

IEEE Signal Processing Magazine

Volume 33 | Number 3 | May 2016



BRAIN ANALYTICS

Enhancing Understanding of Brain Function

The Meaning of the Signal Processing Life

Signal Processing Helps Put Robot Users in Control

Filter Design Tips & Tricks



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
—and get your FIRST MONTH FREE!**

www.ieee.org/go/freemonth



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

Contents

Volume 33 | Number 3 | May 2016

FEATURES

Brain Signal Analytics

12 FROM THE EDITORS

Z. Jane Wang, Wade Trappe,
and Shuguang Cui

14 THE CONNECTED BRAIN

Adeel Razi and Karl J. Friston

36 MAPPING BRAIN ANATOMICAL CONNECTIVITY USING DIFFUSION MAGNETIC RESONANCE IMAGING

Junning Li, Yonggang Shi,
and Arthur W. Toga

52 TIME-VARYING BRAIN CONNECTIVITY IN fMRI DATA

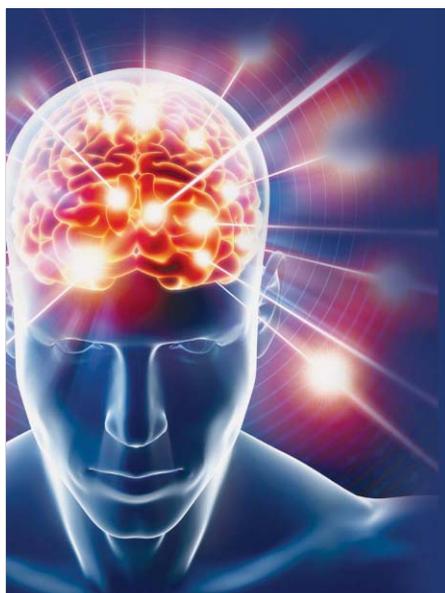
Vince D. Calhoun
and Tülay Adalı

67 THE PROMISE OF HYBRID PET/MRI

Habib Zaidi and
Minerva Becker

86 JOINT BLIND SOURCE SEPARATION FOR NEUROPHYSIOLOGICAL DATA ANALYSIS

Xun Chen, Z. Jane Wang,
and Martin J. McKeown



ON THE COVER

A complementary set of tutorial overview and survey articles demonstrating the importance of incorporating signal processing strategies into the advances in neuroimaging techniques, data analytics, and modeling for brain function is presented in this issue of *IEEE Signal Processing Magazine*. This cluster of feature articles showcases the inherently interdisciplinary nature of brain mapping research and the intriguing signal processing questions.

COVER IMAGE: ©STOCKPHOTO.COM/HENRIK5000

COLUMNS

6 Reader's Choice

Top Downloads in IEEE Xplore

8 Special Reports

Signal Processing Helps Put Robot
Users in Control
John Edwards

108 Tips & Tricks

Improving FIR Filters by Using
Cascade Techniques
*David Shiung, Ya-Yin Yang,
and Chu-Sing Yang*

Narrowband Notch Filter Using
Feedback Structure
*Soo-Chang Pei, Bo-Yi Guo,
and Wen-Yang Lu*

PG. 8



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\$213.00 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 © 2016 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.2016.2523299

DEPARTMENTS

IEEE Signal Processing Magazine

3 From the Editor

Silk Road in the New Millennium
Min Wu

4 President's Message

The Meaning of (the Signal Processing) Life
Rabab Ward

120 Dates Ahead



PG. 120

Check out the calendar in this issue for upcoming conferences!



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

EDITOR-IN-CHIEF

Min Wu—University of Maryland, College Park
U.S.A.

AREA EDITORS

Feature Articles

Shuguang Robert Cui—Texas A&M University,
U.S.A.

Special Issues

Wade Trappe—Rutgers University, U.S.A.

Columns and Forum

Gwenaél Doërr—ContentArmor SAS, France
Kenneth Lam—Hong Kong Polytechnic University,
Hong Kong SAR of China

e-Newsletter

Christian Debes—TU Darmstadt and
AGT International, Germany

Social Media and Outreach

Andres Kwasinski—Rochester Institute
of Technology, U.S.A.

