter in the Presence of Signal-Dependent Interference http://dx.doi.org/10.1109/TSP.2013.2288082 ..... M. M. Naghsh, M. Soltanalian, P. Stoica, M. Modarres-Hashemi, A. De Maio, and A. Aubry 772

Multi-Stream Transmission for Highly Frequency Selective Channels in MIMO-FBMC/OQAM Systems http://dx.doi.org/10.1109/TSP.2013.2293973 ..... M. Caus and A. I. Pérez-Neira 786

Distributed Finite-Horizon Fusion Kalman Filtering for Bandwidth and Energy Constrained Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2294603 ..... B. Chen, W.-A. Zhang, and L. Yu 797

Application of Manifold Separation to Polarimetric Capon Beamformer and Source Tracking http://dx.doi.org/10.1109/TSP.2013.2294598 ..... M. Costa and V. Koivunen 813

Multiuser Diversity for Secrecy Communications Using Opportunistic Jammer Selection: Secure DoF and Jammer Scaling Law http://dx.doi.org/10.1109/TSP.2013.2293979 ..... J. H. Lee and W. Choi 828

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Non-Maximally Decimated Analysis/Synthesis Filter Banks: Applications in Wideband Digital Filtering http://dx.doi.org/10.1109/TSP.2013.2295549 ..... *X. Chen, F. J. Harris, E. Venosa, and B. D. Rao* 852

Convergence Rates of Distributed Nesterov-Like Gradient Methods on Random Networks http://dx.doi.org/10.1109/TSP.2013.2291221 .... *D. Jakovetić, J. M. F. Xavier, and J. M. F. Moura* 868

Learning Overcomplete Dictionaries Based on Atom-by-Atom Updating http://dx.doi.org/10.1109/TSP.2013.2295062 ..... *M. Sadeghi, M. Babaie-Zadeh, and C. Jutten* 883

Maximum Likelihood Laplacian Correlation Channel Estimation in Layered Wyner-Ziv Coding http://dx.doi.org/10.1109/TSP.2013.2295556 ..... *N. Deligiannis, A. Munteanu, S. Wang, S. Cheng, and P. Schelkens* 892

Clustering on Multi-Layer Graphs via Subspace Analysis on Grassmann Manifolds http://dx.doi.org/10.1109/TSP.2013.2295553 ..... *X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov* 905

Decision Directed Estimation of Threshold Voltage Distribution in NAND Flash Memory http://dx.doi.org/10.1109/TSP.2013.2295056 ..... *D.-h. Lee and W. Sung* 919

GESPAR: Efficient Phase Retrieval of Sparse Signals http://dx.doi.org/10.1109/TSP.2013.2297687 ..... *Y. Shechtman, A. Beck, and Y. C. Eldar* 928

Maneuvering Target Detection via Radon-Fractional Fourier Transform-Based Long-Time Coherent Integration http://dx.doi.org/10.1109/TSP.2013.2297682 ..... *X. Chen, J. Guan, N. Liu, and Y. He* 939

Performance Analysis of a Power Limited Spectrum Sharing System With TAS/MRC http://dx.doi.org/10.1109/TSP.2013.2297684 ..... *F. A. Khan, K. Tourki, M.-S. Alouini, and K. A. Qaraqe* 954

On Using the Relative Configuration to Explore Cooperative Localization http://dx.doi.org/10.1109/TSP.2013.2297680 ..... *P. Zhang and Q. Wang* 968

Recovery of Low-Rank Matrices Under Affine Constraints via a Smoothed Rank Function http://dx.doi.org/10.1109/TSP.2013.2295557 .... *M. Malek-Mohammadi, M. Babaie-Zadeh, A. Amini, and C. Jutten* 981

Adaptive Limited Feedback for MISO Wiretap Channels With Cooperative Jamming http://dx.doi.org/10.1109/TSP.2013.2295059 ..... *M. Pei, A. L. Swindlehurst, D. Ma, and J. Wei* 993

Energy-Efficient CoMP Precoding in Heterogeneous Networks http://dx.doi.org/10.1109/TSP.2013.2296279 ..... *Z. Xu, C. Yang, G. Y. Li, Y. Liu, and S. Xu* 1005

List of Reviewers http://dx.doi.org/10.1109/TSP.2013.2294890 1018

CORRECTIONS

Corrections to "Vector Cross-Product Direction-Finding' With an Electromagnetic Vector-Sensor of Six Orthogonally Oriented But Spatially Noncollocating Dipoles/Loops" http://dx.doi.org/10.1109/TSP.2013.2290501 .... *Y. Song, X. Yuan, and K. T. Wong* 1028

EDICS—Editor's Information Classification Scheme http://dx.doi.org/10.1109/TSP.2014.2302616 1031

Information for Authors http://dx.doi.org/10.1109/TSP.2014.2302617 1032

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed | MEDLINE U.S. National Library of Medicine

FEBRUARY 15, 2014 VOLUME 62 NUMBER 4 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Design and Implementation of Signal Processing System

Decision Directed Estimation of Threshold Voltage Distribution in NAND Flash Memory http://dx.doi.org/10.1109/TSP.2013.2295056 ..... *D.-h. Lee and W. Sung* 919

Digital and Multirate Signal Processing

Non-Maximally Decimated Analysis/Synthesis Filter Banks: Applications in Wideband Digital Filtering http://dx.doi.org/10.1109/TSP.2013.2295549 ..... *X. Chen, F. J. Harris, E. Venosa, and B. D. Rao* 852

GESPAR: Efficient Phase Retrieval of Sparse Signals http://dx.doi.org/10.1109/TSP.2013.2297687 ..... *Y. Shechtman, A. Beck, and Y. C. Eldar* 928

Recovery of Low-Rank Matrices Under Affine Constraints via a Smoothed Rank Function http://dx.doi.org/10.1109/TSP.2013.2295557 ..... *M. Malek-Mohammadi, M. Babaie-Zadeh, A. Amini, and C. Jutten* 981

Implementation of Signal Processing Systems

Low-Complexity Motion-Compensated Beamforming Algorithm and Architecture for Synthetic Transmit Aperture in Ultrasound Imaging http://dx.doi.org/10.1109/TSP.2013.2295551 ..... *Y.-H. Chen, Y.-M. Lin, K.-Y. Ho, A.-Y. Wu, and P.-C. Li* 840

Machine Learning

Learning Overcomplete Dictionaries Based on Atom-by-Atom Updating http://dx.doi.org/10.1109/TSP.2013.2295062 ..... *M. Sadeghi, M. Babaie-Zadeh, and C. Jutten* 883

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information**: Individual copies: IEEE Members $20.00 (first copy only), nonmembers $569.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

MIMO Communications & Signal Processing

Adaptive Limited Feedback for MISO Wiretap Channels With Cooperative Jamming http://dx.doi.org/10.1109/TSP.2013.2295059 ..... *M. Pei, A. L. Swindlehurst, D. Ma, and J. Wei* 993

Multi-Stream Transmission for Highly Frequency Selective Channels in MIMO-FBMC/OQAM Systems http://dx.doi.org/10.1109/TSP.2013.2293973 ..... *M. Caus and A. I. Pérez-Neira* 786

Signal Processing for Networks

Clustering on Multi-Layer Graphs via Subspace Analysis on Grassmann Manifolds http://dx.doi.org/10.1109/TSP.2013.2295553 ..... *X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov* 905

Sensor Array and Multichannel Processing

Application of Manifold Separation to Polarimetric Capon Beamformer and Source Tracking http://dx.doi.org/10.1109/TSP.2013.2294598 ..... *M. Costa and V. Koivunen* 813

Maneuvering Target Detection via Radon-Fractional Fourier Transform-Based Long-Time Coherent Integration http://dx.doi.org/10.1109/TSP.2013.2297682 ..... *X. Chen, J. Guan, N. Liu, and Y. He* 939

Signal Processing for Sensor Networks

Maximum Likelihood Laplacian Correlation Channel Estimation in Layered Wyner-Ziv Coding http://dx.doi.org/10.1109/TSP.2013.2295556 ..... *N. Deligiannis, A. Munteanu, S. Wang, S. Cheng, and P. Schelkens* 892

Convergence Rates of Distributed Nesterov-Like Gradient Methods on Random Networks http://dx.doi.org/10.1109/TSP.2013.2291221 ..... *D. Jakovetić, J. M. F. Xavier, and J. M. F. Moura* 868

On Using the Relative Configuration to Explore Cooperative Localization http://dx.doi.org/10.1109/TSP.2013.2297680 ..... *P. Zhang and Q. Wang* 968

Signal Processing for Communications

Multiuser Diversity for Secrecy Communications Using Opportunistic Jammer Selection: Secure DoF and Jammer Scaling Law http://dx.doi.org/10.1109/TSP.2013.2293979 ..... *J. H. Lee and W. Choi* 828

Energy-Efficient CoMP Precoding in Heterogeneous Networks http://dx.doi.org/10.1109/TSP.2013.2296279 ..... *Z. Xu, C. Yang, G. Y. Li, Y. Liu, and S. Xu* 1005

Performance Analysis of a Power Limited Spectrum Sharing System With TAS/MRC http://dx.doi.org/10.1109/TSP.2013.2297684 ..... *F. A. Khan, K. Tourki, M.-S. Alouini, and K. A. Qaraqe* 954

Statistical Signal Processing

A Doppler Robust Design of Transmit Sequence and Receive Filter in the Presence of Signal-Dependent Interference http://dx.doi.org/10.1109/TSP.2013.2288082 ..... *M. M. Naghsh, M. Soltanalian, P. Stoica, M. Modarres-Hashemi, A. De Maio, and A. Aubry* 772

Distributed Finite-Horizon Fusion Kalman Filtering for Bandwidth and Energy Constrained Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2013.2294603 ..... *B. Chen, W.-A. Zhang, and L. Yu* 797

Signal Processing for Wireless Networks

Coalitional Games for Transmitter Cooperation in MIMO Multiple Access Channels http://dx.doi.org/10.1109/TSP.2013.2290496 ..... *S. Yerramalli, R. Jain, and U. Mitra* 757

List of Reviewers http://dx.doi.org/10.1109/TSP.2013.2294890 1018

CORRECTIONS

Corrections to "Vector Cross-Product Direction-Finding' With an Electromagnetic Vector-Sensor of Six Orthogonally Oriented But Spatially Noncollocating Dipoles/Loops" http://dx.doi.org/10.1109/TSP.2013.2290501 .... *Y. Song, X. Yuan, and K. T. Wong* 1028

IEEE/ACM TRANSACTIONS ON

AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society | acm Association for Computing Machinery | www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed | MEDLINE U.S. National Library of Medicine

JANUARY 2014 | VOLUME 22 | NUMBER 1 | ITASFA | (ISSN 2329-9290)

EDITORIAL

Editorial: Expanding the Technical Reach of our Transactions http://dx.doi.org/10.1109/TASLP.2013.2290172 ..... *L. Deng, S. Renals, M. Federico, and M. Ostendorf* 5

REGULAR PAPERS

Objective Intelligibility Measures Based on Mutual Information for Speech Subjected to Speech Enhancement Processing http://dx.doi.org/10.1109/TASL.2013.2281574 ..... *J. Taghia and R. Martin* 6

Cross-Lingual Subspace Gaussian Mixture Models for Low-Resource Speech Recognition http://dx.doi.org/10.1109/TASL.2013.2281575 ..... *L. Lu, A. Ghoshal, and S. Renals* 17

Gaussian Processes for POMDP-Based Dialogue Manager Optimization http://dx.doi.org/10.1109/TASL.2013.2282190 ..... *M. Gašić and S. Young* 28

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 1558-7916) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $307.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Nonlinear Audio Systems Identification Through Audio Input Gaussianization http://dx.doi.org/10.1109/TASL.2013.2282214 ..... I. Mezghani-Marrakchi, G. Mahé, S. Djaziri-Larbi, M. Jaïdane, and M. Turki-Hadj Alouane 41
Multizone Speech Reinforcement http://dx.doi.org/10.1109/TASL.2013.2283100 ..... J. B. Crespo and R. C. Hendriks 54
Performance Study of the MVDR Beamformer as a Function of the Source Incidence Angle http://dx.doi.org/10.1109/TASL.2013.2283104 ..... C. Pan, J. Chen, and J. Benesty 67
Improved Semantic Retrieval of Spoken Content by Document/Query Expansion with Random Walk Over Acoustic Similarity Graphs http://dx.doi.org/10.1109/TASLP.2013.2285469 ..... H.-Y. Lee and L.-S. Lee 80
A New Observation Model in the Logarithmic Mel Power Spectral Domain for the Automatic Recognition of Noisy Reverberant Speech http://dx.doi.org/10.1109/TASLP.2013.2285480 ..... V. Leutnant, A. Krueger, and R. Haeb-Umbach 95
Characterizing Phonetic Transformations and Acoustic Differences Across English Dialects http://dx.doi.org/10.1109/TASLP.2013.2285482 ..... N. F. Chen, S. W. Tam, W. Shen, and J. P. Campbell 110
Estimation of Acoustic Reflection Coefficients Through Pseudospectrum Matching http://dx.doi.org/10.1109/TASLP.2013.2285483 ..... D. Marković, K. Kowalczyk, F. Antonacci, C. Hofmann, A. Sarti, and W. Kellermann 125
Multi-pitch Streaming of Harmonic Sound Mixtures http://dx.doi.org/10.1109/TASLP.2013.2285484 ..... Z. Duan, J. Han, and B. Pardo 138
Temporally Varying Weight Regression: A Semi-Parametric Trajectory Model for Automatic Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2285487 ..... S. Liu and K. C. Sim 151
A Family of Discriminative Manifold Learning Algorithms and Their Application to Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2286906 ..... V. S. Tomar and R. C. Rose 161
Alaryngeal Speech Enhancement Based on One-to-Many Eigenvoice Conversion http://dx.doi.org/10.1109/TASLP.2013.2286917 ..... H. Doi, T. Toda, K. Nakamura, H. Saruwatari, and K. Shikano 172
Converting Neural Network Language Models into Back-off Language Models for Efficient Decoding in Automatic Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2286919 ..... E. Arısoy, S. F. Chen, B. Ramabhadran, and A. Sethy 184
Design, Optimization and Evaluation of a Dual-Radius Spherical Microphone Array http://dx.doi.org/10.1109/TASLP.2013.2286920 ..... C. T. Jin, N. Epain, and A. Parthy 193
Low Frequency Interpolation of Room Impulse Responses Using Compressed Sensing http://dx.doi.org/10.1109/TASLP.2013.2286922 ..... R. Mignot, G. Chardon, and L. Daudet 205
A Study of the Cosine Distance-Based Mean Shift for Telephone Speech Diarization http://dx.doi.org/10.1109/TASLP.2013.2285474 ..... M. Senoussaoui, P. Kenny, T. Stafylakis, and P. Dumouchel 217
Singing Voice Enhancement in Monaural Music Signals Based on Two-stage Harmonic/Percussive Sound Separation on Multiple Resolution Spectrograms http://dx.doi.org/10.1109/TASLP.2013.2287052 ..... H. Tachibana, N. Ono, and S. Sagayama 228
Generalized Spherical Array Beamforming for Binaural Speech Reproduction http://dx.doi.org/10.1109/TASLP.2013.2290499 ..... N. R. Shabtai and B. Rafaely 238
Factorized Sub-Space Estimation for Fast and Memory Effective I-vector Extraction http://dx.doi.org/10.1109/TASLP.2013.2290505 ..... S. Cumani and P. Laface 248
Distributed Delay and Sum Beamformer for Speech Enhancement via Randomized Gossip http://dx.doi.org/10.1109/TASLP.2013.2290861 ..... Y. Zeng and R. C. Hendriks 260
Joint Optimization for Chinese POS Tagging and Dependency Parsing http://dx.doi.org/10.1109/TASLP.2013.2288081 ..... Z. Li, M. Zhang, W. Che, T. Liu, and W. Chen 274

IEEE/ACM TRANSACTIONS ON

AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society | acm Association for Computing Machinery | www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed | MEDLINE U.S. National Library of Medicine

JANUARY 2014 | VOLUME 22 | NUMBER 1 | ITASFA | (ISSN 2329-9290)

EDITORIAL

Editorial: Expanding the Technical Reach of our Transactions http://dx.doi.org/10.1109/TASLP.2013.2290172 ..... *L. Deng, S. Renals, M. Federico, and M. Ostendorf* 5

REGULAR PAPERS

Room Acoustics and Acoustic System Modeling

Low Frequency Interpolation of Room Impulse Responses Using Compressed Sensing http://dx.doi.org/10.1109/TASLP.2013.2286922 ..... *R. Mignot, G. Chardon, and L. Daudet* 205

Loudspeaker and Microphone Array Signal Processing

Estimation of Acoustic Reflection Coefficients Through Pseudospectrum Matching http://dx.doi.org/10.1109/TASLP.2013.2285483 ..... *D. Marković, K. Kowalczyk, F. Antonacci, C. Hofmann, A. Sarti, and W. Kellermann* 125

Design, Optimization and Evaluation of a Dual-Radius Spherical Microphone Array http://dx.doi.org/10.1109/TASLP.2013.2286920 ..... *C. T. Jin, N. Epain, and A. Parthy* 193

Acoustic Sensor Array Processing

Performance Study of the MVDR Beamformer as a Function of the Source Incidence Angle http://dx.doi.org/10.1109/TASL.2013.2283104 .. *C. Pan, J. Chen, and J. Benesty* 67

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 1558-7916) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $307.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Audio Source Separation and Enhancement

Singing Voice Enhancement in Monaural Music Signals Based on Two-stage Harmonic/Percussive Sound Separation on Multiple Resolution Spectrograms http://dx.doi.org/10.1109/TASLP.2013.2287052 ..... *H. Tachibana, N. Ono, and S. Sagayama* 228

Spatial and Multichannel Audio

Generalized Spherical Array Beamforming for Binaural Speech Reproduction http://dx.doi.org/10.1109/TASLP.2013.2290499 ..... *N. R. Shabtai and B. Rafaely* 238

Audio Analysis and Synthesis

Multi-pitch Streaming of Harmonic Sound Mixtures http://dx.doi.org/10.1109/TASLP.2013.2285484 ..... *Z. Duan, J. Han, and B. Pardo* 138

Audio and Acoustic Signal Processing

Nonlinear Audio Systems Identification Through Audio Input Gaussianization http://dx.doi.org/10.1109/TASL.2013.2282214 ..... *I. Mezghani-Marrakchi, G. Mahé, S. Djaziri-Larbi, M. Jaïdane, and M. Turki-Hadj Alouane* 41

Speech Perception and Psychoacoustics

Objective Intelligibility Measures Based on Mutual Information for Speech Subjected to Speech Enhancement Processing http://dx.doi.org/10.1109/TASL.2013.2281574 ..... *J. Taghia and R. Martin* 6

Speech Synthesis and Generation

Alaryngeal Speech Enhancement Based on One-to-Many Eigenvoice Conversion http://dx.doi.org/10.1109/TASLP.2013.2286917 ..... *H. Doi, T. Toda, K. Nakamura, H. Saruwatari, and K. Shikano* 172

Speech Enhancement

Multizone Speech Reinforcement http://dx.doi.org/10.1109/TASL.2013.2283100 ..... *J. B. Crespo and R. C. Hendriks* 54

Distributed Delay and Sum Beamformer for Speech Enhancement via Randomized Gossip http://dx.doi.org/10.1109/TASLP.2013.2290861 ..... *Y. Zeng and R. C. Hendriks* 260

Acoustic Modeling for Automatic Speech Recognition

Temporally Varying Weight Regression: A Semi-Parametric Trajectory Model for Automatic Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2285487 ..... *S. Liu and K. C. Sim* 151

A Family of Discriminative Manifold Learning Algorithms and Their Application to Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2286906 ..... *V. S. Tomar and R. C. Rose* 161

Robust Speech Recognition

A New Observation Model in the Logarithmic Mel Power Spectral Domain for the Automatic Recognition of Noisy Reverberant Speech http://dx.doi.org/10.1109/TASLP.2013.2285480 ..... *V. Leutnant, A. Krueger, and R. Haeb-Umbach* 95

Multilingual Recognition and Identification

Characterizing Phonetic Transformations and Acoustic Differences Across English Dialects http://dx.doi.org/10.1109/TASLP.2013.2285482 ..... *N. F. Chen, S. W. Tam, W. Shen, and J. P. Campbell* 110

Speaker Recognition and Characterization

A Study of the Cosine Distance-Based Mean Shift for Telephone Speech Diarization http://dx.doi.org/10.1109/TASLP.2013.2285474 ..... *M. Senoussaoui, P. Kenny, T. Stafylakis, and P. Dumouchel* 217

Factorized Sub-Space Estimation for Fast and Memory Effective I-vector Extraction http://dx.doi.org/10.1109/TASLP.2013.2290505 ..... *S. Cumani and P. Laface* 248

Resource Constrained Speech Recognition

Cross-Lingual Subspace Gaussian Mixture Models for Low-Resource Speech Recognition http://dx.doi.org/10.1109/TASL.2013.2281575 ..... *L. Lu, A. Ghoshal, and S. Renals* 17

Human Language Acquisition, Development, and Learning

Joint Optimization for Chinese POS Tagging and Dependency Parsing http://dx.doi.org/10.1109/TASLP.2013.2288081 ..... *Z. Li, M. Zhang, W. Che, T. Liu, and W. Chen* 274

Language Modeling

Converting Neural Network Language Models into Back-off Language Models for Efficient Decoding in Automatic Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2286919 ..... *E. Arısoy, S. F. Chen, B. Ramabhadran, and A. Sethy* 184