EDITORIAL BOARD

Mrityunjoy Chakraborty—Indian Institute of
Technology, Kharagpur, India

George Christos—Qualcomm, Inc.,
U.S.A.

Patrick Flandrin—ENS Lyon, France

Mounir Ghogho—University of Leeds,
U. K.

Lina Karam—Arizona State University, U.S.A.

Hamid Krim—North Carolina State University,
U.S.A.

Sven Lončarić—University of Zagreb, Croatia

Brian Lovell—University of Queensland, Australia

Jian Lu—Qihoo 360, China

Henrique (Rico) Malvar—Microsoft Research,
U.S.A.

Yi Ma—ShanghaiTech University, China

Stephen McLaughlin—Heriot-Watt University,
Scotland

Athina Petropulu—Rutgers University,
U.S.A.

Peter Ramadge—Princeton University,
U.S.A.

Shigeki Sagayama—Meiji University, Japan

Erchin Serpedin—Texas A&M University,
U.S.A.

Shihab Shamma—University of Maryland,
U.S.A.

Hing Cheung So—City University of Hong Kong,
Hong Kong

Isabel Trancoso—INESC-ID/Instituto Superior
Técnico, Portugal

Pramod K. Varshney—Syracuse University,
U.S.A.

Z. Jane Wang—The University of British
Columbia, Canada

Gregory Wornell—Massachusetts Institute
of Technology, U.S.A.

Dapeng Wu—University of Florida, U.S.A.

ASSOCIATE EDITORS—COLUMNS AND FORUM

Ivan Bajic—Simon Fraser University, Canada

Rodrigo Capobianco Guido—
São Paulo State University, Brazil

Ching-Te Chiu—National Tsing Hua University,
Taiwan

Michael Gormish—Ricoh Innovations, Inc.

Xiaodong He—Microsoft Research

Danilo Mandic—Imperial College,
U. K.

Aleksandra Mojsilović—

IBM T.J. Watson Research Center

Douglas O'Shaughnessy—INRS, Canada

Fatih Porikli—MERL

Shantanu Rane—PARC, U.S.A.

Saeid Sanei—University of Surrey, U. K.

Roberto Togneri—The University of
Western Australia

Alessandro Vinciarelli—IDIAP-EPFL

Azadeh Vosoughi—University of Central Florida

Stefan Winkler—UIUC/ADSC, Singapore

ASSOCIATE EDITORS—e-NEWSLETTER

Csaba Benedek—Hungarian Academy
of Sciences, Hungary

Paolo Braca—NATO Science and Technology
Organization, Italy

Quan Ding—University of California,
San Francisco, U.S.A.

Pierluigi Failla—Compass Inc, New York,
U.S.A.

Marco Guerriero—General Electric Research,
U.S.A.

Yang Li—Harbin Institute of Technology, China

Yuhong Liu—Penn State University at Altoona,
U.S.A.

Andreas Merentitis—University of Athens, Greece

Michael Muma—TU Darmstadt, Germany

Xiaorong Zhang—San Francisco State University,
U.S.A.

ASSOCIATE EDITOR—SOCIAL MEDIA/OUTREACH

Guijin Wang—Tsinghua University, China

IEEE SIGNAL PROCESSING SOCIETY

Rabab Ward—President

Ali Sayed—President-Elect

Carlo S. Regazzoni—Vice President, Conferences

Konstantinos (Kostas) N. Plataniotis—

Vice President, Membership

Thrasivoulos (Thrasos) N. Pappas—

Vice President, Publications

Walter Kellerman—Vice President,

Technical Directions

IEEE SIGNAL PROCESSING SOCIETY STAFF

Denise Hurley—Senior Manager of Conferences
and Publications

Rebecca Wollman—Publications Administrator

IEEE PERIODICALS MAGAZINES DEPARTMENT

Jessica Barragué, *Managing Editor*

Geraldine Krolin-Taylor, *Senior Managing Editor*

Mark David, *Senior Manager*

Advertising and Business Development

Felicia Spagnoli, *Advertising Production Manager*

Janet Dudar, *Senior Art Director*

Gail A. Schnitzer, Mark Morrissey,

Associate Art Directors

Theresa L. Smith, *Production Coordinator*

Dawn M. Melley, *Editorial Director*

Peter M. Tuohy, *Production Director*

Fran Zappulla, *Staff Director,*

Publishing Operations

Digital Object Identifier 10.1109/MSP.2016.2541418

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.