Discourse and Dialog

Gaussian Processes for POMDP-Based Dialogue Manager Optimization http://dx.doi.org/10.1109/TASL.2013.2282190 ..... *M. Gašić and S. Young* 28

Speech Data Mining and Document Retrieval

Improved Semantic Retrieval of Spoken Content by Document/Query Expansion with Random Walk Over Acoustic Similarity Graphs http://dx.doi.org/10.1109/TASLP.2013.2285469 ..... *H.-Y. Lee and L.-S. Lee* 80

IEEE/ACM TRANSACTIONS ON

AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society | acm Association for Computing Machinery | www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed | MEDLINE U.S. National Library of Medicine

FEBRUARY 2014 | VOLUME 22 | NUMBER 2 | ITASFA | (ISSN 2329-9290)

OVERVIEW ARTICLE

Automatic Chord Estimation from Audio: A Review of the State of the Art http://dx.doi.org/10.1109/TASLP.2013.2294580 ..... M. McVicar, R. Santos-Rodríguez, Y. Ni, and T. D. Bie 556

REGULAR PAPERS

Sequential Summarization: A Full View of Twitter Trending Topics http://dx.doi.org/10.1109/TASL.2013.2282191 ..... D. Gao, W. Li, X. Cai, R. Zhang, and Y. Ouyang 293

A Comparison of Spectro-Temporal Representations of Audio Signals http://dx.doi.org/10.1109/TASL.2013.2283105 ..... P. W. J. van Hengel and J. D. Krijnders 303

Aligned-Parallel-Corpora Based Semi-Supervised Learning for Arabic Mention Detection http://dx.doi.org/10.1109/TASLP.2013.2287055 ..... I. Zitouni and Y. Benajiba 314

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 1558-7916) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $307.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Dissonance Reduction In Polyphonic Audio Using Harmonic Reorganization http://dx.doi.org/10.1109/TASLP.2013.2287056 ..... E. Molina, A. M. Barbancho, L. J. Tardón, and I. Barbancho 325
A Novel Expectation-Maximization Framework for Speech Enhancement in Non-Stationary Noise Environments http://dx.doi.org/10.1109/TASLP.2013.2290497 ..... D. P. K. Lun, T.-W. Shen, and K. C. Ho 335
Cochlear Implant Filterbank Design and Optimization: A Simulation Study http://dx.doi.org/10.1109/TASLP.2013.2290502 ..... S. Cosentino, T. H. Falk, D. McAlpine, and T. Marquardt 347
Location Feature Integration for Clustering-Based Speech Separation in Distributed Microphone Arrays http://dx.doi.org/10.1109/TASLP.2013.2292308 ..... M. Souden, K. Kinoshita, M. Delcroix, and T. Nakatani 354
Estimating Uncertainty to Improve Exemplar-Based Feature Enhancement for Noise Robust Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2292328 ..... H. Kallasjoki, J. F. Gemmeke, and K. J. Palomäki 368
Maximum Likelihood Acoustic Factor Analysis Models for Robust Speaker Verification in Noise http://dx.doi.org/10.1109/TASLP.2013.2292356 ..... T. Hasan and J. H. L. Hansen 381
Speaker Tracking Using Recursive EM Algorithms http://dx.doi.org/10.1109/TASLP.2013.2292361 ..... O. Schwartz and S. Gannot 392
A MAP-based Online Estimation Approach to Ensemble Speaker and Speaking Environment Modeling http://dx.doi.org/10.1109/TASLP.2013.2292362 ..... Y. Tsao, S. Matsuda, C. Hori, H. Kashioka, and C.-H. Lee 403
Latent Semantic Analysis for Multimodal User Input With Speech and Gestures http://dx.doi.org/10.1109/TASLP.2013.2294586 ..... P.-Y. Hui and H. Meng 417
Speech Intelligibility Prediction Based on Mutual Information http://dx.doi.org/10.1109/TASLP.2013.2295914 ..... J. Jensen and C. H. Taal 430
Objective and Subjective Investigation on a Novel Method for Digital Reverberator Parameters Estimation http://dx.doi.org/10.1109/TASLP.2013.2295925 ..... A. Primavera, S. Cecchi, J. Li, and F. Piazza 441
Modeling the Vocal Tract Transfer Function Using a 3D Digital Waveguide Mesh http://dx.doi.org/10.1109/TASLP.2013.2294579 ..... M. Speed, D. Murphy, and D. Howard 453
Theoretical Analysis of Open Spherical Microphone Arrays for Acoustic Intensity Measurements http://dx.doi.org/10.1109/TASLP.2013.2294581 ..... H. Hacıhabiboğlu 465
On the Relative Importance of Individual Components of Chord Recognition Systems http://dx.doi.org/10.1109/TASLP.2013.2295926 ..... T. Cho and J. P. Bello 477
Bayesian Nonparametrics for Microphone Array Processing http://dx.doi.org/10.1109/TASLP.2013.2294582 ..... T. Otsuka, K. Ishiguro, H. Sawada, and H. G. Okuno 493
Linear Estimation Based Primary-Ambient Extraction for Stereo Audio Signals http://dx.doi.org/10.1109/TASLP.2013.2297015 ..... J. He, E.-L. Tan, and W.-S. Gan 505
PEFAC - A Pitch Estimation Algorithm Robust to High Levels of Noise http://dx.doi.org/10.1109/TASLP.2013.2295918 ..... S. Gonzalez and M. Brookes 518
Bayesian Constituent Context Model for Grammar Induction http://dx.doi.org/10.1109/TASLP.2013.2294584 ..... M. Zhang, X. Duan, and W. Chen 531
Feedforward Active Noise Control With a New Variable Tap-Length and Step-Size Filtered-X LMS Algorithm http://dx.doi.org/10.1109/TASLP.2013.2297016 ..... D.-C. Chang and F.-T. Chu 542

EDICS—Editor's Information and Classification Scheme http://dx.doi.org/10.1109/TASLP.2014.2300358 576
Information for Authors http://dx.doi.org/10.1109/TASLP.2014.2300357 578

IEEE/ACM TRANSACTIONS ON

AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

acm Association for Computing Machinery

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

FEBRUARY 2014 VOLUME 22 NUMBER 2 ITASFA (ISSN 2329-9290)

OVERVIEW ARTICLE

Automatic Chord Estimation from Audio: A Review of the State of the Art http://dx.doi.org/10.1109/TASLP.2013.2294580 ..... *M. McVicar, R. Santos-Rodríguez, Y. Ni, and T. D. Bie* 556

REGULAR PAPERS

Loudspeaker and Microphone Array Signal Processing

Theoretical Analysis of Open Spherical Microphone Arrays for Acoustic Intensity Measurements http://dx.doi.org/10.1109/TASLP.2013.2294581 ..... *H. Hacıhabiboğlu* 465

Active Noise Control

Feedforward Active Noise Control With a New Variable Tap-Length and Step-Size Filtered-X LMS Algorithm http://dx.doi.org/10.1109/TASLP.2013.2297016 ..... *D.-C. Chang and F.-T. Chu* 542

Auditory Modeling and Hearing Aids

A Comparison of Spectro-Temporal Representations of Audio Signals http://dx.doi.org/10.1109/TASL.2013.2283105 ..... *P. W. J. van Hengel and J. D. Krijnders* 303

IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 1558-7916) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $307.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of $31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage Paid at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

IEEE

Audio Source Separation and Enhancement

Location Feature Integration for Clustering-Based Speech Separation in Distributed Microphone Arrays http://dx.doi.org/10.1109/TASLP.2013.2292308 ..... *M. Souden, K. Kinoshita, M. Delcroix, and T. Nakatani* 354

Bayesian Nonparametrics for Microphone Array Processing http://dx.doi.org/10.1109/TASLP.2013.2294582 ..... *T. Otsuka, K. Ishiguro, H. Sawada, and H. G. Okuno* 493

Spatial and Multichannel Audio

Objective and Subjective Investigation on a Novel Method for Digital Reverberator Parameters Estimation http://dx.doi.org/10.1109/TASLP.2013.2295925 ..... *A. Primavera, S. Cecchi, J. Li, and F. Piazza* 441

Linear Estimation Based Primary-Ambient Extraction for Stereo Audio Signals http://dx.doi.org/10.1109/TASLP.2013.2297015 ..... *J. He, E.-L. Tan, and W.-S. Gan* 505

Audio Analysis and Synthesis

Dissonance Reduction In Polyphonic Audio Using Harmonic Reorganization http://dx.doi.org/10.1109/TASLP.2013.2287056 ..... *E. Molina, A. M. Barbancho, L. J. Tardón, and I. Barbancho* 325

Content-Based Music Processing

On the Relative Importance of Individual Components of Chord Recognition Systems http://dx.doi.org/10.1109/TASLP.2013.2295926 ..... *T. Cho and J. P. Bello* 477

Speech Production

Modeling the Vocal Tract Transfer Function Using a 3D Digital Waveguide Mesh http://dx.doi.org/10.1109/TASLP.2013.2294579 ..... *M. Speed, D. Murphy, and D. Howard* 453

Speech Analysis

Speech Intelligibility Prediction Based on Mutual Information http://dx.doi.org/10.1109/TASLP.2013.2295914 ..... *J. Jensen and C. H. Taal* 430

PEFAC - A Pitch Estimation Algorithm Robust to High Levels of Noise http://dx.doi.org/10.1109/TASLP.2013.2295918 ..... *S. Gonzalez and M. Brookes* 518

Speech Enhancement

A Novel Expectation-Maximization Framework for Speech Enhancement in Non-Stationary Noise Environments http://dx.doi.org/10.1109/TASLP.2013.2290497 ..... *D. P. K. Lun, T.-W. Shen, and K. C. Ho* 335

Robust Speech Recognition

Estimating Uncertainty to Improve Exemplar-Based Feature Enhancement for Noise Robust Speech Recognition http://dx.doi.org/10.1109/TASLP.2013.2292328 ..... *H. Kallasjoki, J. F. Gemmeke, and K. J. Palomäki* 368

A MAP-based Online Estimation Approach to Ensemble Speaker and Speaking Environment Modeling http://dx.doi.org/10.1109/TASLP.2013.2292362 ..... *Y. Tsao, S. Matsuda, C. Hori, H. Kashioka, and C.-H. Lee* 403

Speaker Recognition and Characterization

Maximum Likelihood Acoustic Factor Analysis Models for Robust Speaker Verification in Noise http://dx.doi.org/10.1109/TASLP.2013.2292356 ..... *T. Hasan and J. H. L. Hansen* 381

Spoken Language Understanding

Aligned-Parallel-Corpora Based Semi-Supervised Learning for Arabic Mention Detection http://dx.doi.org/10.1109/TASLP.2013.2287055 ..... *I. Zitouni and Y. Benajiba* 314

Cochlear Implant Filterbank Design and Optimization: A Simulation Study http://dx.doi.org/10.1109/TASLP.2013.2290502 ..... *S. Cosentino, T. H. Falk, D. McAlpine, and T. Marquardt* 347

Spoken and Multimodal Dialog Systems

Latent Semantic Analysis for Multimodal User Input With Speech and Gestures http://dx.doi.org/10.1109/TASLP.2013.2294586 ..... *P.-Y. Hui and H. Meng* 417

Spoken Document Retrieval and Text Mining

Sequential Summarization: A Full View of Twitter Trending Topics http://dx.doi.org/10.1109/TASL.2013.2282191 ..... *D. Gao, W. Li, X. Cai, R. Zhang, and Y. Ouyang* 293

Human Language Acquisition, Development, and Learning

Bayesian Constituent Context Model for Grammar Induction http://dx.doi.org/10.1109/TASLP.2013.2294584 ..... *M. Zhang, X. Duan, and W. Chen* 531

Acoustic Sensor Array Processing

Speaker Tracking Using Recursive EM Algorithms http://dx.doi.org/10.1109/TASLP.2013.2292361 ..... *O. Schwartz and S. Gannot* 392

EDICS—Editor's Information and Classification Scheme http://dx.doi.org/10.1109/TASLP.2014.2300358 576

Information for Authors http://dx.doi.org/10.1109/TASLP.2014.2300357 578

IEEE TRANSACTIONS ON

IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

JANUARY 2014 VOLUME 23 NUMBER 1 IIPRE4 (ISSN 1057-7149)

PAPERS

Vector-Valued Image Processing by Parallel Level Sets http://dx.doi.org/10.1109/TIP.2013.2277775 ..... *M. J. Ehrhardt and S. R. Arridge* 9
Saliency-Aware Video Compression http://dx.doi.org/10.1109/TIP.2013.2282897 ..... *H. Hadizadeh and I. V. Bajić* 19
Metamer Mismatching http://dx.doi.org/10.1109/TIP.2013.2283148 ..... *A. D. Logvinenko, B. Funt, and C. Godau* 34
Translation Invariant Directional Framelet Transform Combined With Gabor Filters for Image Denoising http://dx.doi.org/10.1109/TIP.2013.2285595 ..... *Y. Shi, X. Yang, and Y. Guo* 44
Motion-Related Resource Allocation in Dynamic Wireless Visual Sensor Network Environments http://dx.doi.org/10.1109/TIP.2013.2286323 ..... *A. V. Katsenou, L. P. Kondi, and K. E. Parsopoulos* 56
Harmonic Active Contours http://dx.doi.org/10.1109/TIP.2013.2286326 ..... *V. Estellers, D. Zosso, X. Bresson, and J.-P. Thiran* 69
Multi-Illuminant Estimation With Conditional Random Fields http://dx.doi.org/10.1109/TIP.2013.2286327 ..... *S. Beigpour, C. Riess, J. van de Weijer, and E. Angelopoulou* 83
Real-Time Scalable Depth Sensing With Hybrid Structured Light Illumination http://dx.doi.org/10.1109/TIP.2013.2286901 ..... *Y. Zhang, Z. Xiong, Z. Yang, and F. Wu* 97
Integrated Segmentation and Interpolation of Sparse Data http://dx.doi.org/10.1109/TIP.2013.2286903 ..... *A. Paiement, M. Mirmehdi, X. Xie, and M. C. K. Hamilton* 110
Linear Time Distances Between Fuzzy Sets With Applications to Pattern Matching and Classification http://dx.doi.org/10.1109/TIP.2013.2286904 ..... *J. Lindblad and N. Sladoje* 126
A Bit Allocation Method for Sparse Source Coding http://dx.doi.org/10.1109/TIP.2013.2286325 ..... *M. Kaaniche, A. Fraysse, B. Pesquet-Popescu, and J.-C. Pesquet* 137
Spatial Statistics of Image Features for Performance Comparison http://dx.doi.org/10.1109/TIP.2013.2286907 ..... *E. Bostanci, N. Kanwal, and A. F. Clark* 153
Shape-Based Normalized Cuts Using Spectral Relaxation for Biomedical Segmentation http://dx.doi.org/10.1109/TIP.2013.2287604 ..... *E. Ruiz Pujadas and M. Reisert* 163
Optimal Camera Planning Under Versatile User Constraints in Multi-Camera Image Processing Systems http://dx.doi.org/10.1109/TIP.2013.2287606 ..... *J. Liu, S. Sridharan, C. Fookes, and T. Wark* 171

IEEE

An Analytical Model for Synthesis Distortion Estimation in 3D Video http://dx.doi.org/10.1109/TIP.2013.2287608 ..... L. Fang, N.-M. Cheung, D. Tian, A. Vetro, H. Sun, and O. C. Au 185
Attention Driven Foveated Video Quality Assessment http://dx.doi.org/10.1109/TIP.2013.2287611 ..... J. You, T. Ebrahimi, and A. Perkis 200
A Weighted Optimization Approach to Time-of-Flight Sensor Fusion http://dx.doi.org/10.1109/TIP.2013.2287613 ..... S. Schwarz, M. Sjöström, and R. Olsson 214
A Novel Eye Localization Method With Rotation Invariance http://dx.doi.org/10.1109/TIP.2013.2287614 ..... Y. Ren, S. Wang, B. Hou, and J. Ma 226
Directionlets Using In-Phase Lifting for Image Representation http://dx.doi.org/10.1109/TIP.2013.2288912 ..... D. Jayachandra and A. Makur 240
Camera Model Identification Based on the Heteroscedastic Noise Model http://dx.doi.org/10.1109/TIP.2013.2290596 ..... T. H. Thai, R. Cogranne, and F. Retraint 250
Multispectral Image Denoising With Optimized Vector Bilateral Filter http://dx.doi.org/10.1109/TIP.2013.2287612 ..... H. Peng, R. Rao, and S. A. Dianat 264
Chroma Intra Prediction Based on Inter-Channel Correlation for HEVC http://dx.doi.org/10.1109/TIP.2013.2288007 ..... X. Zhang, C. Gisquet, E. François, F. Zou, and O. C. Au 274
Human Detection by Quadratic Classification on Subspace of Extended Histogram of Gradients http://dx.doi.org/10.1109/TIP.2013.2264677 ..... A. Satpathy, X. Jiang, and H.-L. Eng 287
Variational Exemplar-Based Image Colorization http://dx.doi.org/10.1109/TIP.2013.2288929 ..... A. Bugeau, V.-T. Ta, and N. Papadakis 298
How to Select Good Neighboring Images in Depth-Map Merging Based 3D Modeling http://dx.doi.org/10.1109/TIP.2013.2290597 ..... S. Shen and Z. Hu 308
Motion-Aware Mesh-Structured Trellis for Correlation Modelling Aided Distributed Multi-View Video Coding http://dx.doi.org/10.1109/TIP.2013.2288913 ..... Y. Huo, T. Wang, R. G. Maunder, and L. Hanzo 319
Filtered Backprojection Reconstruction and Redundancy in Compton Camera Imaging http://dx.doi.org/10.1109/TIP.2013.2288143 ..... V. Maxim 332
Seamless View Synthesis Through Texture Optimization http://dx.doi.org/10.1109/TIP.2013.2289994 ..... W. Sun, O. C. Au, L. Xu, Y. Li, and W. Hu 342
Novel Speed-Up Strategies for Non-Local Means Denoising With Patch and Edge Patch Based Dictionaries http://dx.doi.org/10.1109/TIP.2013.2290871 ..... H. Bhujle and S. Chaudhuri 356
Tension in Active Shapes http://dx.doi.org/10.1109/TIP.2013.2288922 ..... G. Papari 366
Measures of Effective Video Tracking http://dx.doi.org/10.1109/TIP.2013.2288578 ..... T. Nawaz, F. Poiesi, and A. Cavallaro 376
Oriented Image Foresting Transform Segmentation by Seed Competition http://dx.doi.org/10.1109/TIP.2013.2288867 ..... P. A. V. Miranda and L. A. C. Mansilla 389
Frame Rate Up Conversion Based on Variational Image Fusion http://dx.doi.org/10.1109/TIP.2013.2288139 ..... W. H. Lee, K. Choi, and J. B. Ra 399
Adaptive Image Resizing Based on Continuous-Domain Stochastic Modeling http://dx.doi.org/10.1109/TIP.2013.2285597 ..... H. Kirshner, A. Bourquard, J. P. Ward, M. Porat, and M. Unser 413
A MAP-Based Image Interpolation Method via Viterbi Decoding of Markov Chains of Interpolation Functions http://dx.doi.org/10.1109/TIP.2013.2290586 ..... F. Vedadi and S. Shirani 424
A Two Dimensional Optical Input to One Dimensional Serial Pulse Transformation Using Confocal Reflectors http://dx.doi.org/10.1109/TIP.2013.2289933 ..... G. Hulse 439
Hierarchical Prediction and Context Adaptive Coding for Lossless Color Image Compression http://dx.doi.org/10.1109/TIP.2013.2293428 ..... S. Kim and N. I. Cho 445
3D Visual Activity Assessment Based on Natural Scene Statistics http://dx.doi.org/10.1109/TIP.2013.2290592 ..... K. Lee, A. K. Moorthy, S. Lee, and A. C. Bovik 450
Parametric Blur Estimation for Blind Restoration of Natural Images: Linear Motion and Out-of-Focus http://dx.doi.org/10.1109/TIP.2013.2286328 ..... J. P. Oliveira, M. A. T. Figueiredo, and J. M. Bioucas-Dias 466

COMMENTS AND CORRECTIONS

Corrections to ''A Speed-Up Scheme Based on Multiple-Instance Pruning for Pedestrian Detection Using a Support Vector Machine'' http://dx.doi.org/10.1109/TIP.2013.2287104 ..... J. Yu, R. Miyamoto, and T. Onoye 478
Erratum to ''Factors Affecting Enhanced Video Quality Preferences'' http://dx.doi.org/10.1109/TIP.2013.2288513 ..... P. Satgunam, R. L. Woods, P. M. Bronstad, and E. Peli 479

IEEE TRANSACTIONS ON

IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

JANUARY 2014 VOLUME 23 NUMBER 1 IIPRE4 (ISSN 1057-7149)

PAPERS

Statistical-Model Based Methods

Spatial Statistics of Image Features for Performance Comparison http://dx.doi.org/10.1109/TIP.2013.2286907 ..... *E. Bostanci, N. Kanwal, and A. F. Clark* 153

Optimal Camera Planning Under Versatile User Constraints in Multi-Camera Image Processing Systems http://dx.doi.org/10.1109/TIP.2013.2287606 ..... *J. Liu, S. Sridharan, C. Fookes, and T. Wark* 171

Camera Model Identification Based on the Heteroscedastic Noise Model http://dx.doi.org/10.1109/TIP.2013.2290596 ..... *T. H. Thai, R. Cogranne, and F. Retraint* 250

Adaptive Image Resizing Based on Continuous-Domain Stochastic Modeling http://dx.doi.org/10.1109/TIP.2013.2285597 ..... *H. Kirshner, A. Bourquard, J. P. Ward, M. Porat, and M. Unser* 413

3D Visual Activity Assessment Based on Natural Scene Statistics http://dx.doi.org/10.1109/TIP.2013.2290592 ..... *K. Lee, A. K. Moorthy, S. Lee, and A. C. Bovik* 450

Image and Video Representation

A Novel Eye Localization Method With Rotation Invariance http://dx.doi.org/10.1109/TIP.2013.2287614 ..... *Y. Ren, S. Wang, B. Hou, and J. Ma* 226

Human Detection by Quadratic Classification on Subspace of Extended Histogram of Gradients http://dx.doi.org/10.1109/TIP.2013.2264677 ..... *A. Satpathy, X. Jiang, and H.-L. Eng* 287

How to Select Good Neighboring Images in Depth-Map Merging Based 3D Modeling http://dx.doi.org/10.1109/TIP.2013.2290597 ..... *S. Shen and Z. Hu* 308

Perception and Quality Models for Images and Video

Attention Driven Foveated Video Quality Assessment http://dx.doi.org/10.1109/TIP.2013.2287611 ..... *J. You, T. Ebrahimi, and A. Perkis* 200

Partial Differential Equation Based Processing of Images and Video

Vector-Valued Image Processing by Parallel Level Sets http://dx.doi.org/10.1109/TIP.2013.2277775 ..... *M. J. Ehrhardt and S. R. Arridge* 9

IEEE

Harmonic Active Contours http://dx.doi.org/10.1109/TIP.2013.2286326 ..... *V. Estellers, D. Zosso, X. Bresson, and J.-P. Thiran* 69

Multiresolution Processing of Images and Video

Translation Invariant Directional Framelet Transform Combined With Gabor Filters for Image Denoising http://dx.doi.org/10.1109/TIP.2013.2285595 ..... *Y. Shi, X. Yang, and Y. Guo* 44

A Weighted Optimization Approach to Time-of-Flight Sensor Fusion http://dx.doi.org/10.1109/TIP.2013.2287613 ..... *S. Schwarz, M. Sjöström, and R. Olsson* 214

Directionlets Using In-Phase Lifting for Image Representation http://dx.doi.org/10.1109/TIP.2013.2288912 ..... *D. Jayachandra and A. Makur* 240

Restoration and Enhancement

Variational Exemplar-Based Image Colorization http://dx.doi.org/10.1109/TIP.2013.2288929 ..... *A. Bugeau, V.-T. Ta, and N. Papadakis* 298

Novel Speed-Up Strategies for Non-Local Means Denoising With Patch and Edge Patch Based Dictionaries http://dx.doi.org/10.1109/TIP.2013.2290871 ..... *H. Bhujle and S. Chaudhuri* 356

Parametric Blur Estimation for Blind Restoration of Natural Images: Linear Motion and Out-of-Focus http://dx.doi.org/10.1109/TIP.2013.2286328 ..... *J. P. Oliveira, M. A. T. Figueiredo, and J. M. Bioucas-Dias* 466

Interpolation, Super-Resolution, and Mosaicing

Frame Rate Up Conversion Based on Variational Image Fusion http://dx.doi.org/10.1109/TIP.2013.2288139 ..... *W. H. Lee, K. Choi, and J. B. Ra* 399

A MAP-Based Image Interpolation Method via Viterbi Decoding of Markov Chains of Interpolation Functions http://dx.doi.org/10.1109/TIP.2013.2290586 ..... *F. Vedadi and S. Shirani* 424

Biomedical and Biological Image Processing

Integrated Segmentation and Interpolation of Sparse Data http://dx.doi.org/10.1109/TIP.2013.2286903 ..... *A. Paiement, M. Mirmehdi, X. Xie, and M. C. K. Hamilton* 110

Shape-Based Normalized Cuts Using Spectral Relaxation for Biomedical Segmentation http://dx.doi.org/10.1109/TIP.2013.2287604 ..... *E. Ruiz Pujadas and M. Reisert* 163

Lossy Coding of Images and Video

Saliency-Aware Video Compression http://dx.doi.org/10.1109/TIP.2013.2282897 ..... *H. Hadizadeh and I. V. Bajić* 19

A Bit Allocation Method for Sparse Source Coding http://dx.doi.org/10.1109/TIP.2013.2286325 ..... *M. Kaaniche, A. Fraysse, B. Pesquet-Popescu, and J.-C. Pesquet* 137

Chroma Intra Prediction Based on Inter-Channel Correlation for HEVC http://dx.doi.org/10.1109/TIP.2013.2288007 ..... *X. Zhang, C. Gisquet, E. François, F. Zou, and O. C. Au* 274

Lossless Coding of Images and Video

Hierarchical Prediction and Context Adaptive Coding for Lossless Color Image Compression http://dx.doi.org/10.1109/TIP.2013.2293428 ..... *S. Kim and N. I. Cho* 445

Error Resilience and Channel Coding for Image and Video Systems

Motion-Aware Mesh-Structured Trellis for Correlation Modelling Aided Distributed Multi-View Video Coding http://dx.doi.org/10.1109/TIP.2013.2288913 ..... *Y. Huo, T. Wang, R. G. Maunder, and L. Hanzo* 319

Imaging and Video Networks

Motion-Related Resource Allocation in Dynamic Wireless Visual Sensor Network Environments http://dx.doi.org/10.1109/TIP.2013.2286323 ..... *A. V. Katsenou, L. P. Kondi, and K. E. Parsopoulos* 56

Image Scanning and Capture

A Two Dimensional Optical Input to One Dimensional Serial Pulse Transformation Using Confocal Reflectors http://dx.doi.org/10.1109/TIP.2013.2289933 ..... *G. Hulse* 439

Color and Multispectral Imaging

Metamer Mismatching http://dx.doi.org/10.1109/TIP.2013.2283148 ..... *A. D. Logvinenko, B. Funt, and C. Godau* 34

Multi-Illuminant Estimation With Conditional Random Fields http://dx.doi.org/10.1109/TIP.2013.2286327 ..... *S. Beigpour, C. Riess, J. van de Weijer, and E. Angelopoulou* 83

Multispectral Image Denoising With Optimized Vector Bilateral Filter http://dx.doi.org/10.1109/TIP.2013.2287612 ..... *H. Peng, R. Rao, and S. A. Dianat* 264

Hardware and Software Systems for Image and Video Processing

Real-Time Scalable Depth Sensing With Hybrid Structured Light Illumination http://dx.doi.org/10.1109/TIP.2013.2286901 ..... *Y. Zhang, Z. Xiong, Z. Yang, and F. Wu* 97

Tomographic Imaging

Filtered Backprojection Reconstruction and Redundancy in Compton Camera Imaging http://dx.doi.org/10.1109/TIP.2013.2288143 ..... *V. Maxim* 332

Region, Boundary, and Shape Analysis

Linear Time Distances Between Fuzzy Sets With Applications to Pattern Matching and Classification http://dx.doi.org/10.1109/TIP.2013.2286904 ..... *J. Lindblad and N. Sladoje* 126

Tension in Active Shapes http://dx.doi.org/10.1109/TIP.2013.2288922 ..... *G. Papari* 366

Oriented Image Foresting Transform Segmentation by Seed Competition http://dx.doi.org/10.1109/TIP.2013.2288867 ..... *P. A. V. Miranda and L. A. C. Mansilla* 389

Image and Video Mid Level Analysis

Measures of Effective Video Tracking http://dx.doi.org/10.1109/TIP.2013.2288578 ..... *T. Nawaz, F. Poiesi, and A. Cavallaro* 376

Image and Video Synthesis, Rendering, and Visualization

An Analytical Model for Synthesis Distortion Estimation in 3D Video http://dx.doi.org/10.1109/TIP.2013.2287608 ..... *L. Fang, N.-M. Cheung, D. Tian, A. Vetro, H. Sun, and O. C. Au* 185

Seamless View Synthesis Through Texture Optimization http://dx.doi.org/10.1109/TIP.2013.2289994 ..... *W. Sun, O. C. Au, L. Xu, Y. Li, and W. Hu* 342

COMMENTS AND CORRECTIONS

Corrections to ''A Speed-Up Scheme Based on Multiple-Instance Pruning for Pedestrian Detection Using a Support Vector Machine'' http://dx.doi.org/10.1109/TIP.2013.2287104 ..... *J. Yu, R. Miyamoto, and T. Onoye* 478

Erratum to ''Factors Affecting Enhanced Video Quality Preferences'' http://dx.doi.org/10.1109/TIP.2013.2288513 ..... *P. Satgunam, R. L. Woods, P. M. Bronstad, and E. Peli* 479

IEEE TRANSACTIONS ON

IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

FEBRUARY 2014 VOLUME 23 NUMBER 2 IIPRE4 (ISSN 1057-7149)

PAPERS

Fingerprint Compression Based on Sparse Representation http://dx.doi.org/10.1109/TIP.2014.2287996 ..... *G. Shao, Y. Wu, Y. A, X. Liu, and T. Guo* 489
On the Spectrum of the Plenoptic Function http://dx.doi.org/10.1109/TIP.2014.2292363 ..... *C. Gilliam, P.-L. Dragotti, and M. Brookes* 502
Frame-Based Recovery of Corrupted Video Files Using Video Codec Specifications http://dx.doi.org/10.1109/TIP.2014.2285625 ..... *G.-H. Na, K.-S. Shim, K.-W. Moon, S. G. Kong, E.-S. Kim, and J. Lee* 517
Visual Typo Correction by Collocative Optimization: A Case Study on Merchandize Images http://dx.doi.org/10.1109/TIP.2014.2293427 .. *X.-Y. Wei, Z.-Q. Yang, C.-W. Ngo, and W. Zhang* 527
Large Discriminative Structured Set Prediction Modeling With Max-Margin Markov Network for Lossless Image Coding http://dx.doi.org/10.1109/TIP.2014.2293429 ..... *W. Dai, H. Xiong, J. Wang, and Y. F. Zheng* 541
Tree Filtering: Efficient Structure-Preserving Smoothing With a Minimum Spanning Tree http://dx.doi.org/10.1109/TIP.2014.2291328 ..... *L. Bao, Y. Song, Q. Yang, H. Yuan, and G. Wang* 555
Action Recognition Using Nonnegative Action Component Representation and Sparse Basis Selection http://dx.doi.org/10.1109/TIP.2014.2292550 ..... *H. Wang, C. Yuan, W. Hu, H. Ling, W. Yang, and C. Sun* 570
Multiple Description Coding With Randomly and Uniformly Offset Quantizers http://dx.doi.org/10.1109/TIP.2014.2288928 ..... *L. Meng, J. Liang, U. Samarawickrama, Y. Zhao, H. Bai, and A. Kaup* 582
A Unified Methodology for Computing Accurate Quaternion Color Moments and Moment Invariants http://dx.doi.org/10.1109/TIP.2014.2289997 ..... *E. G. Karakasis, G. A. Papakostas, D. E. Koulouriotis, and V. D. Tourassis* 596
Stable and Robust Sampling Strategies for Compressive Imaging http://dx.doi.org/10.1109/TIP.2014.2288004 ..... *F. Krahmer and R. Ward* 612
Learning Category-Specific Dictionary and Shared Dictionary for Fine-Grained Image Categorization http://dx.doi.org/10.1109/TIP.2014.2290593 ..... *S. Gao, I. W.-H. Tsang, and Y. Ma* 623
Including the Size of Regions in Image Segmentation by Region-Based Graph http://dx.doi.org/10.1109/TIP.2014.2289984 ..... *A. Rezvanifar and M. Khosravifard* 635
Neighborhood Supported Model Level Fuzzy Aggregation for Moving Object Segmentation http://dx.doi.org/10.1109/TIP.2014.2285598 .. *P. Chiranjeevi and S. Sengupta* 645

IEEE

Modeling Geometric-Temporal Context With Directional Pyramid Co-Occurrence for Action Recognition http://dx.doi.org/10.1109/TIP.2014.2291319 ..... *C. Yuan, X. Li, W. Hu, H. Ling, and S. J. Maybank* 658
Fast Geodesic Active Fields for Image Registration Based on Splitting and Augmented Lagrangian Approaches http://dx.doi.org/10.1109/TIP.2014.2253473 ..... *D. Zosso, X. Bresson, and J.-P. Thiran* 673
Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index http://dx.doi.org/10.1109/TIP.2014.2293423 ..... *W. Xue, L. Zhang, X. Mou, and A. C. Bovik* 684
Recognizing Gaits Across Views Through Correlated Motion Co-Clustering http://dx.doi.org/10.1109/TIP.2014.2294552 ..... *W. Kusakunniran, Q. Wu, J. Zhang, H. Li, and L. Wang* 696
Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition http://dx.doi.org/10.1109/TIP.2014.2292332 ..... *J. Galbally, S. Marcel, and J. Fierrez* 710
Band-Reweighed Gabor Kernel Embedding for Face Image Representation and Recognition http://dx.doi.org/10.1109/TIP.2014.2292560 ..... *C.-X. Ren, D.-Q. Dai, X.-X. Li, and Z.-R. Lai* 725
Multilinear Graph Embedding: Representation and Regularization for Images http://dx.doi.org/10.1109/TIP.2014.2292303 ..... *Y.-L. Chen and C.-T. Hsu* 741
Global Image Denoising http://dx.doi.org/10.1109/TIP.2014.2293425 ..... *H. Talebi and P. Milanfar* 755
Background-Modeling-Based Adaptive Prediction for Surveillance Video Coding http://dx.doi.org/10.1109/TIP.2014.2294549 ..... *X. Zhang, T. Huang, Y. Tian, and W. Gao* 769
Robust Online Learned Spatio-Temporal Context Model for Visual Tracking http://dx.doi.org/10.1109/TIP.2014.2293430 ..... *L. Wen, Z. Cai, Z. Lei, D. Yi, and S. Z. Li* 785
Compressed-Domain Video Retargeting http://dx.doi.org/10.1109/TIP.2014.2294541 ..... *J. Zhang, S. Li, and C.-C. J. Kuo* 797
Latent Hierarchical Model of Temporal Structure for Complex Activity Classification http://dx.doi.org/10.1109/TIP.2014.2295753 ..... *L. Wang, Y. Qiao, and X. Tang* 810
mCENTRIST: A Multi-Channel Feature Generation Mechanism for Scene Categorization http://dx.doi.org/10.1109/TIP.2014.2295756 ..... *Y. Xiao, J. Wu, and J. Yuan* 823
Scale Adaptive Dictionary Learning http://dx.doi.org/10.1109/TIP.2014.2287602 ..... *C. Lu, J. Shi, and J. Jia* 837
Scaled Heavy-Ball Acceleration of the Richardson-Lucy Algorithm for 3D Microscopy Image Restoration http://dx.doi.org/10.1109/TIP.2014.2291324 ..... *H. Wang and P. C. Miller* 848
Enhanced Decoupled Active Contour Using Structural and Textural Variation Energy Functionals http://dx.doi.org/10.1109/TIP.2014.2295752 ..... *D. Lui, C. Scharfenberger, K. Fergani, A. Wong, and D. A. Clausi* 855
Probability-Based Rendering for View Synthesis http://dx.doi.org/10.1109/TIP.2014.2295716 ..... *B. Ham, D. Min, C. Oh, M. N. Do, and K. Sohn* 870
Topology-Preserving Rigid Transformation of 2D Digital Images http://dx.doi.org/10.1109/TIP.2014.2295751 ..... *P. Ngo, N. Passat, Y. Kenmochi, and H. Talbot* 885
RotationCovariant Texture Learning Using Steerable Riesz Wavelets http://dx.doi.org/10.1109/TIP.2014.2295755 ..... *A. Depeursinge, A. Foncubierta-Rodriguez, D. Van de Ville, and H. Mller* 898
Region-Based Iterative Reconstruction of Structurally Changing Objects in CT http://dx.doi.org/10.1109/TIP.2014.2297024 ..... *G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers* 909
A General Exponential Framework for Dimensionality Reduction http://dx.doi.org/10.1109/TIP.2014.2297020 ..... *S.-J. Wang, S. Yan, J. Yang, C.-G. Zhou, and X. Fu* 920
Transmission Policy Selection for Multi-View Content Delivery Over Bandwidth Constrained Channels http://dx.doi.org/10.1109/TIP.2014.2293419 ..... *J. Chakareski* 931

TECHNICAL CORRESPONDENCE

Configuration-Transition-Based Connected-Component Labeling http://dx.doi.org/10.1109/TIP.2014.2289968 ..... *L. He, X. Zhao, Y. Chao, and K. Suzuki* 943

COMMENTS AND CORRECTIONS

Region, Boundary, and Shape Analysis

Comment on ''Collinear Segment Detection Using HT Neighborhoods'' http://dx.doi.org/10.1109/TIP.2014.2291329 ..... *P. S. Rahmdel, D. Shi, and R. Comley* 952
EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIP.2014.2301101 956
Information for Authors http://dx.doi.org/10.1109/TIP.2014.2301102 957

IEEE TRANSACTIONS ON

IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

FEBRUARY 2014 VOLUME 23 NUMBER 2 IIPRE4 (ISSN 1057-7149)

PAPERS

Statistical-Model Based Methods

A General Exponential Framework for Dimensionality Reduction http://dx.doi.org/10.1109/TIP.2014.2297020 ..... *S.-J. Wang, S. Yan, J. Yang, C.-G. Zhou, and X. Fu* 920

Image and Video Representation

A Unified Methodology for Computing Accurate Quaternion Color Moments and Moment Invariants http://dx.doi.org/10.1109/TIP.2014.2289997 ..... *E. G. Karakasis, G. A. Papakostas, D. E. Koulouriotis, and V. D. Tourassis* 596

Learning Category-Specific Dictionary and Shared Dictionary for Fine-Grained Image Categorization http://dx.doi.org/10.1109/TIP.2014.2290593 ..... *S. Gao, I. W.-H. Tsang, and Y. Ma* 623

Neighborhood Supported Model Level Fuzzy Aggregation for Moving Object Segmentation http://dx.doi.org/10.1109/TIP.2014.2285598 ..... *P. Chiranjeevi and S. Sengupta* 645

Modeling Geometric-Temporal Context With Directional Pyramid Co-Occurrence for Action Recognition http://dx.doi.org/10.1109/TIP.2014.2291319 ..... *C. Yuan, X. Li, W. Hu, H. Ling, and S. J. Maybank* 658

Multilinear Graph Embedding: Representation and Regularization for Images http://dx.doi.org/10.1109/TIP.2014.2292303 ..... *Y.-L. Chen and C.-T. Hsu* 741

Perception and Quality Models for Images and Video

Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index http://dx.doi.org/10.1109/TIP.2014.2293423 ..... *W. Xue, L. Zhang, X. Mou, and A. C. Bovik* 684

Linear and Nonlinear Filtering of Images and Video

Tree Filtering: Efficient Structure-Preserving Smoothing With a Minimum Spanning Tree http://dx.doi.org/10.1109/TIP.2014.2291328 ..... *L. Bao, Y. Song, Q. Yang, H. Yuan, and G. Wang* 555

Topology-Preserving Rigid Transformation of 2D Digital Images http://dx.doi.org/10.1109/TIP.2014.2295751 ..... *P. Ngo, N. Passat, Y. Kenmochi, and H. Talbot* 885

IEEE

RotationCovariant Texture Learning Using Steerable Riesz Wavelets http://dx.doi.org/10.1109/TIP.2014.2295755 ..... A. Depeursinge, A. Foncubierta-Rodriguez, D. Van de Ville, and H. Mller 898

Multiresolution Processing of Images and Video

Band-Reweighed Gabor Kernel Embedding for Face Image Representation and Recognition http://dx.doi.org/10.1109/TIP.2014.2292560 .. C.-X. Ren, D.-Q. Dai, X.-X. Li, and Z.-R. Lai 725
Compressed-Domain Video Retargeting http://dx.doi.org/10.1109/TIP.2014.2294541 ..... J. Zhang, S. Li, and C.-C. J. Kuo 797

Restoration and Enhancement

Frame-Based Recovery of Corrupted Video Files Using Video Codec Specifications http://dx.doi.org/10.1109/TIP.2014.2285625 ..... G.-H. Na, K.-S. Shim, K.-W. Moon, S. G. Kong, E.-S. Kim, and J. Lee 517
Global Image Denoising http://dx.doi.org/10.1109/TIP.2014.2293425 ..... H. Talebi and P. Milanfar 755
Scaled Heavy-Ball Acceleration of the Richardson-Lucy Algorithm for 3D Microscopy Image Restoration http://dx.doi.org/10.1109/TIP.2014.2291324 ..... H. Wang and P. C. Miller 848

Lossy Coding of Images and Video

Fingerprint Compression Based on Sparse Representation http://dx.doi.org/10.1109/TIP.2014.2287996 ..... G. Shao, Y. Wu, Y. A, X. Liu, and T. Guo 489
Multiple Description Coding With Randomly and Uniformly Offset Quantizers http://dx.doi.org/10.1109/TIP.2014.2288928 ..... L. Meng, J. Liang, U. Samarawickrama, Y. Zhao, H. Bai, and A. Kaup 582
Background-Modeling-Based Adaptive Prediction for Surveillance Video Coding http://dx.doi.org/10.1109/TIP.2014.2294549 ..... X. Zhang, T. Huang, Y. Tian, and W. Gao 769

Lossless Coding of Images and Video

Large Discriminative Structured Set Prediction Modeling With Max-Margin Markov Network for Lossless Image Coding http://dx.doi.org/10.1109/TIP.2014.2293429 ..... W. Dai, H. Xiong, J. Wang, and Y. F. Zheng 541