FROM THE EDITOR

Min Wu | Editor-in-Chief | minwu@umd.edu

Silk Road in the New Millennium

I am writing this editorial just days before heading to Shanghai, China, to attend ICASSP 2016. This is the first time that the IEEE Signal Processing Society (SPS) has held its flagship conference, the 41st in the series, in Mainland China. ICIP will follow the footsteps of ICASSP and be held in Beijing, the capital city of China, in September 2017. And, for the first time, ICASSP will be held in South Korea, in 2018. In a sense, major SPS conferences are paving the way for the SPS to develop a strong presence in Asia and encourage engagement between communities in the East and the West, playing a role of what the Silk Road once achieved.

For colleagues in North America and Europe, traveling to conferences in Asia often takes nearly a whole day. When we frown about the long travels for these major conferences increasingly being held in Asia, have we thought about when the tables were turned? Colleagues from Asia (as well as such regions as Australia, New Zealand, and South America) have had to travel that far for the vast majority of past conferences that were held in North America or Europe!

Growing up in China, Asia is undoubtedly special to me. Emotional attachment aside, we have seen a rapid growth of SPS membership and our magazine's readership in Asia, now accounting for nearly 30% of total SPS members. I want to share with you what I recently learned from two efforts related

to Asia, as they brought enlightening thoughts about the spectrum of future SPS activities and magazine contents.

The first event in my ICASSP trip is to organize the final competition of this year's IEEE Signal Processing Cup (SP Cup) global competitions. The SP Cup provides undergraduate students with an opportunity to form teams and work together to solve a challenging and interesting real-world problem using signal processing methods. SP Cup 2016 has seen participations from 28 countries, covering every habitable continent. For three years in a row, we have seen enthusiastic involvement of undergraduate students from Asia. Not only are more than half of the team submissions in 2016 from Asia, but at least two out of three finalists in each SP Cup so far came from Asia.

The second effort I'd like to share addresses a traditional conference model that has focused on publishing the latest original research work. With its selective nature on paper acceptance as well as a nontrivial registration fee, attendees would be limited to primarily researchers and doctoral students in well-funded research institutions. But last summer, I learned of a grassroots effort by an enthusiastic group of mid-career Chinese colleagues. They initiated an annual gathering of Vision and Learning Seminars (VALSE), which aims to help researchers, practitioners, and graduate students in China acquire the latest knowledge and elevate the overall technical levels. To lower the engagement barrier, VALSE is free to attend and uses local language in oral communications; to reduce overhead

and seek to complement, instead of compete for papers with highly established conferences, VALSE does not have submission of new papers, and instead the focus is on invited talks by active researchers and rising stars and highlights of recently accepted papers in major journals and conferences. Within just six years, VALSE has energized the computer vision community in China and attracted more than 1,000 participants and numerous industry sponsors. It also utilizes social and online platforms to engage people at different career stages and is extending its activities to provide webinars and learning resources throughout the year.

Seeing the encouraging responses toward these activity models, I wonder: can a new model of community learning in a geographic region address what's missing between traditional research conferences and one-time local Chapter events? Will the SP Cup complement the traditional classroom learning and connect signal processing students around the world to develop a global community? *IEEE Signal Processing Magazine* has technical readership spanning broad backgrounds and needs, so what can we learn from the enthusiastic responses of undergraduate and professionals in Asia toward these events? Perhaps insights on these questions will shine a light on this modern-times Silk Road between the East and the West for the signal processing community.

SP

PRESIDENT'S MESSAGE

Rabab Ward | SPS President | rababw@ece.ubc.ca

The Meaning of (the Signal Processing) Life

When I told a friend of mine that I'd soon be starting my term as president of the IEEE Signal Processing Society (SPS), she gave me a look that's probably familiar to many of you—"What is *that*, and why should I care?" Two great questions from the uninitiated that underscored the importance of our efforts these past two years in raising awareness about signal processing (SP) among the general public. To help identify new methods for increasing our visibility over the next two years, I took a few steps back and had a fresh look at the definition and purpose of our chosen field, by reading the "Scope/Mission" section of our webpage.