Imaging and Video Networks

Transmission Policy Selection for Multi-View Content Delivery Over Bandwidth Constrained Channels http://dx.doi.org/10.1109/TIP.2014.2293419 ..... J. Chakareski 931

Stereoscopic and Multiview Processing and Display

Probability-Based Rendering for View Synthesis http://dx.doi.org/10.1109/TIP.2014.2295716 ..... B. Ham, D. Min, C. Oh, M. N. Do, and K. Sohn 870

Tomographic Imaging

Region-Based Iterative Reconstruction of Structurally Changing Objects in CT http://dx.doi.org/10.1109/TIP.2014.2297024 ..... G. Van Eyndhoven, K. J. Batenburg, and J. Sijbers 909

Radar Imaging, Remote Sensing, and Geophysical Imaging

Stable and Robust Sampling Strategies for Compressive Imaging http://dx.doi.org/10.1109/TIP.2014.2288004 ..... F. Krahmer and R. Ward 612

Region, Boundary, and Shape Analysis

Including the Size of Regions in Image Segmentation by Region-Based Graph http://dx.doi.org/10.1109/TIP.2014.2289984 ..... A. Rezvanifar and M. Khosravifard 635

Image and Video Mid Level Analysis

Action Recognition Using Nonnegative Action Component Representation and Sparse Basis Selection http://dx.doi.org/10.1109/TIP.2014.2292550 ..... H. Wang, C. Yuan, W. Hu, H. Ling, W. Yang, and C. Sun 570
Fast Geodesic Active Fields for Image Registration Based on Splitting and Augmented Lagrangian Approaches http://dx.doi.org/10.1109/TIP.2014.2253473 ..... D. Zosso, X. Bresson, and J.-P. Thiran 673
Robust Online Learned Spatio-Temporal Context Model for Visual Tracking http://dx.doi.org/10.1109/TIP.2014.2293430 ..... L. Wen, Z. Cai, Z. Lei, D. Yi, and S. Z. Li 785

Image and Video Interpretation and Understanding

Latent Hierarchical Model of Temporal Structure for Complex Activity Classification http://dx.doi.org/10.1109/TIP.2014.2295753 ..... L. Wang, Y. Qiao, and X. Tang 810
mCENTRIST: A Multi-Channel Feature Generation Mechanism for Scene Categorization http://dx.doi.org/10.1109/TIP.2014.2295756 ..... Y. Xiao, J. Wu, and J. Yuan 823
Scale Adaptive Dictionary Learning http://dx.doi.org/10.1109/TIP.2014.2287602 ..... C. Lu, J. Shi, and J. Jia 837

Enhanced Decoupled Active Contour Using Structural and Textural Variation Energy Functionals http://dx.doi.org/10.1109/TIP.2014.2295752 ..... *D. Lui, C. Scharfenberger, K. Fergani, A. Wong, and D. A. Clausi* 855

Image and Video Biometric Analysis

Recognizing Gaits Across Views Through Correlated Motion Co-Clustering http://dx.doi.org/10.1109/TIP.2014.2294552 ..... *W. Kusakunniran, Q. Wu, J. Zhang, H. Li, and L. Wang* 696

Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint, and Face Recognition http://dx.doi.org/10.1109/TIP.2014.2292332 ..... *J. Galbally, S. Marcel, and J. Fierrez* 710

Image and Video Storage and Retrieval

Visual Typo Correction by Collocative Optimization: A Case Study on Merchandize Images http://dx.doi.org/10.1109/TIP.2014.2293427 ..... *X.-Y. Wei, Z.-Q. Yang, C.-W. Ngo, and W. Zhang* 527

Image and Video Synthesis, Rendering, and Visualization

On the Spectrum of the Plenoptic Function http://dx.doi.org/10.1109/TIP.2014.2292363 ..... *C. Gilliam, P.-L. Dragotti, and M. Brookes* 502

TECHNICAL CORRESPONDENCE

Region, Boundary, and Shape Analysis

Configuration-Transition-Based Connected-Component Labeling http://dx.doi.org/10.1109/TIP.2014.2289968 ..... *L. He, X. Zhao, Y. Chao, and K. Suzuki* 943

COMMENTS AND CORRECTIONS

Region, Boundary, and Shape Analysis

Comment on ''Collinear Segment Detection Using HT Neighborhoods'' http://dx.doi.org/10.1109/TIP.2014.2291329 ..... *P. S. Rahmdel, D. Shi, and R. Comley* 952

EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIP.2014.2301101 956

Information for Authors http://dx.doi.org/10.1109/TIP.2014.2301102 957

IEEE TRANSACTIONS ON

INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

JANUARY 2014 VOLUME 9 NUMBER 1 ITIFA6 (ISSN 1556-6013)

PAPERS

Non-Interactive Key Establishment for Bundle Security Protocol of Space DTNs http://dx.doi.org/10.1109/TIFS.2013.2289993 ..... *X. Lv, Y. Mu, and H. Li* 5

Gabor Ordinal Measures for Face Recognition http://dx.doi.org/10.1109/TIFS.2013.2290064 ..... *Z. Chai, Z. Sun, H. Méndez-Vázquez, R. He, and T. Tan* 14

Building a Scalable System for Stealthy P2P-Botnet Detection http://dx.doi.org/10.1109/TIFS.2013.2290197 ..... *J. Zhang, R. Perdisci, W. Lee, X. Luo, and U. Sarfraz* 27

Designing an Efficient Image Encryption-Then-Compression System via Prediction Error Clustering and Random Permutation http://dx.doi.org/10.1109/TIFS.2013.2291625 ..... *J. Zhou, X. Liu, O. C. Au, and Y. Y. Tang* 39

Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions http://dx.doi.org/10.1109/TIFS.2013.2291969 ..... *J. Lu, G. Wang, and P. Moulin* 51

A High Performance Fingerprint Matching System for Large Databases Based on GPU http://dx.doi.org/10.1109/TIFS.2013.2291220 ..... *P. D. Gutiérrez, M. Lastra, F. Herrera, and J. M. Benæetez* 62

Data-Centric OS Kernel Malware Characterization http://dx.doi.org/10.1109/TIFS.2013.2291964 ..... *J. Rhee, R. Riley, Z. Lin, X. Jiang, and D. Xu* 72

Digital Image Sharing by Diverse Image Media http://dx.doi.org/10.1109/TIFS.2013.2292509 ..... *K.-H. Lee and P.-L. Chiu* 88

Catch Me If You Can: Evaluating Android Anti-Malware Against Transformation Attacks http://dx.doi.org/10.1109/TIFS.2013.2290431 .... *V. Rastogi, Y. Chen, and X. Jiang* 99

Top-k Query Result Completeness Verification in Tiered Sensor Networks http://dx.doi.org/10.1109/TIFS.2013.2291326 ..... *C.-M. Yu, G.-K. Ni, I.-Y. Chen, E. Gelenbe, and S.-Y. Kuo* 109

Privacy-Preserving Authenticated Key-Exchange Over Internet http://dx.doi.org/10.1109/TIFS.2013.2293457 ..... *A. C.-C. Yao and Y. Zhao* 125

IEEE

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags

IEEE TRANSACTIONS ON
INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

JANUARY 2014 VOLUME 9 NUMBER 1 ITIFA6 (ISSN 1556-6013)

PAPERS

Anonymization and Data Privacy

Top-k Query Result Completeness Verification in Tiered Sensor Networks http://dx.doi.org/10.1109/TIFS.2013.2291326 ..... *C.-M. Yu, G.-K. Ni, I.-Y. Chen, E. Gelenbe, and S.-Y. Kuo* 109

Applied Cryptography

Designing an Efficient Image Encryption-Then-Compression System via Prediction Error Clustering and Random Permutation http://dx.doi.org/10.1109/TIFS.2013.2291625 ..... *J. Zhou, X. Liu, O. C. Au, and Y. Y. Tang* 39

Digital Image Sharing by Diverse Image Media http://dx.doi.org/10.1109/TIFS.2013.2292509 ..... *K.-H. Lee and P.-L. Chiu* 88

Biometrics

Gabor Ordinal Measures for Face Recognition http://dx.doi.org/10.1109/TIFS.2013.2290064 ..... *Z. Chai, Z. Sun, H. Méndez-Vázquez, R. He, and T. Tan* 14

Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions http://dx.doi.org/10.1109/TIFS.2013.2291969 ..... *J. Lu, G. Wang, and P. Moulin* 51

A High Performance Fingerprint Matching System for Large Databases Based on GPU http://dx.doi.org/10.1109/TIFS.2013.2291220 ..... *P. D. Gutiérrez, M. Lastra, F. Herrera, and J. M. Benætez* 62

Network Security

Non-Interactive Key Establishment for Bundle Security Protocol of Space DTNs http://dx.doi.org/10.1109/TIFS.2013.2289993 ..... *X. Lv, Y. Mu, and H. Li* 5

Building a Scalable System for Stealthy P2P-Botnet Detection http://dx.doi.org/10.1109/TIFS.2013.2290197 ..... *J. Zhang, R. Perdisci, W. Lee, X. Luo, and U. Sarfraz* 27

Data-Centric OS Kernel Malware Characterization http://dx.doi.org/10.1109/TIFS.2013.2291964 ..... *J. Rhee, R. Riley, Z. Lin, X. Jiang, and D. Xu* 72

Privacy-Preserving Authenticated Key-Exchange Over Internet http://dx.doi.org/10.1109/TIFS.2013.2293457 ..... *A. C.-C. Yao and Y. Zhao* 125

Security of Large Networked Systems

Catch Me If You Can: Evaluating Android Anti-Malware Against Transformation Attacks http://dx.doi.org/10.1109/TIFS.2013.2290431 ..... *V. Rastogi, Y. Chen, and X. Jiang* 99

IEEE

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

FEBRUARY 2014 VOLUME 9 NUMBER 2 ITIFA6 (ISSN 1556-6013)

PAPERS

An Efficient Parallel Approach for Sclera Vein Recognition http://dx.doi.org/10.1109/TIFS.2013.2291314 ..... *Y. Lin, E. Y. Du, Z. Zhou, and N. L. Thomas* 147

An Information-Theoretic Security Evaluation of a Class of Randomized Encryption Schemes http://dx.doi.org/10.1109/TIFS.2013.2294763 ..... *F. Oggier and M. J. Mihaljević* 158

Adaptive Self-Embedding Scheme With Controlled Reconstruction Performance http://dx.doi.org/10.1109/TIFS.2013.2295154 ..... *P. Korus and A. Dziech* 169

A General Framework of Wiretap Channel With Helping Interference and State Information http://dx.doi.org/10.1109/TIFS.2013.2295031 ..... *P. Xu, Z. Ding, X. Dai, and K. K. Leung* 182

Growing Grapes in Your Computer to Defend Against Malware http://dx.doi.org/10.1109/TIFS.2013.2291066 ..... *Z. Shan and X. Wang* 196

Role-Dependent Privacy Preservation for Secure V2G Networks in the Smart Grid http://dx.doi.org/10.1109/TIFS.2013.2295032 ..... *H. Liu, H. Ning, Y. Zhang, Q. Xiong, and L. T. Yang* 208

The Silence of the LANs: Efficient Leakage Resilience for IPsec VPNs http://dx.doi.org/10.1109/TIFS.2013.2289978 ..... *S. Schulz, V. Varadharajan, and A.-R. Sadeghi* 221

Learning Multi-Boosted HMMs for Lip-Password Based Speaker Verification http://dx.doi.org/10.1109/TIFS.2013.2293025 ..... *X. Liu and Y.-M. Cheung* 233

Physical Layer Security of Maximal Ratio Combining in Two-Wave With Diffuse Power Fading Channels http://dx.doi.org/10.1109/TIFS.2013.2296991 ..... *L. Wang, N. Yang, M. Elkashlan, P. L. Yeoh, and J. Yuan* 247

Toward Optimal Fusion Algorithms With Security Against Wolves and Lambs in Biometrics http://dx.doi.org/10.1109/TIFS.2013.2296993 ..... *T. Murakami, K. Takahashi, and K. Matsuura* 259

Cooperative Key Agreement for Wireless Networking: Key Rates and Practical Protocol Design http://dx.doi.org/10.1109/TIFS.2013.2293113 ..... *N. Wang, N. Zhang, and T. A. Gulliver* 272

Double Trouble: Differentiating Identical Twins by Face Recognition http://dx.doi.org/10.1109/TIFS.2013.2296373 ..... *J. R. Paone, P. J. Flynn, P. J. Philips, K. W. Bowyer, R. W. V. Bruegge, P. J. Grother, G. W. Quinn, M. T. Pruitt, and J. M. Grant* 285

Firewall Policy Reconnaissance: Techniques and Analysis http://dx.doi.org/10.1109/TIFS.2013.2296874 ..... *M. Q. Ali, E. Al-Shaer, and T. Samak* 296

Perceptual Encryption of H.264 Videos: Embedding Sign-Flips Into the Integer-Based Transforms http://dx.doi.org/10.1109/TIFS.2013.2293955 ..... *B. Zeng, S.-K. Au Yeung, S. Zhu, and M. Gabbouj* 309

Design and Performance Analysis of a Virtual Ring Architecture for Smart Grid Privacy http://dx.doi.org/10.1109/TIFS.2013.2296441 ..... *M. Badra and S. Zeadally* 321

EDICS-Editors Information Classification Scheme http://dx.doi.org/10.1109/TIFS.2014.2300722 330

Information for Authors http://dx.doi.org/10.1109/TIFS.2014.2300721 331

IEEE

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

FEBRUARY 2014 VOLUME 9 NUMBER 2 ITIFA6 (ISSN 1556-6013)

PAPERS

Anonymization and Data Privacy

Design and Performance Analysis of a Virtual Ring Architecture for Smart Grid Privacy http://dx.doi.org/10.1109/TIFS.2013.2296441 ..... *M. Badra and S. Zeadally* 321

Applied Cryptography

Perceptual Encryption of H.264 Videos: Embedding Sign-Flips Into the Integer-Based Transforms http://dx.doi.org/10.1109/TIFS.2013.2293955 ..... *B. Zeng, S.-K. Au Yeung, S. Zhu, and M. Gabbouj* 309

Biometrics

An Efficient Parallel Approach for Sclera Vein Recognition http://dx.doi.org/10.1109/TIFS.2013.2291314 ..... *Y. Lin, E. Y. Du, Z. Zhou, and N. L. Thomas* 147

Learning Multi-Boosted HMMs for Lip-Password Based Speaker Verification http://dx.doi.org/10.1109/TIFS.2013.2293025 ..... *X. Liu and Y.-M. Cheung* 233

Toward Optimal Fusion Algorithms With Security Against Wolves and Lambs in Biometrics http://dx.doi.org/10.1109/TIFS.2013.2296993 ..... *T. Murakami, K. Takahashi, and K. Matsuura* 259

Double Trouble: Differentiating Identical Twins by Face Recognition http://dx.doi.org/10.1109/TIFS.2013.2296373 ..... *J. R. Paone, P. J. Flynn, P. J. Philips, K. W. Bowyer, R. W. V. Bruegge, P. J. Grother, G. W. Quinn, M. T. Pruitt, and J. M. Grant* 285

Communication and Physical-Layer Security

A General Framework of Wiretap Channel With Helping Interference and State Information http://dx.doi.org/10.1109/TIFS.2013.2295031 ..... *P. Xu, Z. Ding, X. Dai, and K. K. Leung* 182

Physical Layer Security of Maximal Ratio Combining in Two-Wave With Diffuse Power Fading Channels http://dx.doi.org/10.1109/TIFS.2013.2296991 ..... *L. Wang, N. Yang, M. Elkashlan, P. L. Yeoh, and J. Yuan* 247

Cooperative Key Agreement for Wireless Networking: Key Rates and Practical Protocol Design http://dx.doi.org/10.1109/TIFS.2013.2293113 ..... *N. Wang, N. Zhang, and T. A. Gulliver* 272

Information Theoretic Security

An Information-Theoretic Security Evaluation of a Class of Randomized Encryption Schemes http://dx.doi.org/10.1109/TIFS.2013.2294763 ..... *F. Oggier and M. J. Mihaljević* 158

Multimedia Content Hash

Adaptive Self-Embedding Scheme With Controlled Reconstruction Performance http://dx.doi.org/10.1109/TIFS.2013.2295154 ..... *P. Korus and A. Dziech* 169

IEEE

Network Security

Growing Grapes in Your Computer to Defend Against Malware http://dx.doi.org/10.1109/TIFS.2013.2291066 ..... *Z. Shan and X. Wang* 196
Role-Dependent Privacy Preservation for Secure V2G Networks in the Smart Grid http://dx.doi.org/10.1109/TIFS.2013.2295032 ..... *H. Liu, H. Ning, Y. Zhang, Q. Xiong, and L. T. Yang* 208
Firewall Policy Reconnaissance: Techniques and Analysis http://dx.doi.org/10.1109/TIFS.2013.2296874 ..... *M. Q. Ali, E. Al-Shaer, and T. Samak* 296

Steganography and Covert Communications

The Silence of the LANs: Efficient Leakage Resilience for IPsec VPNs http://dx.doi.org/10.1109/TIFS.2013.2289978 ..... *S. Schulz, V. Varadharajan, and A.-R. Sadeghi* 221
EDICS-Editors Information Classification Scheme http://dx.doi.org/10.1109/TIFS.2014.2300722 330
Information for Authors http://dx.doi.org/10.1109/TIFS.2014.2300721 331

CALL FOR PAPERS

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

Special Issue on Biometric Spoofing and Countermeasures

Guest Editors

Nicholas Evans EURECOM, France (evans@eurecom.fr)
Sébastien Marcel Idiap Research Institute, Switzerland (marcel@idiap.ch)
Arun Ross Michigan State University, USA (rossarun@cse.msu.edu)
Stan Z. Li Chinese Academy of Sciences, China (szli@nlpr.ia.ac.cn)

While biometrics technology has revolutionized approaches to person authentication and has evolved to play a critical role in personal, national and global security, the potential for the technology to be fooled or 'spoofed' is widely acknowledged. Efforts to study such threats and to develop countermeasures are now well underway resulting in some promising solutions. While progress with respect to each biometric modality has attained varying degrees of maturity, there are some notable shortcomings in research methodologies. Current spoofing studies focus on specific, known attacks. Existing countermeasures designed to detect and deflect such attacks are often based on unrealistic *a priori* knowledge and typically learned using training data produced using exactly the same spoofing method that is to be detected. Current countermeasures thus have questionable application in practical scenarios where the nature of the attack can never be known. This special issue will focus on the latest research on the topic of biometric spoofing and countermeasures, with a particular emphasis on novel methodologies and generalized spoofing countermeasures that have the potential to protect biometric systems against varying or previously unseen attacks. The aim is to further the state-of-the-art in this field, to stimulate interactions between the biometrics and information forensic communities, to encourage the development of reliable methodologies in spoofing and countermeasure assessment and solutions, and to promote the development of generalized countermeasures. Papers on biometric obfuscation (e.g., fingerprint or face alteration) and relevant countermeasures will also be considered in the special issue. Novel contributions related to both traditional biometric modalities such as face, iris, fingerprint, and voice, and other modalities such as vasculature and electrophysiological signals will be considered. The focus includes, but is not limited to, the following topics related to spoofing and anti-spoofing countermeasures in biometrics:

- vulnerability analysis with an emphasis on previously unconsidered spoofing attacks;
- theoretical models for attack vectors;
- advanced machine learning and pattern recognition algorithms for anti-spoofing;
- information theoretic approaches to quantify spoofing vulnerability;
- spoofing and anti-spoofing in mobile devices;
- generalized countermeasures;
- challenge-response countermeasures;
- sensor-based solutions to spoof attacks;
- biometric obfuscation schemes;
- information forensic approaches to spoofing detection;
- new evaluation protocols, datasets, and performance metrics;
- reproducible research (public databases, open source software and experimental setups).

Submission Procedure: Manuscripts are to be submitted according to the Information for Authors at http://www.signalprocessingsociety.org/publications/periodicals/forensics/forensics-authors-info/ using the IEEE online manuscript system, Manuscript Central. Papers must not have appeared or be under review elsewhere. Manuscripts by the guest editors submitted to this SI will be handled by the EIC of IEEE-TIFS.