After defining what SP is, the mission statement addresses the "Why should I care" question: "... signal processing is a core technology for addressing critical societal challenges that include healthcare, energy systems, sustainability, transportation, entertainment, education, communication, collaboration, defense, and security." In simple terms, SP impacts all facets of life, and it's uniquely positioned to solve a wide array of "social challenges" by "enabling technology for the generation, transformation, and interpretation of information" in a broad variety of disciplines. Those are some very important reasons why my friends and society at large should care about SPS.

Digital Object Identifier 10.1109/MSP.2016.2525498
Date of publication: 27 April 2016

How does the SPS go about reaching that goal? Our mission statement asserts that we will "advance and disseminate state-of-the-art scientific information and resources; educate the signal processing community; and provide a venue for people to interact and exchange ideas." I'd like to share some of my own strategies for fulfilling our mission statement and growing the scope of our Society during my two-year tenure. Your input is critical and most welcome. Meanwhile, I've identified five key objectives.

First, I'd like to engage a broader pool of members. The SPS is built on the efforts of hundreds of dedicated volunteers. Yet, as our membership numbers grow, the needs of our community expands, and the reach of our disciplines evolves, the importance of engaging committed volunteers increases exponentially. We need volunteers to make our initiatives happen, and we need to continuously define and refine our next steps. Given our international nature, we must rely on our conferences and publications but more so on our networking channels, specifically our regional Chapters. To engage all of our members from around the globe, we must pay special attention to growing and nurturing our worldwide

membership and its roots among the local Chapters.

The voices of our young members are particularly important for the health and longevity of our Society. They are savvy to social media and are uniquely positioned to engage with other young colleagues and spearhead innovative campaigns that will tap into a wider demographic and a global audience. One idea to drum up interest among

young signal processors could be a competition for members to produce a 1-min video about what the SPS should look like in ten years.

Our young members play a key role in helping us reach another important

goal: raising SP awareness to attract young innovative minds to the SPS. This will also help secure our Society's future and raise our profile with a demographic that is uniquely attracted to idealistic social causes—a perfect fit for the goal of solving social challenges. Outreach to students and young professionals is critical, and mentorships can benefit both established and fledgling practitioners. It's our responsibility to ensure that young minds around the world understand what SP is about. It's our responsibility to provide all members—especially young people—the insights and tools so that they can

To engage all of our members from around the globe, we must pay special attention to growing and nurturing our worldwide membership and its roots among the local Chapters.

learn more about SP. Once we pique interest, we must support and empower these young people to carve a professional path and make their own contributions to benefit the field and society at large. This underscores the necessity that our Society provides continuing education to our SPS members—a new effort that SPS is now taking serious steps to address.

Another goal is to enhance the SPS's value for current and potential members. I'd like to go above and beyond our current services, which include our highly respected conferences and workshops, our newsletter, and our highly acclaimed magazine and publications. We should also include the needs of potential members in our planning. Our SP community is much larger than twice the number of our members, encompassing nonmembers who attend our conferences, follow our publications, author articles in our journals, etc. Identifying these nonmembers and assessing and meeting their needs should be a future priority.

The SPS has already started several great initiatives to provide services to members and to the SP community at large. SigPort and SigView are two such great initiatives. SigPort allows researchers to post and archive their manuscripts, reports, theses, and supporting materials. We must assess the utility of this system and find out from members how to make it more valuable to all of our authors. For example, how can we kick-start a lively comments section alongside our members' published papers? Can it be more interactive? Is "crowd reviewing" possible? Is it useful? SigView is already attracting the interest of graduate students and practitioners. We need to hear from users regarding ways that we can build on its current success to further serve their needs and expand the circle of benefits. We've launched such initiatives, so let's monitor and assess them on a continual basis, fine-tuning them as necessary to ensure their value, success, and user friendliness.