Schedule:
Submission deadline: 1st June 2014
First Review: 15th September 2014
Revisions Due: 1st November 2014
Final Decision: 15th December 2014
Final manuscript due: 15th January 2015
Tentative publication date: 1st April 2015

2014 International Workshop on Multimedia Signal Processing (MMSP 2014)

in Jakarta, Indonesia, Sept 22 - 24 , 2014

http://mmsp2014.ilearning.me

The 16th International Workshop on Multimedia Signal Processing will be held in STMIK Raharja, Jakarta, Indonesia, on September 22-24, 2014. This year's event has a new theme "Internet-of-Thing (IoT) Multimedia". The theme covers multiple research areas of multimedia signal processing including processing, analysis, networking, security, database and visualization with emphasis on issues, challenges and solutions arising from their synergy with the emerging IoT paradigm. To recognize high quality research papers, MMSP 2014 will be giving out awards for the best paper, best student paper and best poster. Nominations for the awards will be based on the significance and potential impact to the scientific community. The topics for regular sessions include, but are not limited to:

Multimedia coding
- Distributed/centralized source coding for sensor networks
- Scalable and low delay source coding
- Error/loss resilient source coding
- Standards

Multimedia networking
- Multimedia communication over IoT
- Green multimedia communication
- Wireless multimedia sensor networks
- QoS control and scheduling for real time applications
- Networked audio and video quality monitoring and management
- Cross-layer optimization

Multimedia Signal Processing
- Multimedia signal de-noising and enhancement
- Multimedia scene analysis and event detection
- Multimedia security and forensics
- Joint audio-visual processing and analysis
- Multimedia database and content management

Virtual reality and 3D
- Augmented, mixed and virtual reality and 3D imaging
- 2D and 3D graphics/geometry coding and animation
- Distributed virtual reality and immersive communications
- 3D audio and video processing and rendering

Multimedia systems and applications
- Multimedia system for traffic and public security
- Mobile and location aware multimedia services
- Cloud-assisted multimedia systems
- eHealthcare and telemedicine
- Multimodal collaboration
- Social networking

IoT systems and applications
- IoT signal processing
- IoT systems and network
- Interaction of IoT and humans
- IoT applications
- IoT related standards

Indonesia

Indonesia is the largest archipelago country in the world. With over 17,000 islands covering a land size of about 2 million square km, Indonesia is multicultural, has more than 700 indigenous languages and presents numerous world's top tourist attraction places such as Bali, Lake Toba, Temple of Borobudur, and etc. Famous for its diverse artistic dancing and handicraft workpiece, Indonesia is notably rich in cuisine and is a nature-preserving place that one must visit. Its capital in Jakarta is the largest and most populous city in Indonesia and is the country's economic, cultural and political centre. As the gateway of the country, Jakarta often is a stop-over to many popular tourist destination and is also one of the best places to shop in South East Asia.

Top 10% paper Award :

This award is granted to as many as 10% of the total paper submissions. To be selected from both oral and poster sessions, these papers are high quality accepted papers. Selection will be done by a panel of judges and is based on originality, technical contributions, and presentation during the workshop. The award includes a certificate and recognition on the official MMSP website.

Demo & Sketch

Authors are invited to submit a 1-2 page description of demonstrable prototype in the form of software or hardware in the above-mentioned area. In addition, authors can also submit a 2-4 page description of late breaking results or work that is not yet sufficiently mature for a full paper. Both Demo & Sketch works will be included in the conference brochure. The Demo & Sketch forum provides an excellent opportunity to get feedback from world-leading experts so the works can be refined and extended for future publication.

Important Dates

Proposals for Special Sessions :	February 20, 2014
Full Paper (4-6 pages) Submissions (Regular and Special Sessions):	April 4, 2014
Notification to Authors :	May 12, 2014
Camera-ready Paper Submission :	June 2, 2014
Author registration deadline :	June 12, 2014
Sketch & Demo Submission :	June 22, 2014
Technical Program :	June 22, 2014
Author Notification :	July 2, 2014
Early registration deadline :	July 12, 2014

Find us :

General Chairs
Susanto Rahardja
National University of Singapore

Zhengyou Zhang
Microsoft Research

Technical Program Chairs
Fernando Pereira
Instituto Superior Tecnico - Instito de Telecomunicatoes, Lisbon

Alexander Loui
Kodak Research

Overview Chair
Mohan Kankanhalli
National University of Singapore

Demo & Sketch Chairs
Norliza Mohd Noor
Universiti Teknologi Malaysia

Ayu Purwarianti
Institut Teknologi Bandung

Special Session Chairs
Rongshan Yu
Institute for Infocomm Research

Anthony Ventro
Mitsubishi Electric Research Lab

Local Arrangement Chairs
Untung Rahardja
STMIK Raharja

Lekha Chaisorn
National University of Singapore

Finance Chair
Chia-Wen Lin
National Tsing Hua University

Publicity Chair
Hitoshi Kiya
Tokyo Metropolitan University

Publication Chair
Cha Zhang
Microsoft Research

Registration Chair
Wei Yao
Institute for Infocomm Research

America Liaison
Nam Ling
Santa Clara University

Asia/Australia Liaison
Oscar Au
Hongkong University of Science & Technology

Europe Liaison
Enrico Magli
Politecnico di Torino

Sponsorship Chair
Abas Sunarya
STMIK Raharja

IEEE Signal Processing Society

IEEE TRANSACTIONS ON MULTIMEDIA

A PUBLICATION OF
THE IEEE CIRCUITS AND SYSTEMS SOCIETY
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY

http://www.signalprocessingsociety.org/tmm/

JANUARY 2014 VOLUME 16 NUMBER 1 ITMUF8 (ISSN 1520-9210)

PAPERS

Audio/Video Analysis and Synthesis

Automatic Estimation of Multiple Motion Fields From Video Sequences Using a Region Matching Based Approach http://dx.doi.org/10.1109/TMM.2013.2281023 ..... *M. V. Afonso, J. C. Nascimento, and J. S. Marques* 1

CAVVA: Computational Affective Video-in-Video Advertising http://dx.doi.org/10.1109/TMM.2013.2282128 ..... *K. Yadati, H. Katti, and M. Kankanhalli* 15

Watermarking, Encryption, and Data Hiding

Visual Protection of HEVC Video by Selective Encryption of CABAC Binstrings http://dx.doi.org/10.1109/TMM.2013.2281029 ..... *Z. Shahid and W. Puech* 24

3-D Audio/Video Processing

Creating the Sydney York Morphological and Acoustic Recordings of Ears Database http://dx.doi.org/10.1109/TMM.2013.2282134 ..... *C. T. Jin, P. Guillon, N. Epain, R. Zolfaghari, A. van Schaik, A. I. Tew, C. Hetherington, and J. Thorpe* 37

System Design Methodology and Tools

Example-Based Human Motion Extrapolation and Motion Repairing Using Contour Manifold http://dx.doi.org/10.1109/TMM.2013.2283844 ..... *N. C. Tang, C.-T. Hsu, M.-F. Weng, T.-Y. Lin, and H.-Y. M. Liao* 47

Low-Power Digital and Analog Circuits for Multimedia

An Adaptive Thread Scheduling Mechanism With Low-Power Register File for Mobile GPUs http://dx.doi.org/10.1109/TMM.2013.2281584 ..... *C.-C. Hsiao, S.-L. Chu, and C.-C. Hsieh* 60

Display Technology for Multimedia

Generalized Equalization Model for Image Enhancement http://dx.doi.org/10.1109/TMM.2013.2283453 ..... *H. Xu, G. Zhai, X. Wu, and X. Yang* 68

Self-Learning Based Image Decomposition With Applications to Single Image Denoising http://dx.doi.org/10.1109/TMM.2013.2284759 ..... *D.-A. Huang, L.-W. Kang, Y.-C. F. Wang, and C.-W. Lin* 83

Weakly Supervised Photo Cropping http://dx.doi.org/10.1109/TMM.2013.2286817 ..... *L. Zhang, M. Song, Y. Yang, Q. Zhao, C. Zhao, and N. Sebe* 94

Architectures and Design Techniques

Dynamic Load Balancing for Real-Time Video Encoding on Heterogeneous CPU+GPU Systems http://dx.doi.org/10.1109/TMM.2013.2284892 ..... *S. Momcilovic, A. Ilic, N. Roma, and L. Sousa* 108

3-D Presentation

Robust Semi-Automatic Depth Map Generation in Unconstrained Images and Video Sequences for 2D to Stereoscopic 3D Conversion http://dx.doi.org/10.1109/TMM.2013.2283451 ..... *R. Phan and D. Androutsos* 122

System Performance

A Simple and Efficient Re-Scrambling Scheme for DTV Programs http://dx.doi.org/10.1109/TMM.2013.2283848 ..... *Y. Liu, J. Duan, Q. Tang, and Y. Zhang* 137

Knowledge and Semantics Modeling for Multimedia Databases

Learning Effective Event Models to Recognize a Large Number of Human Actions http://dx.doi.org/10.1109/TMM.2013.2283846 ..... *J. Wu and D. Hu* 147

Multimedia Search and Retrieval

Exploiting Click Constraints and Multi-view Features for Image Re-ranking http://dx.doi.org/10.1109/TMM.2013.2284755 ..... *J. Yu, Y. Rui, and B. Chen* 159

A Generative Model for Concurrent Image Retrieval and ROI Segmentation http://dx.doi.org/10.1109/TMM.2013.2286083 ..... *I. González-Díaz, C. E. Baz-Hormigos, and F. Díaz-de-María* 169

Social and Web Multimedia

Socialized Mobile Photography: Learning to Photograph With Social Context via Mobile Devices http://dx.doi.org/10.1109/TMM.2013.2283468 ..... *W. Yin, T. Mei, C. W. Chen, and S. Li* 184

Broadcasting Oneself: Visual Discovery of Vlogging Styles http://dx.doi.org/10.1109/TMM.2013.2284893 ..... *O. Aran, J.-I. Biel, and D. Gatica-Perez* 201

Error Resilience and Concealment

Motion Vector Recovery for Video Error Concealment by Using Iterative Dynamic-Programming Optimization http://dx.doi.org/10.1109/TMM.2013.2281587 ..... *W.-N. Lie, C.-M. Lee, C.-H. Yeh, and Z.-W. Gao* 216

Multimedia Steaming and Transport

UMSM: A Traffic Reduction Method on Multi-View Video Streaming for Multiple Users http://dx.doi.org/10.1109/TMM.2013.2281588 ..... *T. Fujihashi, Z. Pan, and T. Watanabe* 228

CBM: Online Strategies on Cost-Aware Buffer Management for Mobile Video Streaming http://dx.doi.org/10.1109/TMM.2013.2284894 ..... *J. He, Z. Xue, D. Wu, D. O. Wu, and Y. Wen* 242

Multimedia Algorithms, Systems, and Interfaces

Fashion Parsing With Weak Color-Category Labels http://dx.doi.org/10.1109/TMM.2013.2285526 ..... *S. Liu, J. Feng, C. Domokos, H. Xu, J. Huang, Z. Hu, and S. Yan* 253

CORRESPONDENCES

Compression and Coding

Low Complexity Adaptive View Synthesis Optimization in HEVC Based 3D Video Coding http://dx.doi.org/10.1109/TMM.2013.2284751 ..... *S. Ma, S. Wang, and W. Gao* 266

Display Technology for Multimedia

Effective Video Retargeting With Jittery Assessment http://dx.doi.org/10.1109/TMM.2013.2286112 ..... *B. Yan, B. Yuan, and B. Yang* 272

Multimedia Steaming and Transport

Random Network Coding for Multimedia Delivery Services in LTE/LTE-Advanced http://dx.doi.org/10.1109/TMM.2013.2282129 ..... *D. Vukobratović, C. Khirallah, V. Stanković, and J. S. Thompson* 277

EDICS—Editors Classification Scheme http://dx.doi.org/10.1109/TMM.2013.2294429 283

Information for Authors http://dx.doi.org/10.1109/TMM.2013.2294880 284

ANNOUNCEMENTS

Call for Papers—Journal of Selected Topics in Signal Processing, Special Issue on Signal Processing for Situational Awareness from Networked Sensors and Social Media http://dx.doi.org/10.1109/TMM.2013.2294430 286

Call for Papers—IEEE Transactions on Information Forensics and Security, Special Issue on Facial Biometrics in the Wild. http://dx.doi.org/10.1109/TMM.2013.2294746 287

IEEE TRANSACTIONS ON MULTIMEDIA

A PUBLICATION OF
THE IEEE CIRCUITS AND SYSTEMS SOCIETY
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY

http://www.signalprocessingsociety.org/tmm/

FEBRUARY 2014 VOLUME 16 NUMBER 2 ITMUF8 (ISSN 1520-9210)

PAPERS

Audio/Video Analysis and Synthesis

Semi-Supervised Multiple Feature Analysis for Action Recognition http://dx.doi.org/10.1109/TMM.2013.2293060 ..... *S. Wang, Z. Ma, Y. Yang, X. Li, C. Pang, and A. G. Hauptmann* 289

A Data-Driven Approach for Facial Expression Retargeting in Video http://dx.doi.org/10.1109/TMM.2013.2293064 ..... *K. Li, Q. Dai, R. Wang, Y. Liu, F. Xu, and J. Wang* 299

Watermarking, Encryption, and Data Hiding

Adaptive Watermarking and Tree Structure Based Image Quality Estimation http://dx.doi.org/10.1109/TMM.2013.2291658 ..... *S. Wang, D. Zheng, J. Zhao, W. J. Tam, and F. Speranza* 311

3-D Audio/Video Processing

Efficient Viewer-Centric Depth Adjustment Based on Virtual Fronto-Parallel Planar Projection in Stereo 3D Images http://dx.doi.org/10.1109/TMM.2013.2286567 ..... *H. Park, H. Lee, and S. Sull* 326

3-D Presentation

Point Cloud Encoding for 3D Building Model Retrieval http://dx.doi.org/10.1109/TMM.2013.2286580 ..... *J.-Y. Chen, C.-H. Lin, P.-C. Hsu, and C.-H. Chen* 337

System Software for Multimedia

Towards Mobile Document Image Retrieval for Digital Library http://dx.doi.org/10.1109/TMM.2013.2293063 ..... *L.-Y. Duan, R. Ji, Z. Chen, T. Huang, and W. Gao* 346

Quality Assessment and User Experience

An Analytical Approach for Voice Capacity Estimation Over WiFi Network Using ITU-T E-Model http://dx.doi.org/10.1109/TMM.2013.2291212 ..... *M. A. R. Siddique, J. Kamruzzaman, and M. J. Hossain* 360

Online HodgeRank on Random Graphs for Crowdsourceable QoE Evaluation http://dx.doi.org/10.1109/TMM.2013.2292568 ..... *Q. Xu, J. Xiong, Q. Huang, and Y. Yao* 373

Subjective Test Methodology and Databases

Multimodal Interactive Continuous Scoring of Subjective 3D Video Quality of Experience http://dx.doi.org/10.1109/TMM.2013.2292592 ..... *T. Kim, J. Kang, S. Lee, and A. C. Bovik* 387

Content Description and Annotation

Multi-Label Learning With Fused Multimodal Bi-Relational Graph http://dx.doi.org/10.1109/TMM.2013.2291218 ..... *J. Xu, V. Jagadeesh, and B. S. Manjunath* 403

Discriminative Structure Learning for Semantic Concept Detection With Graph Embedding http://dx.doi.org/10.1109/TMM.2013.2291657 ..... *M. Jian, C. Jung, and Y. Zheng* 413

Multimedia Search and Retrieval

Sparse Multi-Modal Hashing http://dx.doi.org/10.1109/TMM.2013.2291214 ..... *F. Wu, Z. Yu, Y. Yang, S. Tang, Y. Zhang, and Y. Zhuang* 427

Gaze-Based Relevance Feedback for Realizing Region-Based Image Retrieval http://dx.doi.org/10.1109/TMM.2013.2291535 ..... *G. T. Papadopoulos, K. C. Apostolakis, and P. Daras* 440

Resource Allocation for Personalized Video Summarization http://dx.doi.org/10.1109/TMM.2013.2291967 ..... *F. Chen, C. De Vleeschouwer, and A. Cavallaro* 455

Representative Discovery of Structure Cues for Weakly-Supervised Image Segmentation http://dx.doi.org/10.1109/TMM.2013.2293424 ..... *L. Zhang, Y. Gao, Y. Xia, K. Lu, J. Shen, and R. Ji* 470

Realtime Communication and Video Conferencing

On the Quality of Service of Cloud Gaming Systems http://dx.doi.org/10.1109/TMM.2013.2291532 ..... *K.-T. Chen, Y.-C. Chang, H.-J. Hsu, D.-Y. Chen, C.-Y. Huang, and C.-H. Hsu* 480

Multimedia Streaming and Transport

Correlation-Aware Packet Scheduling in Multi-Camera Networks http://dx.doi.org/10.1109/TMM.2013.2291531 ..... *L. Toni, T. Maugey, and P. Frossard* 496

Investigating Redundant Internet Video Streaming Traffic on iOS Devices: Causes and Solutions http://dx.doi.org/10.1109/TMM.2013.2293312 ..... *Y. Liu, Q. Wei, L. Guo, B. Shen, S. Chen, and Y. Lan* 510

Distributive/Cooperative Networks and Communication

Band Codes for Energy-Efficient Network Coding With Application to P2P Mobile Streaming http://dx.doi.org/10.1109/TMM.2013.2285518 ..... *A. Fiandrotti, V. Bioglio, M. Grangetto, R. Gaeta, and E. Magli* 521

Wireless/Mobile Multimedia

Multipath Video Real-Time Streaming by Field-Based Anycast Routing http://dx.doi.org/10.1109/TMM.2013.2293315 ..... *M. Kserawi, S. Jung, D. Lee, J. Sung, and J.-K. K. Rhee* 533

Multimedia and Crowdsourcing

Best Practices for QoE Crowdtesting: QoE Assessment With Crowdsourcing http://dx.doi.org/10.1109/TMM.2013.2291663 ..... *T. Hoßfeld, C. Keimel, M. Hirth, B. Gardlo, J. Habigt, K. Diepold, and P. Tran-Gia* 541

CORRESPONDENCES

Compression and Coding

A Fast HEVC Inter CU Selection Method Based on Pyramid Motion Divergence http://dx.doi.org/10.1109/TMM.2013.2291958 ..... *J. Xiong, H. Li, Q. Wu, and F. Meng* 559

Subjective Test Methodology and Databases

On Designing Paired Comparison Experiments for Subjective Multimedia Quality Assessment http://dx.doi.org/10.1109/TMM.2013.2292590 ..... *J.-S. Lee* 564

Multimedia Algorithms, Systems, and Interfaces

Solving a Special Type of Jigsaw Puzzles: Banknote Reconstruction From a Large Number of Fragments http://dx.doi.org/10.1109/TMM.2013.2291968 ..... *H. Li, Y. Zheng, S. Zhang, and J. Cheng* 571

EDICS—Editors Classification Scheme http://dx.doi.org/10.1109/TMM.2014.2300395 579

Information for Authors http://dx.doi.org/10.1109/TMM.2014.2300394 580

ANNOUNCEMENTS

Call for Papers—JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING Special Issue on Signal Processing for Situational Awareness from Networked Sensors and Social Media http://dx.doi.org/10.1109/TMM.2014.2300371 582

Call for Papers—IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY Special Issue on Biometric Spoofing and Countermeasures http://dx.doi.org/10.1109/TMM.2014.2300375 583

Call for Nominations—IEEE TRANSACTIONS ON MULTIMEDIA 2014 Best Paper Award http://dx.doi.org/10.1109/TMM.2014.2300391 584

IEEE JOURNAL OF
SELECTED TOPICS IN SIGNAL PROCESSING

IEEE Signal Processing Society

www.ieee.org/sp/index.html

FEBRUARY 2014 VOLUME 8 NUMBER 1 IJSTGY (ISSN 1932-4553)

EDITORIAL

Introduction to the Special Issue on Non-Cooperative Localization Networks http://dx.doi.org/10.1109/JSTSP.2013.2291302 ..... R. M. Buehrer, C. R. Anderson, R. K. Martin, N. Patwari, and M. G. Rabbat 2

PAPERS

Ichnaea: A Low-Overhead Robust WLAN Device-Free Passive Localization System http://dx.doi.org/10.1109/JSTSP.2013.2287480 ..... A. Saeed, A. E. Kosba, and M. Youssef 5

A Bayesian Approach to Device-Free Localization: Modeling and Experimental Assessment http://dx.doi.org/10.1109/JSTSP.2013.2286772 .. S. Savazzi, M. Nicoli, F. Carminati, and M. Riva 16

Breathfinding: A Wireless Network That Monitors and Locates Breathing in a Home http://dx.doi.org/10.1109/JSTSP.2013.2287473 ..... N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca 30

Device-Free Person Detection and Ranging in UWB Networks http://dx.doi.org/10.1109/JSTSP.2013.2281780 ..... Y. Kilic, H. Wymeersch, A. Meijerink, M. J. Bentum, and W. G. Scanlon 43

Propagation Modeling for Radio Frequency Tomography in Wireless Networks http://dx.doi.org/10.1109/JSTSP.2013.2287471 ..... B. R. Hamilton, X. Ma, R. J. Baxley, and S. M. Matechik 55

Radio Tomography for Roadside Surveillance http://dx.doi.org/10.1109/JSTSP.2013.2286774 ..... C. R. Anderson, R. K. Martin, T. O. Walker, and R. W. Thomas 66

A Novel Joint RFID and Radar Sensor Network for Passive Localization: Design and Performance Bounds http://dx.doi.org/10.1109/JSTSP.2013.2287174 ..... N. Decarli, F. Guidi, and D. Dardari 80

Indoor Tracking With RFID Systems http://dx.doi.org/10.1109/JSTSP.2013.2286972 .... L. Geng, M. F. Bugallo, A. Athalye, and P. M. Djurić 96

Range-Angle Localization of Targets by A Double-Pulse Frequency Diverse Array Radar http://dx.doi.org/10.1109/JSTSP.2013.2285528 .... W.-Q. Wang and H. Shao 106

Joint Transmission and Reception Diversity Smoothing for Direction Finding of Coherent Targets in MIMO Radar http://dx.doi.org/10.1109/JSTSP.2013.2285520 ..... W. Zhang, W. Liu, J. Wang, and S. Wu 115

Target Tracking for UWB Multistatic Radar Sensor Networks http://dx.doi.org/10.1109/JSTSP.2013.2286771 ..... B. Sobhani, E. Paolini, A. Giorgetti, M. Mazzotti, and M. Chiani 125

Multipath Exploitation in Non-LOS Urban Synthetic Aperture Radar http://dx.doi.org/10.1109/JSTSP.2013.2287185 ..... P. Setlur, T. Negishi, N. Devroye, and D. Erricolo 137

Simultaneous Target and Multipath Positioning http://dx.doi.org/10.1109/JSTSP.2013.2289949 ..... L. Li and J. L. Krolik 153

Information for Authors http://dx.doi.org/10.1109/JSTSP.2014.2299197 166

ANNOUNCEMENTS

Call for Papers—Special Issue on Signal Processing for Situational Awareness from Networked Sensors and Social Media http://dx.doi.org/10.1109/JSTSP.2014.2299394 168

IEEE Journal on Selected Topics in Signal Processing

Special Issue on Interactive Media Processing for Immersive Communication

Interpersonal communication is an essential and intrinsic element of the human way of life. Yet, for interpersonal communication among individuals connected via high-speed data networks, the prevalent technologies are limited to exchanging 2D video and audio captured by a single camera and microphone. These technologies fail to provide a level of immersivity necessary for "in the same room" sense of presence, due to limitations such as gaze mismatch and lack of depth perception in the rendered scene.