We already have a few platforms in place that could stimulate interest

among young people. Summer schools are one such venue. Another medium is our very successful outreach: *IEEE Signal Processing Magazine*. Should we use it to reach a wider audience, beyond our communities, as *IEEE Spectrum* has already done? Is the use of social media channels and our website the most effective way to interact with our members and to reach a broader audience? The Student Cup is proving to be a very

powerful tool to engage young students in SP. What are the best approaches to define challenges for the Cup that would attract even a much wider pool of participants and inspire them to find innovative solutions? What would it take for the Student Cup to become the premier "filter" to identify SP future leaders? Can students help educate others, perhaps by creating and publishing "SP for Dummies"-style videos?

Another priority is to create a culture of meaningful partnership between academic and industrial SP communities. It's an ideal way for us to reach our stated goal of harnessing SP as an enabling technology to address society's critical challenges. We have a much better opportunity to meet this goal with mutually beneficial partnerships between the innovators at the cutting edge and the industries and organizations with first-hand knowledge and understanding of these social issues and challenges. We've done a lot of talking about the need to build bridges. Let's start making these partnerships happen. The SPS's Chapters, conferences, and technical committees are in an ideal position to reach out to industry and ensure ongoing dialogues so that industry can share its various needs and goals and researchers can share their recent discoveries and novel ideas—a collaborative process that will motivate and benefit everyone. This will also enhance the magazine's efforts in

becoming an even more powerful mechanism of communication between researchers and practitioners.

My fifth goal for the SPS is to improve the administrative and financial efficiency of the Society. As we continue to strategize methods for adding value to our core mission initiatives, we must ensure that we have a reliable, efficient, transparent infrastructure supporting our operations. This includes ongoing improvement of our online presence

(including smart-phone applications and other mechanisms for continuous member feedback), controlling and reducing our publication costs, preparing solutions to counter-

act possible impact on our finances from open access and alternative publishing archives, and, last but not least, continuing to be the best custodians of all our financial affairs.

Each and every one of the aforementioned initiatives builds on the contributions of previous SPS staff and volunteers, and they cannot be fulfilled without the continued efforts of our various committees, including our technical committees, our various boards, including our Board of Governors, our Society officers, and current staff and volunteers. I am genuinely grateful to my predecessors for having left the Society in great shape. I am also grateful to all of the volunteers for their dedication and commitment. I'm ever-mindful of the fact that SPS will only grow through the direct engagement of its members. We on the governing boards are committed to doing our best, but we need you with us to help us ensure that SPS provides the greatest value to its members, to the field of SP, and to the public. I know that together we can address and solve the Society's critical challenges.




The SPS has already started several great initiatives to provide services to members and to the SP community at large.

rate reduction is shown for phone recognition and large vocabulary voice search.

October 2014

Robust Sound Event Classification Using Deep Neural Networks

McLoughlin, I.; Zhang, H.; Xie, Z.; Song, Y.; Xiao, W.

This paper outlines a sound event classification framework that compares auditory image front end features with spectrogram image-based front-end features, using a support vector machine and deep neural network classifiers. Performance is evaluated on a standard robust classification task in different levels of corrupting noise, and with several system enhancements.

March 2015

Keyword Extraction and Clustering for Document Recommendation in Conversations

Habibi, M.; Popescu-Belis, A.

The authors extract key words from an automatic speech recognition system and use a topic model that favors diversity in the key word set and derive topically separated queries to run on Wikipedia. It attempts to maximize the probability of making at least one relevant recognition. Testing with human judges shows improvement over word frequency or topic similarity methods.

April 2015

A Regression Approach to Speech Enhancement Based on Deep Neural Networks

Xu, Y.; Du, J.; Dai, L.-R.; Lee, C.-H.

This paper finds a map from a noisy speech signal to a clean one using deep neural networks. Dropout and noise-aware training strategies lead to robust performance to highly nonstationary noise.

January 2015

From Feedforward to Recurrent LSTM Neural Networks for Language Modeling

Sundermeyer, M.; Ney, H.; Schluter, R.

This paper compares count models to feedforward, recurrent, and long



IMAGE LICENSED BY INGRAM PUBLISHING

short-term memory neural network variants on two large-vocabulary speech recognition tasks. The increased computational complexity requires efficient search methods for the neural networks. Performance is evaluated in terms of perplexity and word error rate.

March 2015

Sentence Compression for Aspect-Based Sentiment Analysis

Che, W.; Zhao, Y.; Guo, H.; Su, Z.; Liu, T.

The method described in this paper compresses sentences by removing information unrelated to sentiment using a conditional random field model. The shorter sentences are easier to parse allowing fine-grained aspect-based sentiment analysis.

December 2015

An Overview of Noise-Robust Automatic Speech Recognition

Li, J.; Deng, L.; Gong, Y.; Haeb-Umbach, R.

This paper provides a thorough overview of modern noise-robust techniques for ASR developed over the past 30 years. Noise robust techniques are analyzed using five criteria: 1) feature-domain versus model-domain processing, 2) the use of prior knowledge about the acoustic environment distortion, 3) the use of explicit environment-distortion models, 4) deterministic versus uncertainty processing, and 5) the

use of acoustic models trained jointly with the same feature enhancement or model adaptation process used in the testing stage.

April 2014

Single Frequency Filtering Approach for Discriminating Speech and Nonspeech

Aneesa, G.; Yegnanarayana, B.

The mean and variance of the noise-compensated weighted envelopes are computed across frequency at each time instant. Because the variance of the spectral information across frequency is

higher for speech and lower for many types of noises, the method obtains better performance than adaptive multi-rate VAD2.

April 2015

This issue's "Reader's Choice" column lists the top ten papers most downloaded for the past year.

Voice Conversion Using RNN Pre-Trained by Recurrent Temporal Restricted Boltzmann Machines

Nakashika, T.; Takiguchi, T.; Ariki, Y.

This paper converts voices using one recurrent temporal restricted Boltzmann machine for the source and one for the destination speaker to convert both speaker signals to the model parameters, then a neural network to convert between the models' parameters. This method compares favorably with Gaussian mixture model methods.

March 2015

SP

SPECIAL REPORTS

John Edwards

Signal Processing Helps Put Robot Users in Control

According to BI Intelligence, the global robotics market—long dominated by industrial and logistics uses—is finally beginning to see a shift toward consumer and office applications. There will be a US\$1.5 billion market for consumer and business robots by 2019, predicts the technology market research firm, headquartered in London.

Another technology research organization, Juniper Research, expects that more than one in ten U.S. households will own a consumer robot by the end of the decade, up from fewer than one in 25 in 2015, primarily “task-oriented” robots assigned to take over household chores, such as lawn mowing or vacuum cleaning.

According to Juniper Research, located in Basingstoke, Hampshire, England, devices such as iRobot’s Roomba robot vacuum and Droplet Robotics’ Sprinkler, already help make life more convenient for consumers and, despite design compromises, are likely to usher in a new era of housekeeping.

Yet, before businesses and consumers can begin routinely using robots capable of handling even more complex tasks, a new generation of easy and intuitive control methodologies will have to be designed, tested, and implemented. Prospective users will also have to be convinced that robot controls are both reliable and safe. This will not be easy, however, in light of the fact that even as software algorithms continue to improve, user mistrust remains high.

A 2014 University of Pennsylvania study demonstrated that software is

given little room for error by users before trust is seriously eroded [1]. In fact, the study concluded that humans naturally tend to place more trust in another human’s judgment, even when an algorithm is shown to outperform its flesh and blood counterpart.

“It just works” was a slogan promoted by Apple several years ago in an effort to convince skeptical Windows PC users to switch to the company’s purportedly simpler and easier to use Macintosh computer. Robot-control developers are now taking that same catch phrase to heart as they develop and refine promising new ways of making robots follow their users’ instructions quickly, obediently, and faithfully.

Thought control

Simple, intuitive robot control is essential for people who depend on robotic-enabled assistive devices, such as wheelchairs and performance-enhancing exoskeletons.

A team of researchers at Ecole Polytechnique Fédérale de Lausanne’s Defitech Foundation Chair in Brain-Machine Interfaces are focusing on a revolutionary brain-machine approach with the goal of restoring a sense of independence to the disabled. The researchers are working on a technology that would allow users to remotely control a robot with one’s thoughts.