With the recent advances in sensing technologies and accompanying data analysis tools, one can now acquire a large collection of media data describing both the sender's physical environment (e.g., texture/depth images and audio captured from arrays of cameras and microphones) and the manner in which the sender is perceiving the presented media (e.g., gaze & head movements). This means that although the viewer's display capabilities may remain limited (e.g., 2D display, stereo speakers), the sense of immersion can be greatly enhanced through innovative human-centric interaction with the presented media (e.g., gazed-corrected view, seamless view-switching and/or spatial audio corresponding to viewer's tracked head position, haptic vibrations in response to loud audio events). This enhanced media interaction must be designed within the context of delay- and loss-prone networks for delivery of delay-sensitive data, towards the ultimate goal of improving immersive communication beyond 2D video communication. In particular, the main technical challenges are: i) efficient acquisition and processing of observer's sensory data, ii) compact/robust representation of media data for network transport given the receiver's current patterns of media consumption and display limitations, iii) real-time human-centric media interaction for an enriched immersive experience. Further, evaluation methodologies and metrics for immersive communication systems must be subjectively accurate for the range of systems deploying multiple sensory input and output devices.

We invite authors to address aspects of immersive communication related to interactive media processing, such as the following. Please note that submission of pure video compression papers with no direct connection to media interaction is not encouraged.

- ***Interactive Visual Communication*** (e.g., efficient view switching systems enabling motion parallax, 3D visual representation with desirable properties like flexible decoding, robustness, etc.).
- ***Acquisition & Reconstruction of Media Data for Immersivity*** (e.g., depth data acquisition and pre-processing, microphone array design for spatial audio, hybrid approaches to depth estimation, gaze/head tracking and prediction).
- ***Multi-modal Media Interaction*** (e.g., multi-modal responses to detected communication events, visual interaction (defocusing, saliency-based content adaptation) based on gaze patterns and/or pupil size).
- ***Streaming/Transport of Immersive Media Data*** (e.g., media-specific, delay-sensitive FEC, multiple descriptions for multi-path transmission, multi-modal loss concealment strategies).
- ***Applications of Immersive Communication*** (e.g., systems for tele-medicine/education, immersive video conferencing).
- ***Quality Issues in Immersive Communication*** (e.g., evaluation methodologies and metrics for immersive systems).

Prospective authors should visit http://www.signalprocessingsociety.org/publications/periodicals/jstsp/ for information on paper submission. Manuscripts should be submitted at http://mc.manuscriptcentral.com/jstsp-ieee.

Manuscript Submission	First Review Due	Revised Manuscript	Second Review Due	Final Manuscript
April 2, 2014	July 1, 2014	September 1, 2014	November 1, 2014	December 1, 2014

Guest Editors:

Gene Cheung, National Institute of Informatics, Japan (cheung@nii.ac.jp)

Dinei Florencio, Microsoft Research, USA (dinei@micrsoft.com)

Patrick Le Callet, University of Nantes, France (patrick.lecallet@univ-nantes.fr)

Chia-Wen Lin, National Tsing-Hwa University, Taiwan (cwlin@ee.nthu.edu.tw)

Enrico Magli, Politecnico di Torino, Italy (enrico.magli@polito.it)

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

JANUARY 2014 VOLUME 21 NUMBER 1 ISPLEM (ISSN 1070-9908)

Editorial http://dx.doi.org/10.1109/LSP.2013.2290876 ..... *P. Willett* 5

LETTERS

Explicit Shift-Invariant Dictionary Learning http://dx.doi.org/10.1109/LSP.2013.2288788 ..... *C. Rusu, B. Dumitrescu, and S. A. Tsaftaris* 6

On the Chi Square and Higher-Order Chi Distances for Approximating f-Divergences http://dx.doi.org/10.1109/LSP.2013.2288355 ..... *F. Nielsen and R. Nock* 10

Lattice Structure for Regular Linear Phase Paraunitary Filter Bank With Odd Decimation Factor http://dx.doi.org/10.1109/LSP.2013.2285435 ..... *B. Li and X. Gao* 14

A Decentralized Positioning Method Based on Recursive Weighted Least Absolute Value Optimization for Wireless Sensor Networks http://dx.doi.org/10.1109/LSP.2013.2288037 ..... *D.-S. Wu and C.-L. Wang* 18

Extended Layered Depth Image Representation in Multiview Navigation http://dx.doi.org/10.1109/LSP.2013.2288014 ..... *U. Takyar, T. Maugey, and P. Frossard* 22

Sparse Direction of Arrival Estimation Using Co-Prime Arrays with Off-Grid Targets http://dx.doi.org/10.1109/LSP.2013.2289740 ..... *Z. Tan and A. Nehorai* 26

A Robust Particle Filtering Algorithm With Non-Gaussian Measurement Noise Using Student-t Distribution http://dx.doi.org/10.1109/LSP.2013.2289975 ..... *D. Xu, C. Shen, and F. Shen* 30

Visually Lossless Strategies to Decode and Transmit JPEG2000 Imagery http://dx.doi.org/10.1109/LSP.2013.2290317 ..... *L. Jiménez-Rodríguez, F. Aulí-Llinàs, and M. W. Marcellin* 35

Joint Optimization of Transmit and Receive Beamforming in Active Arrays http://dx.doi.org/10.1109/LSP.2013.2289325 ..... *J. Liu, H. Li, and B. Himed* 39

Statistical and Adaptive Signal Processing

Spherical Simplex-Radial Cubature Kalman Filter http://dx.doi.org/10.1109/LSP.2013.2290381 ..... *S. Wang, J. Feng, and C. K. Tse* 43

Computing the Cramer–Rao Bound of Markov Random Field Parameters: Application to the Ising and the Potts Models http://dx.doi.org/10.1109/LSP.2013.2290329 ..... *M. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret* 47

IEEE SIGNAL PROCESSING LETTERS (ISSN 1070–9908) is published quarterly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $304.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE SIGNAL PROCESSING LETTERS, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

Salient Object Detection via Random Forest http://dx.doi.org/10.1109/LSP.2013.2290547 ..... *S. Du and S. Chen* 51
A Generalized Chinese Remainder Theorem for Two Integers http://dx.doi.org/10.1109/LSP.2013.2289326 ..... *L. Xiao and X.-G. Xia* 55
Robust Adaptive Beamforming Using a Low-Complexity Shrinkage-Based Mismatch Estimation Algorithm http://dx.doi.org/10.1109/LSP.2013.2290948 ..... *H. Ruan and R. C. de Lamare* 60
An Experimental Study on Speech Enhancement Based on Deep Neural Networks http://dx.doi.org/10.1109/LSP.2013.2291240 ..... *Y. Xu, J. Du, L.-R. Dai, and C.-H. Lee* 65
Digital Suppression of Power Amplifier Spurious Emissions at Receiver Band in FDD Transceivers http://dx.doi.org/10.1109/LSP.2013.2291032 ..... *A. Kiayani, L. Anttila, and M. Valkama* 69
Pixel-Based Inter Prediction in Coded Texture Assisted Depth Coding http://dx.doi.org/10.1109/LSP.2013.2291941 ..... *S. Li, J. Lei, C. Zhu, L. Yu, and C. Hou* 74
Low-Rank Neighbor Embedding for Single Image Super-Resolution http://dx.doi.org/10.1109/LSP.2013.2286417 ..... *X. Chen and C. Qi* 79
Achieving Perfect Secrecy by pdf-Bandlimited Jamming http://dx.doi.org/10.1109/LSP.2013.2292094 ..... *S. Marano and V. Matta* 83
Co-Saliency Detection Based on Hierarchical Segmentation http://dx.doi.org/10.1109/LSP.2013.2292873 ..... *Z. Liu, W. Zou, L. Li, L. Shen, and O. Le Meur* 88
Noise-Robust Texture Description Using Local Contrast Patterns via Global Measures http://dx.doi.org/10.1109/LSP.2013.2293335 ..... *T. Song, H. Li, F. Meng, Q. Wu, B. Luo, B. Zeng, and M. Gabbouj* 93
Optimization of Bilateral Filter Parameters via Chi-Square Unbiased Risk Estimate http://dx.doi.org/10.1109/LSP.2013.2293592 ..... *Y. Chen and Y. Shu* 97
Learning Boundary and Appearance for Video Object Cutout http://dx.doi.org/10.1109/LSP.2013.2293778 ..... *S. Chen, Q. Zhou, and H. Ding* 101
Improving Time-Scale Modification of Music Signals Using Harmonic-Percussive Separation http://dx.doi.org/10.1109/LSP.2013.2294023 ..... *J. Driedger, M. Müller, and S. Ewert* 105
Hierarchical Sparse Signal Recovery by Variational Bayesian Inference http://dx.doi.org/10.1109/LSP.2013.2292589 ..... *L. Wang, L. Zhao, G. Bi, and C. Wan* 110
A Conic Quadratic Programming Approach to Physical Layer Multicasting for Large-Scale Antenna Arrays http://dx.doi.org/10.1109/LSP.2013.2293840 ..... *L.-N. Tran, M. F. Hanif, and M. Juntti* 114
A Safe Approximation Approach to Secrecy Outage Design for MIMO Wiretap Channels http://dx.doi.org/10.1109/LSP.2013.2293884 ..... *Q. Li, W.-K. Ma, and A. M.-C. So* 118
Referenceless Measure of Blocking Artifacts by Tchebichef Kernel Analysis http://dx.doi.org/10.1109/LSP.2013.2294333 ..... *L. Li, H. Zhu, G. Yang, and J. Qian* 122

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

JANUARY 2014 VOLUME 21 NUMBER 1 ISPLEM (ISSN 1070-9908)

Editorial http://dx.doi.org/10.1109/LSP.2013.2290876 ..... *P. Willett* 5

LETTERS

Audio and Acoustic Signal Processing

An Experimental Study on Speech Enhancement Based on Deep Neural Networks http://dx.doi.org/10.1109/LSP.2013.2291240 ..... *Y. Xu, J. Du, L.-R. Dai, and C.-H. Lee* 65

Improving Time-Scale Modification of Music Signals Using Harmonic-Percussive Separation http://dx.doi.org/10.1109/LSP.2013.2294023 ..... *J. Driedger, M. Müller, and S. Ewert* 105

Digital and Multirate Signal Processing

Lattice Structure for Regular Linear Phase Paraunitary Filter Bank With Odd Decimation Factor http://dx.doi.org/10.1109/LSP.2013.2285435 ..... *B. Li and X. Gao* 14

A Generalized Chinese Remainder Theorem for Two Integers http://dx.doi.org/10.1109/LSP.2013.2289326 ..... *L. Xiao and X.-G. Xia* 55

Image and Multidimensional Signal Processing

Extended Layered Depth Image Representation in Multiview Navigation http://dx.doi.org/10.1109/LSP.2013.2288014 ..... *U. Takyar, T. Maugey, and P. Frossard* 22

Visually Lossless Strategies to Decode and Transmit JPEG2000 Imagery http://dx.doi.org/10.1109/LSP.2013.2290317 ..... *L. Jiménez-Rodríguez, F. Aulí-Llinàs, and M. W. Marcellin* 35

Salient Object Detection via Random Forest http://dx.doi.org/10.1109/LSP.2013.2290547 ..... *S. Du and S. Chen* 51

Pixel-Based Inter Prediction in Coded Texture Assisted Depth Coding http://dx.doi.org/10.1109/LSP.2013.2291941 ..... *S. Li, J. Lei, C. Zhu, L. Yu, and C. Hou* 74

Low-Rank Neighbor Embedding for Single Image Super-Resolution http://dx.doi.org/10.1109/LSP.2013.2286417 ..... *X. Chen and C. Qi* 79

Co-Saliency Detection Based on Hierarchical Segmentation http://dx.doi.org/10.1109/LSP.2013.2292873 ..... *Z. Liu, W. Zou, L. Li, L. Shen, and O. Le Meur* 88

IEEE SIGNAL PROCESSING LETTERS (ISSN 1070–9908) is published quarterly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $304.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2013 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE SIGNAL PROCESSING LETTERS, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

Noise-Robust Texture Description Using Local Contrast Patterns via Global Measures http://dx.doi.org/10.1109/LSP.2013.2293335 ..... T. Song, H. Li, F. Meng, Q. Wu, B. Luo, B. Zeng, and M. Gabbouj 93
Optimization of Bilateral Filter Parameters via Chi-Square Unbiased Risk Estimate http://dx.doi.org/10.1109/LSP.2013.2293592 ..... Y. Chen and Y. Shu 97
Learning Boundary and Appearance for Video Object Cutout http://dx.doi.org/10.1109/LSP.2013.2293778 ..... S. Chen, Q. Zhou, and H. Ding 101
Referenceless Measure of Blocking Artifacts by Tchebichef Kernel Analysis http://dx.doi.org/10.1109/LSP.2013.2294333 ..... L. Li, H. Zhu, G. Yang, and J. Qian 122

Machine Learning and Statistical Signal Processing

Explicit Shift-Invariant Dictionary Learning http://dx.doi.org/10.1109/LSP.2013.2288788 ..... C. Rusu, B. Dumitrescu, and S. A. Tsaftaris 6
Hierarchical Sparse Signal Recovery by Variational Bayesian Inference http://dx.doi.org/10.1109/LSP.2013.2292589 ..... L. Wang, L. Zhao, G. Bi, and C. Wan 110

Sensor Array and Multichannel Signal Processing

Sparse Direction of Arrival Estimation Using Co-Prime Arrays with Off-Grid Targets http://dx.doi.org/10.1109/LSP.2013.2289740 ..... Z. Tan and A. Nehorai 26
Joint Optimization of Transmit and Receive Beamforming in Active Arrays http://dx.doi.org/10.1109/LSP.2013.2289325 ..... J. Liu, H. Li, and B. Himed 39
Robust Adaptive Beamforming Using a Low-Complexity Shrinkage-Based Mismatch Estimation Algorithm http://dx.doi.org/10.1109/LSP.2013.2290948 ..... H. Ruan and R. C. de Lamare 60

Signal Processing for Communications

A Decentralized Positioning Method Based on Recursive Weighted Least Absolute Value Optimization for Wireless Sensor Networks http://dx.doi.org/10.1109/LSP.2013.2288037 ..... D.-S. Wu and C.-L. Wang 18
Achieving Perfect Secrecy by pdf-Bandlimited Jamming http://dx.doi.org/10.1109/LSP.2013.2292094 ..... S. Marano and V. Matta 83
A Conic Quadratic Programming Approach to Physical Layer Multicasting for Large-Scale Antenna Arrays http://dx.doi.org/10.1109/LSP.2013.2293840 ..... L.-N. Tran, M. F. Hanif, and M. Juntti 114
A Safe Approximation Approach to Secrecy Outage Design for MIMO Wiretap Channels http://dx.doi.org/10.1109/LSP.2013.2293884 ..... Q. Li, W.-K. Ma, and A. M.-C. So 118

Statistical and Adaptive Signal Processing

On the Chi Square and Higher-Order Chi Distances for Approximating f-Divergences http://dx.doi.org/10.1109/LSP.2013.2288355 ..... F. Nielsen and R. Nock 10
A Robust Particle Filtering Algorithm With Non-Gaussian Measurement Noise Using Student-t Distribution http://dx.doi.org/10.1109/LSP.2013.2289975 ..... D. Xu, C. Shen, and F. Shen 30
Spherical Simplex-Radial Cubature Kalman Filter http://dx.doi.org/10.1109/LSP.2013.2290381 ..... S. Wang, J. Feng, and C. K. Tse 43
Computing the Cramer–Rao Bound of Markov Random Field Parameters: Application to the Ising and the Potts Models http://dx.doi.org/10.1109/LSP.2013.2290329 ..... M. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret 47
Digital Suppression of Power Amplifier Spurious Emissions at Receiver Band in FDD Transceivers http://dx.doi.org/10.1109/LSP.2013.2291032 ..... A. Kiayani, L. Anttila, and M. Valkama 69

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

FEBRUARY 2014 VOLUME 21 NUMBER 2 ISPLEM (ISSN 1070-9908)

LETTERS

Localization of More Sources Than Sensors via Jointly-Sparse Bayesian Learning http://dx.doi.org/10.1109/LSP.2013.2294862 ..... *O. Balkan, K. Kreutz-Delgado, and S. Makeig* 131

Optimal Prediction Likelihood Tree Based Source- Channel ML Decoder for Wireless Sensor Networks http://dx.doi.org/10.1109/LSP.2013.2294794 ..... *C. Manoj and A. K. Jagannatham* 135

Improved Unitary Root-MUSIC for DOA Estimation Based on Pseudo-Noise Resampling http://dx.doi.org/10.1109/LSP.2013.2294676 ..... *C. Qian, L. Huang, and H. C. So* 140

Sigma Point Belief Propagation http://dx.doi.org/10.1109/LSP.2013.2290192 ..... *F. Meyer, O. Hlinka, and F. Hlawatsch* 145

Adaptive Support of Spatial–Temporal Neighbors for Depth Map Sequence Up-sampling http://dx.doi.org/10.1109/LSP.2013.2295252 ..... *M.-K. Kang, D.-Y Kim, and K.-J. Yoon* 150

No-Reference Quality Assessment of JPEG Images via a Quality Relevance Map http://dx.doi.org/10.1109/LSP.2013.2296038 ..... *S. A. Golestaneh and D. M. Chandler* 155

Estimating Speaker Height and Subglottal Resonances Using MFCCs and GMMs http://dx.doi.org/10.1109/LSP.2013.2295397 ..... *H. Arsikere, S. M. Lulich, and A. Alwan* 159

Optimal Non-regenerative Linear MIMO Relay for Orthogonal Space Time Codes http://dx.doi.org/10.1109/LSP.2013.2295919 ..... *M. H. Shariat and S. Gazor* 163

Joint Bayesian Estimation of Close Subspaces from Noisy Measurements http://dx.doi.org/10.1109/LSP.2013.2296138 ..... *O. Besson, N. Dobigeon, and J.-Y. Tourneret* 168

3-D Face Recognition Using Curvelet Local Features http://dx.doi.org/10.1109/LSP.2013.2295119 ..... *S. Elaiwat, M. Bennamoun, F. Boussaid, and A. El-Sallam* 172

New Feedback Topology Designs with Reduced CSI Overhead for MIMO Interference Alignment http://dx.doi.org/10.1109/LSP.2013.2296492 ..... *J. Jin, L. Li, H. Tian, Q. Wang, and X.-C. Gao* 176

The Significant Gains From Optimally Processed Multiple Signals of Opportunity and Multiple Receive Stations in Passive Radar http://dx.doi.org/10.1109/LSP.2013.2296369 ..... *Q. He and R. S. Blum* 180

Maximum Phase Modeling for Sparse Linear Prediction of Speech http://dx.doi.org/10.1109/LSP.2013.2296944 ..... *T. Drugman* 185

Online Discriminative Structured Output SVM Learning for Multi-Target Tracking http://dx.doi.org/10.1109/LSP.2013.2296602 ..... *Y. Xu, L. Qin, G. Li, and Q. Huang* 190

IEEE SIGNAL PROCESSING LETTERS (ISSN 1070–9908) is published quarterly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $304.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE SIGNAL PROCESSING LETTERS, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

Rate-adaptive Compact Fisher Codes for Mobile Visual Search http://dx.doi.org/10.1109/LSP.2013.2296532 ..... J. Lin, L.-Y. Duan, Y. Huang, S. Luo, T. Huang, and W. Gao 195
Rank-Two Beamformed Secure Multicasting for Wireless Information and Power Transfer http://dx.doi.org/10.1109/LSP.2013.2296941 .... B. Zhu, J. Ge, Y. Huang, Y. Yang, and M. Lin 199
Transceiver Design for Hybrid One-Way and Two-Way Relay Networks http://dx.doi.org/10.1109/LSP.2013.2296975 ..... M. Wang, P. Wang, Y. Li, Z. Zhong, F. Wang, and B. Vucetic 204
Decision Fusion With Unknown Sensor Detection Probability http://dx.doi.org/10.1109/LSP.2013.2295054 ..... D. Ciuonzo and P. Salvo Rossi 208
A Statistical-CSI-Based Scheme for Multiple Description Coding Multicast in CRNs http://dx.doi.org/10.1109/LSP.2013.2297338 ..... K. Yang, W. Xu, S. Li, J. Lin, and W. Wu 213
Sparsity-Aware Sensor Selection: Centralized and Distributed Algorithms http://dx.doi.org/10.1109/LSP.2013.2297419 ..... H. Jamali-Rad, A. Simonetto, and G. Leus 217
Alternating Rate Profile Optimization in Single Stream MIMO Interference Channels http://dx.doi.org/10.1109/LSP.2013.2297351 ..... R. Mochaourab, P. Cao, and E. Jorswieck 221
Sparsity-Aware Adaptive Algorithms Based on Alternating Optimization and Shrinkage http://dx.doi.org/10.1109/LSP.2014.2298116 ..... R. C. de Lamare and R. Sampaio-Neto 225
Sparse Array-Based Room Transfer Function Estimation for Echo Cancellation http://dx.doi.org/10.1109/LSP.2013.2297832 ..... A. Yellepeddi and D. Florencio 230
Adaptive Transmit and Receive Beamforming for Interference Mitigation http://dx.doi.org/10.1109/LSP.2014.2298497 ..... Z. Chen, H. Li, G. Cui, and M. Rangaswamy 235
About QLMS Derivations http://dx.doi.org/10.1109/LSP.2014.2299066 ..... Q. Barthélemy, A. Larue, and J. I. Mars 240
A Failure-Aware Explicit Shape Regression Model for Facial Landmark Detection in Video http://dx.doi.org/10.1109/LSP.2013.2295231 ... M. Zhang, L. Tao, Y. Zheng, and Y. Du 244