“Our work focuses on the design of neuroprostheses—robots and exoskele-

tons that human users control directly by voluntarily modulating their brain activity,” says research leader José del R. Millán. “At the core of these neuroprostheses, there is a brain-computer interface (BCI), a system that records neural signals and decodes them in order to transform the user’s intention into appropriate commands to operate practical devices for motor-disabled people such as wheelchairs, telepresence robots, hand and lower-limb exoskeletons” (Figure 1).

An important aspect of a BCI is the capability to distinguish between different patterns of brain activity, with each being associated with a particular intention or mental task. “Adaptation is a key component of a BCI because, on the one hand, users must learn to modulate their neural activity to generate distinct brain patterns while, on the other hand, machine-learning techniques need to discover the individual brain patterns

characterizing the user’s intention,” Millán says. “In essence, a BCI is a two-learner system that must engage in a mutual adaptation process.”

Millán says that the researchers are focusing their attention on electrical brain signals origi-

nating directly from neuron activity, since their high temporal resolution is suitable for operating robotic devices in real time. “Brain signals for a BCI can be recorded from single neurons, using microelectrode arrays implanted in the

Simple, intuitive robot control is essential for people who depend on robotic-enabled assistive devices, such as wheelchairs and performance-enhancing exoskeletons.

brain, or as the concerted activity of neuronal populations of different sizes, depending on the position of the electrodes—either implanted in the brain, on the surface of the brain or outside the scalp.” He notes that a combination of these approaches may be necessary to achieve the ultimate goal of controlling a neuroprosthesis as easily and precisely as able-bodied people control their natural limbs.

In tests involving nine disabled and ten healthy people in Italy, Germany, and Switzerland, participants wore an electrode-studded hat that detected their brain signals. The individuals then instructed a robot to move in various ways. “Each of the nine subjects with disabilities managed to remotely control the robot with ease after fewer than ten days of training,” Millán says. The tests ultimately revealed no difference in piloting ability between healthy and disabled participants.

The researchers also believe that at least some degree of robot autonomy should be available to supplement and complement user control. A robotic mobility device, for example, should be able to avoid obstacles by itself, even when it is not told to, Millán notes. To avoid becoming overly tired, the user should be able to take a break from giving instructions, allowing the robot to continue on its current path until it receives an order to stop or change course. “In this way, control over the robot is shared between the human and the computer, allowing the pilot to rest while navigating,” Millán says.

Signal processing is a critical component in several areas of BCI research. “The first is to increase the signal-to-noise ratio of the recorded signals, which have very low amplitude on the order of microvolts,” Millán says. “Typical tools at this stage are filtering in the frequency and spatial domain, as we record from many electrodes at a relatively high sampling rate.” Signal processing also helps to extract features that may reflect different neuronal processes, each associated to an aspect of the user’s intent. “For instance, imagination and execution of a movement gives rise to rhythmic activity in different frequencies in the sensorimo-



FIGURE 1. A robotic wheelchair equipped with a BCI that records neural signals and decodes them in order to transform user’s intentions into appropriate commands. (Photo courtesy of Ecole Polytechnique Fédérale de Lausanne’s Defitech Foundation Chair in Brain-Machine Interfaces.)

tor cortex of the corresponding body part,” Millán explains. “Out of many potential candidates, machine-learning techniques select those features that improve decoding.”

Millán says that one of the biggest signal processing-related challenges that researchers face is the intrinsic variability of brain signals, which makes it difficult to decode a user’s intentions from within a short time window. “We deal with this issue by using statistical approaches that combine evidence accumulated over time to robustify the final decision,” he says.

Millán is confident that the BCI will ultimately be used by robotic mobility system manufacturers worldwide, but not for at least several more years. “It will require large trials to demonstrate the robustness and reliability of the technology,” he says. “This is not only time-consuming, but will also require substantial financial resources to cover all the necessary personnel to run the trials.”

Finger control

Researchers at the Georgia Institute of Technology (Georgia Tech) believe that businesses and consumers will soon have the ability to control entire fleets

of robots with just the flick of a finger. Their new tablet-based system is designed to be used by almost anyone, including people with no technical training. To make a swarm of robots do his or her bidding, the user simply taps the tablet display to control where a beam of red light appears on a floor. The robots then will then move toward the illuminated area, constantly communicating with each other and deciding how to evenly cover the lit space.