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

FEBRUARY 2014 VOLUME 21 NUMBER 2 ISPLEM (ISSN 1070-9908)

LETTERS

Audio and Acoustic Signal Processing

Sparse Array-Based Room Transfer Function Estimation for Echo Cancellation http://dx.doi.org/10.1109/LSP.2013.2297832 ..... *A. Yellepeddi and D. Florencio* 230

Image and Multidimensional Signal Processing

Adaptive Support of Spatial–Temporal Neighbors for Depth Map Sequence Up-sampling http://dx.doi.org/10.1109/LSP.2013.2295252 ..... *M.-K. Kang, D.-Y Kim, and K.-J. Yoon* 150

No-Reference Quality Assessment of JPEG Images via a Quality Relevance Map http://dx.doi.org/10.1109/LSP.2013.2296038 ..... *S. A. Golestaneh and D. M. Chandler* 155

3-D Face Recognition Using Curvelet Local Features http://dx.doi.org/10.1109/LSP.2013.2295119 ..... *S. Elaiwat, M. Bennamoun, F. Boussaid, and A. El-Sallam* 172

Online Discriminative Structured Output SVM Learning for Multi-Target Tracking http://dx.doi.org/10.1109/LSP.2013.2296602 ..... *Y. Xu, L. Qin, G. Li, and Q. Huang* 190

Rate-adaptive Compact Fisher Codes for Mobile Visual Search http://dx.doi.org/10.1109/LSP.2013.2296532 ..... *J. Lin, L.-Y. Duan, Y. Huang, S. Luo, T. Huang, and W. Gao* 195

A Failure-Aware Explicit Shape Regression Model for Facial Landmark Detection in Video http://dx.doi.org/10.1109/LSP.2013.2295231 ... *M. Zhang, L. Tao, Y. Zheng, and Y. Du* 244

Machine Learning and Statistical Signal Processing

Localization of More Sources Than Sensors via Jointly-Sparse Bayesian Learning http://dx.doi.org/10.1109/LSP.2013.2294862 ..... *O. Balkan, K. Kreutz-Delgado, and S. Makeig* 131

IEEE SIGNAL PROCESSING LETTERS (ISSN 1070–9908) is published quarterly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members $20.00 (first copy only), nonmembers $304.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2014 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals Postage at New York, NY and at additional mailing offices. **Postmaster:** Send address changes to IEEE SIGNAL PROCESSING LETTERS, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit http://www.ieee.org/nondiscrimination. Printed in U.S.A.

Sensor Array and Multichannel Signal Processing

Improved Unitary Root-MUSIC for DOA Estimation Based on Pseudo-Noise Resampling http://dx.doi.org/10.1109/LSP.2013.2294676 ..... *C. Qian, L. Huang, and H. C. So* 140

The Significant Gains From Optimally Processed Multiple Signals of Opportunity and Multiple Receive Stations in Passive Radar http://dx.doi.org/10.1109/LSP.2013.2296369 ..... *Q. He and R. S. Blum* 180

Sparsity-Aware Sensor Selection: Centralized and Distributed Algorithms http://dx.doi.org/10.1109/LSP.2013.2297419 ..... *H. Jamali-Rad, A. Simonetto, and G. Leus* 217

Adaptive Transmit and Receive Beamforming for Interference Mitigation http://dx.doi.org/10.1109/LSP.2014.2298497 ..... *Z. Chen, H. Li, G. Cui, and M. Rangaswamy* 235

Signal Processing for Communications

Optimal Prediction Likelihood Tree Based Source- Channel ML Decoder for Wireless Sensor Networks http://dx.doi.org/10.1109/LSP.2013.2294794 ..... *C. Manoj and A. K. Jagannatham* 135

Optimal Non-regenerative Linear MIMO Relay for Orthogonal Space Time Codes http://dx.doi.org/10.1109/LSP.2013.2295919 ..... *M. H. Shariat and S. Gazor* 163

New Feedback Topology Designs with Reduced CSI Overhead for MIMO Interference Alignment http://dx.doi.org/10.1109/LSP.2013.2296492 ..... *J. Jin, L. Li, H. Tian, Q. Wang, and X.-C. Gao* 176

Rank-Two Beamformed Secure Multicasting for Wireless Information and Power Transfer http://dx.doi.org/10.1109/LSP.2013.2296941 ..... *B. Zhu, J. Ge, Y. Huang, Y. Yang, and M. Lin* 199

Transceiver Design for Hybrid One-Way and Two-Way Relay Networks http://dx.doi.org/10.1109/LSP.2013.2296975 ..... *M. Wang, P. Wang, Y. Li, Z. Zhong, F. Wang, and B. Vucetic* 204

A Statistical-CSI-Based Scheme for Multiple Description Coding Multicast in CRNs http://dx.doi.org/10.1109/LSP.2013.2297338 ..... *K. Yang, W. Xu, S. Li, J. Lin, and W. Wu* 213

Alternating Rate Profile Optimization in Single Stream MIMO Interference Channels http://dx.doi.org/10.1109/LSP.2013.2297351 ..... *R. Mochaourab, P. Cao, and E. Jorswieck* 221

Speech Processing

Estimating Speaker Height and Subglottal Resonances Using MFCCs and GMMs http://dx.doi.org/10.1109/LSP.2013.2295397 ..... *H. Arsikere, S. M. Lulich, and A. Alwan* 159

Maximum Phase Modeling for Sparse Linear Prediction of Speech http://dx.doi.org/10.1109/LSP.2013.2296944 ..... *T. Drugman* 185

Statistical and Adaptive Signal Processing

Sigma Point Belief Propagation http://dx.doi.org/10.1109/LSP.2013.2290192 ..... *F. Meyer, O. Hlinka, and F. Hlawatsch* 145

Joint Bayesian Estimation of Close Subspaces from Noisy Measurements http://dx.doi.org/10.1109/LSP.2013.2296138 ..... *O. Besson, N. Dobigeon, and J.-Y. Tourneret* 168

Decision Fusion With Unknown Sensor Detection Probability http://dx.doi.org/10.1109/LSP.2013.2295054 ..... *D. Ciuonzo and P. Salvo Rossi* 208

Sparsity-Aware Adaptive Algorithms Based on Alternating Optimization and Shrinkage http://dx.doi.org/10.1109/LSP.2014.2298116 ..... *R. C. de Lamare and R. Sampaio-Neto* 225

About QLMS Derivations http://dx.doi.org/10.1109/LSP.2014.2299066 ..... *Q. Barthélemy, A. Larue, and J. I. Mars* 240

MLSP2014

IEEE International Workshop on

MACHINE LEARNING FOR SIGNAL PROCESSING

SEPTEMBER 21-24, 2014 REIMS, FRANCE

MLSP2014.CONWIZ.DK

Schedule

Special session call for proposal	March 5, 2014
Submission of full paper:	May 5, 2014
Notification of acceptance:	June 27, 2014
Camera-ready paper:	April 5, 2014
Advance registration before:	July 24, 2014
Camera-ready paper:	July 25, 2014

Organizing committee

General Chair	Mamadou Mboup. Univ. Reims France
Program Chairs	Tülay Adali, Univ. Maryland BC, USA Eric Moreau, Univ. of Toulon France
Data Competition Chair	Vince Calhoun, Univ. of New Mexico USA
Special Session Chair	Jean-Yves Tourneret, INP - ENSEEIHT Toulouse, France
Publicity Chair	Marc Van Hulle, KU Leuven, Belgium
Web and Publication Chairs	Jan Larsen. Tech. Univ. of Denmark, Denmark Kevin Guelton, Univ. of Reims, France
Local Arrangement Chairs	Valeriu Vrabie, Univ. of Reims, France Hassan Fenniri, Univ. of Reims, France

Sponsored by

IEEE Signal Processing Society

CALL FOR PAPERS

The 24th MLSP workshop in the series of workshops organized by the IEEE Signal Processing Society MLSP Technical Committee will present the most recent and exciting advances in machine learning for signal processing through keynote talks, tutorials as well as special and regular single-track sessions. Prospective authors are invited to submit papers on relevant algorithms and applications including, but not limited to:

- Learning theory and techniques
- Graphical models and kernel methods
- Data-driven adaptive systems and models
- Pattern recognition and classification
- Distributed, Bayesian, subspace/manifold and sparsity-aware learning
- Multiset data analysis and multimodal data fusion
- Perceptual signal processing in audio, image and video
- Cognitive information processing
- Multichannel adaptive and nonlinear signal processing
- Applications, including: speech and audio, image and video, music biomedical signals and images, communications, bioinformatics, biometrics, computational intelligence, genomic signals and sequences, social networks, games, smart grid

Data Analysis and Signal Processing Competition is being organized in conjunction with the workshop. The goal of the competition is to advance the current state-of-the-art in theoretical and practical aspects of signal processing domains. Winners will be announced and awards given at the workshop.

MLSP 2014 seeks proposals for **Special Sessions**, which are expected to address research in focused, emerging, or interdisciplinary areas of particular interest, not covered already by traditional MLSP sessions.

The **MLSP Best Student Paper Award** will be granted to the best overall paper for which a student is the principal author and presenter.

Prospective authors are invited to submit a double column paper of up to six pages using the electronic submission procedure at http://mlsp2014.conwiz.dk. Accepted papers will be published on memory sticks to be distributed at the workshop. The presented papers will be published in and indexed by IEEE Xplore.

IEEE
SignalProcessing
MAGAZINE
[VOLUME 31 NUMBER 2 MARCH 2014]
DSP AND CODING FOR MULTI-Tb/s OPTICAL TRANSPORT
RECENT ADVANCES
MELODY EXTRACTION FROM POLYPHONIC MUSIC SIGNALS
ON THE RECURSIVE SOLUTION TO THE NORMAL EQUATIONS
HAPTICS AND SIGNAL PROCESSING
THE HOPPING DISCRETE FOURIER TRANSFORM
IEEE Signal Processing Society
IEEE

CONTENTS

[VOLUME 31 NUMBER 2]

SPECIAL SECTION—DSP AND CODING FOR MULTI-Tb/s OPTICAL TRANSPORT

15 **FROM THE GUEST EDITORS**
Ivan B. Djordjevic, William Shieh, Xiang Liu, and Moshe Nazarathy

16 **DIGITAL SIGNAL PROCESSING TECHNIQUES ENABLING MULTI-Tb/s SUPERCHANNEL TRANSMISSION**
Xiang Liu, S. Chandrasekhar, and Peter J. Winzer

25 **MIMO SIGNAL PROCESSING FOR MODE-DIVISION MULTIPLEXING**
Sercan Ö. Arık, Joseph M. Kahn, and Keang-Po Ho

35 **EFFICIENT CLOCK AND CARRIER RECOVERY ALGORITHMS FOR SINGLE-CARRIER COHERENT OPTICAL SYSTEMS**
Xiang Zhou

46 **DIGITAL FIBER NONLINEARITY COMPENSATION**
Liang B. Du, Danish Rafique, Antonio Napoli, Bernhard Spinnler, Andrew D. Ellis, Maxim Kuschnerov, and Arthur J. Lowery

57 **CYCLE SLIP COMPENSATION WITH POLARIZATION BLOCK CODING FOR COHERENT OPTICAL TRANSMISSION**
Tsuyoshi Yoshida, Takashi Sugihara, Kazuyuki Ishida, and Takashi Mizuochi

70 **SUBBANDED DSP ARCHITECTURES BASED ON UNDERDECIMATED FILTER BANKS FOR COHERENT OFDM RECEIVERS**
Moshe Nazarathy and Alex Tolmachev

82 **ADVANCED DSP TECHNIQUES ENABLING HIGH SPECTRAL EFFICIENCY AND FLEXIBLE TRANSMISSIONS**
Alan Pak Tao Lau, Yuliang Gao, Qi Sui, Dawei Wang, Qunbi Zhuge, Mohamed H. Morsy-Osman, Mathieu Chagnon, Xian Xu, Chao Lu, and David V. Plant

93 **CODED MODULATION FOR FIBER-OPTIC NETWORKS**
Lotfollah Beygi, Erik Agrell, Joseph M. Kahn, and Magnus Karlsson

104 **MULTIDIMENSIONAL SIGNALING AND CODING ENABLING MULTI-Tb/s OPTICAL TRANSPORT AND NETWORKING**
Ivan B. Djordjevic, Milorad Cvijetic, and Changyu Lin

FEATURE

118 **MELODY EXTRACTION FROM POLYPHONIC MUSIC SIGNALS**
Justin Salamon, Emilia Gómez, Daniel P.W. Ellis, and Gaël Richard

COLUMNS

4 **FROM THE EDITOR**
Peer Reviews: What Would Be the Best Model?
Abdelhak Zoubir

6 **PRESIDENT'S MESSAGE**
The Only Constant Is Change
Alex Acero

11 **SPECIAL REPORTS**
Touching Research: Haptics and Signal Processing
John Edwards

135 **SP TIPS & TRICKS**
The Hopping Discrete Fourier Transform
Chun-Su Park and Sung-Jea Ko

140 **LECTURE NOTES**
On the Recursive Solution to the Normal Equations: Some Other Results
Manuel D. Ortigueira

DEPARTMENTS

8 **SOCIETY NEWS**
SPS Fellows and Award Winners Recognized

143 **DATES AHEAD**

Digital Object Identifier 10.1109/MSP.2014.2299114

IEEE ORDER FORM FOR REPRINTS

Purchasing IEEE Papers in Print is easy, cost-effective and quick.
Complete this form, tear it out, and either fax it (24 hours a day) to 732-981-8062 or mail it back to us.

PLEASE FILL OUT THE FOLLOWING

Author: _____

Publication Title: _____

Paper Title: _____

_____

RETURN THIS FORM TO:
IEEE Publishing Services
445 Hoes Lane
Box 1331
Piscataway, NJ 08855-1331
Call Reprint Department at (732) 562-3941 for questions regarding this form
(732) 981-8062 - FAX

PLEASE SEND ME

☐ 50 ☐ 100 ☐ 200 ☐ 300 ☐ 400 ☐ 500 or _____ (in multiples of 50) reprints.
☐ YES ☐ NO Self-covering/title page required. COVER PRICE: $74 per 100, $39 per 50.
☐ $58.00 Air Freight must be added for all orders being shipped outside the U.S.
☐ $18.50 must be added for all USA shipments to cover the cost of UPS shipping and handling.

PAYMENT

☐ Check enclosed. Payable on a bank in the USA.
☐ Charge my: ☐ Visa ☐ Mastercard ☐ Amex ☐ Diners Club

Account # _____ Exp. date _____

Cardholder's Name (please print): _____

_____

☐ Bill me (you must attach a purchase order) Purchase Order Number _____

Send Reprints to:

_____
_____
_____
_____

Bill to address, if different:

_____
_____
_____
_____

Because information and papers are gathered from various sources, there may be a delay in receiving your reprint request. This is especially true with postconference publications. Please provide us with contact information if you would like notification of a delay of more than 12 weeks.

Telephone: _____ Fax: _____ Email Address: _____

2011 REPRINT PRICES (without covers)

Number of Text Pages

	1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44	45-48
50	$126	$211	$243	$245	$285	$340	$371	$408	$440	$477	$510	$543
100	$242	$423	$476	$492	$570	$680	$742	$817	$885	$953	$1021	$1088

Larger quantities can be ordered. Email reprints@ieee.org with specific details.

Tax Applies on shipments of regular reprints to CA, DC, FL, MI, NJ, NY, OH and Canada (GST Registration no. 12534188).
Prices are based on black & white printing. Please call us for full color price quote, if applicable.

Authorized Signature: _____ Date: _____

2014 IEEE MEMBERSHIP APPLICATION

(students and graduate students must apply online)

IEEE Advancing Technology for Humanity

Start your membership immediately: Join online www.ieee.org/join

Please complete both sides of this form, typing or **printing in capital letters**. Use only English characters and abbreviate only if more than 40 characters and spaces per line. We regret that incomplete applications cannot be processed.

1 Name & Contact Information

Please PRINT your name as you want it to appear on your membership card and IEEE correspondence. As a key identifier for the IEEE database, circle your last/surname.

☐ Male ☐ Female Date of birth (Day/Month/Year) ____/____/____

Title First/Given Name Middle Last/Family Surname

▼ Primary Address ☐ Home ☐ Business (All IEEE mail sent here)

Street Address

City State/Province

Postal Code Country

Primary Phone

Primary E-mail

▼ Secondary Address ☐ Home ☐ Business

Company Name Department/Division

Street Address City State/Province

Postal Code Country

Secondary Phone

Secondary E-mail

To better serve our members and supplement member dues, your postal mailing address is made available to carefully selected organizations to provide you with information on technical services, continuing education, and conferences. Your e-mail address is not rented by IEEE. Please check box only if you do not want to receive these postal mailings to the selected address. ☐

2 Attestation

I have graduated from a three- to five-year academic program with a university-level degree.
☐ Yes ☐ No

This program is in one of the following fields of study:
☐ Engineering
☐ Computer Sciences and Information Technologies
☐ Physical Sciences
☐ Biological and Medical Sciences
☐ Mathematics
☐ Technical Communications, Education, Management, Law and Policy
☐ Other (please specify): _____

This academic institution or program is accredited in the country where the institution is located. ☐ Yes ☐ No ☐ Do not know

I have _____ years of professional experience in teaching, creating, developing, practicing, or managing within the following field:

☐ Engineering
☐ Computer Sciences and Information Technologies
☐ Physical Sciences
☐ Biological and Medical Sciences
☐ Mathematics
☐ Technical Communications, Education, Management, Law and Policy
☐ Other (please specify): _____

3 Please Tell Us About Yourself

Select the numbered option that best describes yourself. This information is used by IEEE magazines to verify their annual circulation. Please enter numbered selections in the boxes provided.

A. Primary line of business → ☐

1. Computers
2. Computer peripheral equipment
3. Software
4. Office and business machines
5. Test, measurement and instrumentation equipment
6. Communications systems and equipment
7. Navigation and guidance systems and equipment
8. Consumer electronics/appliances
9. Industrial equipment, controls and systems
10. ICs and microprocessors
11. Semiconductors, components, sub-assemblies, materials and supplies
12. Aircraft, missiles, space and ground support equipment
13. Oceanography and support equipment
14. Medical electronic equipment
15. OEM incorporating electronics in their end product (not elsewhere classified)
16. Independent and university research, test and design laboratories and consultants (not connected with a mfg. co.)
17. Government agencies and armed forces
18. Companies using and/or incorporating any electronic products in their manufacturing, processing, research or development activities
19. Telecommunications services, telephone (including cellular)
20. Broadcast services (TV, cable, radio)
21. Transportation services (airline, railroad, etc.)
22. Computer and communications and data processing services
23. Power production, generation, transmission and distribution
24. Other commercial users of electrical, electronic equipment and services (not elsewhere classified)
25. Distributor (reseller, wholesaler, retailer)
26. University, college/other educational institutions, libraries
27. Retired
28. Other_____

B. Principal job function → ☐

1. General and corporate management
2. Engineering management
3. Project engineering management
4. Research and development management
5. Design engineering management—analog
6. Design engineering management—digital
7. Research and development engineering
8. Design/development engineering—analog
9. Design/development engineering—digital
10. Hardware engineering
11. Software design/development
12. Computer science
13. Science/physics/mathematics
14. Engineering (not elsewhere specified)
15. Marketing/sales/purchasing
16. Consulting
17. Education/teaching
18. Retired
19. Other_____

C. Principal responsibility → ☐

1. Engineering and scientific management
2. Management other than engineering
3. Engineering design
4. Engineering
5. Software: science/mngmnt/engineering
6. Education/teaching
7. Consulting
8. Retired
9. Other_____

D. Title → ☐

1. Chairman of the Board/President/CEO
2. Owner/Partner
3. General Manager
4. VP Operations
5. VP Engineering/Dir. Engineering
6. Chief Engineer/Chief Scientist
7. Engineering Management
8. Scientific Management
9. Member of Technical Staff
10. Design Engineering Manager
11. Design Engineer
12. Hardware Engineer
13. Software Engineer
14. Computer Scientist
15. Dean/Professor/Instructor
16. Consultant
17. Retired
18. Other_____

Are you now or were you ever a member of IEEE?
☐ Yes ☐ No If yes, provide, if known:

Membership Number Grade Year Expired

4 Please Sign Your Application

I hereby apply for IEEE membership and agree to be governed by the IEEE Constitution, Bylaws, and Code of Ethics. I understand that IEEE will communicate with me regarding my individual membership and all related benefits. **Application must be signed.**

Signature Date

Over Please

Information for Authors

(Updated March 2012)

The IEEE TRANSACTIONS are published monthly covering advances in the theory and application of signal processing. The scope is reflected in the EDICS: the Editor's Information and Classification Scheme. Please consider the journal with the most appropriate scope for your submission.

Authors are encouraged to submit manuscripts of Regular papers (papers which provide a complete disclosure of a technical premise), or Correspondences (brief items that describe a use for or magnify the meaning of a single technical point, or provide comment on a paper previously published in the TRANSACTIONS). Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere.

Every manuscript must (a) provide a clearly defined statement of the problem being addressed, (b) state why it is important to solve the problem, and (c) give an indication as to how the current solution fits into the history of the problem.

By submission/resubmission of your manuscript to this TRANSACTIONS, you are acknowledging that you accept the rules established for publication of manuscripts, including agreement to pay all overlength page charges, color charges, and any other charges and fees associated with publication of the manuscript. Such charges are not negotiable and cannot be suspended.

New and revised manuscripts should be prepared following the "New Manuscript Submission" guidelines below, and submitted to the online manuscript system ScholarOne Manuscripts. After acceptance, finalized manuscripts should be prepared following the "Final Manuscript Submission Guidelines" below. Do not send original submissions or revisions directly to the Editor-in-Chief or Associate Editors; they will access your manuscript electronically via the ScholarOne Manuscripts system.

New Manuscript Submission. Please follow the next steps.

1. *Account in ScholarOne Manuscripts*. If necessary, create an account in the on-line submission system ScholarOne Manuscripts. Please check first if you already have an existing account which is based on your e-mail address and may have been created for you when you reviewed or authored a previous paper.
2. *Electronic Manuscript*. Prepare a PDF file containing your manuscript in double-spaced format (one full blank line between lines of type) using a font size of 11 points or larger, having a margin of at least 1 inch on all sides. For a regular paper, the manuscript may not exceed 30 double-spaced pages, including title; names of authors and their complete contact information; abstract; text; all images, figures and tables; and all references.

   Upload your manuscript as a PDF file "manuscript.pdf" to the ScholarOne Manuscripts site, then proofread your submission, confirming that all figures and equations are visible in your document before you "SUBMIT" your manuscript. Proofreading is critical; once you submit your manuscript, the manuscript cannot be changed in any way. You may also submit your manuscript as a PostScript or MS Word file. The system has the capability of converting your files to PDF, however it is your responsibility to confirm that the conversion is correct and there are no font or graphics issues prior to completing the submission process.
3. *Double-Column Version of Manuscript*. You are required to also submit a roughly formatted version of the manuscript in single-spaced, double column IEEE format (10 points for a regular submission or 9 points for a Correspondence) using the IEEE style files (it is allowed to let long equations stick out). *If accepted for publication,* over length page charges are levied beginning with the 11th published page of your manuscript. You are, therefore, advised to be conservative in your submission. This double-column version submitted will serve as a confirmation of the approximate publication length of the manuscript and gives an additional confirmation of your understanding that over length page charges will be paid when billed upon publication.

   Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOneManuscripts site.
4. *Additional Material for Review*. Please upload pdf versions of all items in the reference list which are not publicly available, such as unpublished (submitted) papers. Other materials for review such as supplementary tables and figures, audio fragments and QuickTime movies may be uploaded as well. Reviewers will be able to view these files only if they have the appropriate software on their computers. Use short filenames without spaces or special characters. When the upload of each file is completed, you will be asked to provide a description of that file.
5. *Submission*. After uploading all files and proofreading them, submit your manuscript by clicking "Submit." A confirmation of the successful submission will open on screen containing the manuscript tracking number and will be followed with an e-mail confirmation to the corresponding and all contributing authors. Once you click "Submit," your manuscript cannot be changed in any way.
6. *Copyright Form and Consent Form*. By policy, IEEE owns the copyright to the technical contributions it publishes on behalf of the interests of the IEEE, its authors, and their employers; and to facilitate the appropriate reuse of this material by others. To comply with the IEEE copyright policies, authors are required to sign and submit a completed "IEEE Copyright and Consent Form" prior to publication by the IEEE.

   The IEEE recommends authors to use an effective electronic copyright form (eCF) tool within the ScholarOne Manuscripts system. You will be redirected to the "IEEE Electronic Copyright Form" wizard at the end of your original submission; please simply sign the eCF by typing your name at the proper location and click on the "Submit" button.

Correspondence Items. Correspondence items are short disclosures with a reduced scope or significance that typically describe a use for or magnify the meaning of a single technical point, or provide brief comments on material previously published in the TRANSACTIONS. These items may not exceed 12 pages in double-spaced format (3 pages for Comments), using 11 point type, with margins of 1 inch minimum on all sides, and including: title, names and contact information for authors, abstract, text, references, and an appropriate number of illustrations and/or tables. Correspondence items are submitted in the same way as regular manuscripts (see "New Manuscript Submission" above for instructions).

Manuscript Length. Papers published on or after 1 January 2007 can now be up to 10 pages, and any paper in excess of 10 pages will be subject to over length page charges. The IEEE Signal Processing Society has determined that the standard manuscript length shall be no more than 10 published pages (double-column format, 10 point type) for a regular submission, or 6 published pages (9 point type) for a Correspondence item, respectively. Manuscripts that exceed these limits will incur mandatory over length page charges, as discussed below. Since changes recommended as a result of peer review may require additions to the manuscript, it is strongly recommended that you practice economy in preparing original submissions.

Exceptions to the 30-page (regular paper) or 12-page (Correspondences) manuscript length may, under extraordinary circumstances, be granted by the Editor-in-Chief. However, such exception does not obviate your requirement to pay any and all over length or additional charges that attach to the manuscript.

Resubmission of Previously Rejected Manuscripts. Authors of rejected manuscripts are allowed to resubmit their manuscripts only once. The Signal Processing Society strongly discourages resubmission of rejected manuscripts more than once. At the time of submission, you will be asked whether you consider your manuscript to be a new submission or a resubmission of an earlier rejected manuscript. If you choose to submit a new version of your manuscript, you will be asked to submit supporting documents detailing how your new version addresses all of the reviewers' comments.

Full details of the resubmission process can be found in the Signal Processing Society "Policy and Procedures Manual" at http://www.signalprocessingsociety.org/about/governance/policy-procedure/. Also, please refer to the decision letter and your Author Center on the on-line submission system.

Author Misconduct.

Author Misconduct Policy: Plagiarism includes copying someone else's work without appropriate credit, using someone else's work without clear delineation of citation, and the uncited reuse of an authors previously published work that also involves other authors. Plagiarism is unacceptable.

Self-plagiarism involves the verbatim copying or reuse of an authors own prior work without appropriate citation; it is also unacceptable. Self-plagiarism includes duplicate submission of a single journal manuscript to two different journals, and submission of two different journal manuscripts which overlap substantially in language or technical contribution.

Authors may only submit original work that has not appeared elsewhere in a journal publication, nor is under review for another journal publication. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper. If authors have used their own previously published work as a basis for a new submission, they are required to cite the previous work and very briefly indicate how the new submission offers substantively novel contributions beyond those of the previously published work.

It is acceptable for conference papers to be used as the basis for a more fully developed journal submission. Still, authors are required to cite related prior work; the papers cannot be identical; and the journal publication must include novel aspects.

Author Misconduct Procedures: The procedures that will be used by the Signal Processing Society in the investigation of author misconduct allegations are described in the IEEE SPS Policies and Procedures Manual.

Author Misconduct Sanctions: The IEEE Signal Processing Society will apply the following sanctions in any case of plagiarism, or in cases of self-plagiarism that involve an overlap of more than 25% with another journal manuscript:

1) immediate rejection of the manuscript in question;
2) immediate withdrawal of all other submitted manuscripts by any of the authors, submitted to any of the Society's publications (journals, conferences, workshops), except for manuscripts that also involve innocent co-authors; immediate withdrawal of all other submitted manuscripts by any of the authors, submitted to any of the Society's publications (journals, conferences, workshops), except for manuscripts that also involve innocent co-authors;
3) prohibition against each of the authors for any new submissions, either individually, in combination with the authors of the plagiarizing manuscript, or in combination with new co-authors, to all of the Society's publications (journals, conferences, workshops). The prohibition shall continue for one year from notice of suspension.

Further, plagiarism and self-plagiarism may also be actionable by the IEEE under the rules of Member Conduct.

Submission Format.

Authors are encouraged to prepare manuscripts employing the on-line style files developed by IEEE. All manuscripts accepted for publication will require the authors to make final submission employing these style files. The style files are available on the web at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2 under "Template for all Transactions." (LaTeX and MS Word).

Authors using LaTeX: the two PDF versions of the manuscript needed for submission can both be produced by the IEEEtran.cls style file. A double-spaced document is generated by including \documentclass[11pt,draftcls,onecolumn]{IEEEtran} as the first line of the manuscript source file, and a single-spaced double-column document for estimating the publication page charges via \documentclass[10pt,twocolumn,twoside]{IEEEtran} for a regular submission, or \documentclass[9pt,twocolumn,twoside]{IEEEtran} for a Correspondence item.

- *Title page and abstract:* The first page of the manuscript shall contain the title, names and contact information for all authors (full mailing address, institutional affiliations, phone, fax, and e-mail), the abstract, and the EDICS. An asterisk * should be placed next to the name of the Corresponding Author who will serve as the main point of contact for the manuscript during the review and publication processes.

  An abstract should have not more than 200 words for a regular paper, or 50 words for a Correspondence item. The abstract should indicate the scope of the paper or Correspondence, and summarize the author's conclusions. This will make the abstract, by itself, a useful tool for information retrieval.
- *EDICS:* All submissions must be classified by the author with an EDICS (Editors' Information Classification Scheme) selected from the list of EDICS published online at http://www.signalprocessingsociety.org/publications/periodicals/tsp/TSP-EDICS/
- NOTE: EDICS are necessary to begin the peer review process. Upon submission of a new manuscript, please choose the EDICS categories that best suit your manuscript. Failure to do so will likely result in a delay of the peer review process.
- The EDICS category should appear on the first page—i.e., the title and abstract page—of the manuscript.
- *Illustrations and tables:* Each figure and table should have a caption that is intelligible without requiring reference to the text. Illustrations/tables may be worked into the text of a newly-submitted manuscript, or placed at the end of the manuscript. (However, for the final submission, illustrations/tables must be submitted separately and not interwoven with the text.)

  Illustrations in color may be used but, unless the final publishing will be in color, the author is responsible that the corresponding grayscale figure is understandable.

  In preparing your illustrations, note that in the printing process, most illustrations are reduced to single-column width to conserve space. This may result in as much as a 4:1 reduction from the original. Therefore, make sure that all words are in a type size that will reduce to a minimum of 9 points or 3/16 inch high in the printed version. Only the major grid lines on graphs should be indicated.
- *Abbreviations:* This TRANSACTIONS follows the practices of the IEEE on units and abbreviations, as outlined in the Institute's published standards. See http://www.ieee.org/portal/cms_docs_iportals/iportals/publications/authors/transjnl/auinfo07.pdf for details.
- *Mathematics:* All mathematical expressions must be legible. Do not give derivations that are easily found in the literature; merely cite the reference.

Final Manuscript Submission Guidelines.

Upon formal acceptance of a manuscript for publication, instructions for providing the final materials required for publication will be sent to the Corresponding Author. Finalized manuscripts should be prepared in LaTeX or MS Word, and are required to use the style files established by IEEE, available at http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2.

Instructions for preparing files for electronic submission are as follows:

- Files must be self-contained; that is, there can be no pointers to your system setup.
- Include a header to identify the name of the TRANSACTIONS, the name of the author, and the software used to format the manuscript.
- Do not import graphics files into the text file of your finalized manuscript (although this is acceptable for your initial submission). If submitting on disk, use a separate disk for graphics files.
- Do not create special macros.
- Do not send PostScript files of the text.
- File names should be lower case.
- Graphics files should be separate from the text, and not contain the caption text, but include callouts like "(a)," "(b)."
- Graphics file names should be lower case and named fig1.eps, fig2.tif, etc.
- Supported graphics types are EPS, PS, TIFF, or graphics created using Word, Powerpoint, Excel or PDF. Not acceptable is GIF, JPEG, WMF, PNG, BMP or any other format (JPEG is accepted for author photographs only). The provided resolution needs to be at least 600 dpi (400 dpi for color).
- Please indicate explicitly if certain illustrations should be printed in color; note that this will be at the expense of the author. Without other indications, color graphics will appear in color in the online version, but will be converted to grayscale in the print version.

Additional instructions for preparing, verifying the quality, and submitting graphics are available via http://www.ieee.org/publications_standards/publications/ authors/authors_journals.html.

Multimedia Materials.

IEEE Xplore can publish multimedia files and Matlab code along with your paper. Alternatively, you can provide the links to such files in a README file that appears on Xplore along with your paper. For details, please see http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect6 under "Multimedia." To make your work reproducible by others, the TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper.

Page Charges.

Voluntary Page Charges. Upon acceptance of a manuscript for publication, the author(s) or his/her/their company or institution will be asked to pay a charge of $110 per page to cover part of the cost of publication of the first ten pages that comprise the standard length (six pages, in the case of Correspondences).

Mandatory Page Charges. The author(s) or his/her/their company or institution will be billed $220 per each page in excess of the first ten published pages for regular papers and six published pages for correspondence items. These are mandatory page charges and the author(s) will be held responsible for them. They are not negotiable or voluntary. The author(s) signifies his willingness to pay these charges simply by submitting his/her/their manuscript to the TRANSACTIONS. The Publisher holds the right to withhold publication under any circumstance, as well as publication of the current or future submissions of authors who have outstanding mandatory page charge debt.

Color Charges. Color figures which appear in color only in the electronic (Xplore) version can be used free of charge. In this case, the figure will be printed in the hardcopy version in grayscale, and the author is responsible that the corresponding grayscale figure is intelligible. Color reproduction in print is expensive, and all charges for color are the responsibility of the author. The estimated costs are as follows. There will be a charge of $62.50 for each figure; this charge may be subject to change without notification. In addition, there are printing preparation charges which may be estimated as follows: color reproductions on four or fewer pages of the manuscript: a total of approximately $1045; color reproductions on five pages through eight pages: a total of approximately $2090; color reproductions on nine through 12 pages: a total of approximately $3135, and so on. Payment of fees on color reproduction is not negotiable or voluntary, and the author's agreement to publish the manuscript in the TRANSACTIONS is considered acceptance of this requirement.

To find the Information for Authors for IEEE Signal Processing Letters, the IEEE Journal of Selected Topics in Signal Processing or the IEEE Signal Processing Magazine, please refer to the IEEE Signal Processing website at www.signalprocessingsociety.org.

2014 IEEE SIGNAL PROCESSING SOCIETY MEMBERSHIP APPLICATION

(Current and reinstating IEEE members joining SPS complete areas 1, 2, 8, 9.)

Mail to: **IEEE OPERATIONS CENTER, ATTN: Louis Curcio, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA**
or Fax to (732) 981-0225 (credit card payments only.)
For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org

IEEE

1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: SEND MAIL TO: ☐ Home Address OR ☐ Business/School Address
If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE | FIRST OR GIVEN NAME | MIDDLE NAME | SURNAME/LAST NAME

HOME ADDRESS

CITY | STATE/PROVINCE | POSTAL CODE | COUNTRY

2.

Are you now or were you ever a member of IEEE? ☐ Yes ☐ No
If yes, please provide, if known:

MEMBERSHIP NUMBER_____

Grade_____ **Year Membership Expired:**_____

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name

Department/Division

Title/Position | Years in Current Position

Years in the Profession Since Graduation | ☐ PE State/Province

Street Address

City | State/Province | Postal Code | Country

4. EDUCATION

A baccalaureate degree from an IEEE recognized educational program assures assignment of "Member" grade. For others, additional information and references may be necessary for grade assignment.

A.

Baccalaureate Degree Received | Program/Course of Study

College/University | Campus

State/Province | Country | Mo./Yr. Degree Received

B.

Highest Technical Degree Received | Program/Course of Study

College/University | Campus

State/Province | Country | Mo./Yr. Degree Received

5.

Full signature of applicant

6. DEMOGRAPHIC INFORMATION – ALL APPLICANTS -

Date Of Birth _____ (Day Month Year) ☐ Male ☐ Female

7. CONTACT INFORMATION

Office Phone/Office Fax | Home Phone/Home Fax

Office E-Mail | Home E-Mail

8. 2014 IEEE MEMBER RATES

IEEE DUES Residence	16 Aug 12-28 Feb 14 Pay Full Year	1 Mar 13-15 Aug 14 Pay Half Year**
United States	$187.00 ☐	$93.50 ☐
Canada (incl. GST)	$168.10 ☐	$84.05 ☐
Canada (incl. HST for NB, NF and ON)	$179.46 ☐	$89.73 ☐
Canada (incl. HST for Nova Scotia)	$182.30 ☐	$91.15 ☐
Canada (incl. HST for PEI)	$180.88 ☐	$90.44 ☐
Africa, Europe, Middle East	$155.00 ☐	$77.50 ☐
Latin America	$146.00 ☐	$73.00 ☐
Asia, Pacific	$147.00 ☐	$73.50 ☐

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 13%, 14% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)

Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)

U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to **Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin.** Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.

IEEE Signal Processing Society

2014 SPS MEMBER RATES

		16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*		$ 20.00 ☐	$ 10.00 ☐
Fee includes: **IEEE Signal Processing Magazine** (electronic and digital), **Inside Signal Processing eNewsletter** (electronic) and **IEEE Signal Processing Society Content Gazette** (electronic).			
Add $15 to enhance SPS Membership and also receive:		$ 15.00 ☐	$ 7.50 ☐
IEEE Signal Processing Magazine (print) and **SPS Digital Library**: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.			
Publications available only with SPS membership:			
Signal Processing, IEEE Transactions on:	Print	$190.00 ☐	$ 95.00 ☐
Audio, Speech, and Language Proc., IEEE/ACM Trans. on:	Print	$145.00 ☐	$ 72.50 ☐
Image Processing, IEEE Transactions on:	Print	$188.00 ☐	$ 94.00 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$163.00 ☐	$ 81.50 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$160.00 ☐	$ 80.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 33.00 ☐	$ 16.50 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 55.00 ☐	$ 27.50 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$ 65.00 ☐	$ 32.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
Computing in Science & Engineering:	Electonic and Digital	$ 45.00 ☐	$ 22.50 ☐
	Print & Electronic	$ 49.00 ☐	$ 24.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 74.00 ☐	$ 37.00 ☐
	Electronic	$ 53.00 ☐	$ 26.50 ☐
	Print & Electronic	$ 89.00 ☐	$ 44.50 ☐
Mobile Computing, IEEE Trans. on:	ELE/Print Abstract/CD-ROM	$ 36.00 ☐	$ 18.00 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 42.00 ☐	$ 21.00 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 45.00 ☐	$ 22.50 ☐
	Print & Electronic	$ 49.00 ☐	$ 24.50 ☐
Network Science and Engrg, IEEE Trans. on:	Electronic	$ 32.00 ☐	$ 16.00 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 25.00 ☐	$ 12.50 ☐
	Electronic	$ 25.00 ☐	$ 12.50 ☐
	Print & Electronic	$ 40.00 ☐	$ 20.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 45.00 ☐	$ 22.50 ☐
	Print and Electronic	$ 49.00 ☐	$ 24.50 ☐
IEEE Biometrics Compendium	Online	$ 30.00 ☐	$ 15.00 ☐
IEEE Sensors Journal:	Print	$150.00 ☐	$ 75.00 ☐
	Electronic	$ 50.00 ☐	$ 25.00 ☐
	Print & Electronic	$190.00 ☐	$ 95.00 ☐
Smart Grid, IEEE Transactions on:	Print	$100.00 ☐	$ 50.00 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$120.00 ☐	$ 60.00 ☐
IEEE Transactions on Engineering Management:	Print & Electronic	$ 39.00 ☐	$ 19.50 ☐
IEEE Engineering Management Review:	Print & Electronic	$ 29.00 ☐	$ 14.50 ☐
IEEE Technology Management Package:	Print & Electronic	$ 60.00 ☐	$ 30.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$ 87.00 ☐	$ 43.50 ☐
	Electronic	$ 42.00 ☐	$ 21.00 ☐
	Print & Electronic	$ 87.00 ☐	$ 43.50 ☐
IEEE Wireless Communications Letters:	Print	$ 80.00 ☐	$ 40.00 ☐
	Electronic	$ 18.00 ☐	$ 9.00 ☐
	Print & Electronic	$ 95.00 ☐	$ 47.50 ☐

*IEEE membership required or requested
Affiliate application to join SP Society only.

Amount Paid $_____

9.

IEEE Membership Dues (See pricing in Section 8) $_____
Signal Processing Society Fees $_____

Canadian residents pay 5% GST or 13% HST
Reg. No. 125634188 on Society payment(s) & pubs only Tax $_____

AMOUNT PAID WITH APPLICATION TOTAL $_____

Prices subject to change without notice.

☐ **Check or money order enclosed Payable to IEEE on a U.S. Bank**
☐ **American Express** ☐ **VISA** ☐ **MasterCard**
☐ **Diners Club**

Exp. Date Mo./Yr.

Cardholder 5 Digit Zip Code Billing Statement Address/USA Only

Full signature of applicant using credit card Date

10. WERE YOU REFERRED?

☐ Yes ☐ No If yes, please provide the follow information:
Member Recruiter Name:_____
IEEE Recruiter's Member Number (Required):_____

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®