If the user swipes a finger across the tablet display to drag the light across the floor, the robots will obediently follow. If the operator places two fingers in different locations on the tablet, the robots will begin splitting into teams and repeat the even-covering process within the two specified areas. “Basically the robots move around so as to balance how much light they are responsible for in the sense that they should end up with the same amount of light in the different areas of responsibility,” says Magnus Egerstedt, Schlumberger Professor in Georgia Tech’s School of Electrical and Computer Engineering (Figure 2).

“A few years ago, we ran user studies in the lab and found that people were generally quite bad at controlling large



FIGURE 2. Magnus Egerstedt, a professor in Georgia Tech's School of Electrical and Computer Engineering, with some of the robots he is using in his research into controlling swarms of machines that coordinate their own activities to accomplish tasks. (Photo courtesy of Georgia Tech's School of Electrical and Computer Engineering/Rob Felt.)

teams,” Egerstedt says. “So we started looking for easier ways of giving people control over large teams and the result was the tablet interface, which allows human operators to simply swipe desired shapes on the tablet and then have the robots respond accordingly.”

The new algorithm promises easy control over large teams of robots, even by a single operator, in applications such as manufacturing, agriculture and search and rescue, Egerstedt says. “The core algorithms behind this technology make robots move around in a distributed manner by only paying attention to neighboring robots,” he explains.

“For this to work, the robots need to measure not only their own positions, but also where other robots are relative to themselves.”

Signal processing lies at the heart of the system. “This requires a lot of signal processing in that the robots must make sense of noisy and sometimes intermittent sensor measurements,” Egerstedt says. “The sensor measurements are both sampled and filtered for smoothness, so sampling and filtering are key.”

The Georgia Tech model differs from most other robotic coverage algorithms in that it is not static. “It is flexible enough to allow robots to change

their minds effectively, rather than just performing the single job they’re programmed to do,” Egerstedt says.

Inside a lab, the robots can easily measure light intensities. Yet relying on light beams is hardly practical out on a farm field or in a disaster area. “What the light is really representing is areas of interest—the more light there is, the more interesting stuff there is,” Egerstedt says. “This is called a density function and, when deployed, we are envisioning that the tablet simply communicates what the density function looks like to the robots and they then execute the same algorithm without the need to actually measure any light.”

Egerstedt says that one of the biggest challenges his team has encountered so far is a multitarget matching problem—keeping track of which signal corresponds to which robot. “This has really given us the most headaches from a signal processing point of view,” he states.

“Technically, what the robots are doing is solving a ‘dynamic coverage problem,’ which is the name we use for the technology when we publish papers or give technical talks about it,” Egerstedt says. The project began in January 2014. “We have no firm end date; we will keep going until we have found the perfect solution,” Egerstedt adds.

Complex collaboration

One of the biggest challenges developers face as they attempt to get robots to work together is the fact that the human world is packed with uncertainty. A new robot management system developed by researchers at the Massachusetts Institute of Technology’s (MIT’s) Computer Science and Artificial Intelligence Laboratory (CSAIL) combines existing control programs to enable multiagent systems, such as teams or swarms of robots, to collaborate effectively and reliably in a variety of complex ways, whether they are working in a warehouse, searching earthquake rubble for survivors, or flying through the air delivering packages to homes.

According to Jonathan P. How, the Richard Cockburn Maclaurin Professor of Aeronautics and Astronautics at MIT,

the new system factors in uncertainty, such as a communication link will suddenly drop or that a specific algorithm will unintentionally direct a robot into a wall or other type of dead end, and automatically plan around such a situation occurring. “What the technology solves is the problem of how to get agents to work together when communications is an issue,” says How, who developed the system with student Chris Maynor. “The project goal is to find decision-making algorithms that can solve the problem and then decide which agent should do what and when it should do it,” How says.

As the project moved forward, the researchers decided that it was important to have robots to view specific tasks in much the way people do. Humans, after all, do not have to think about every single footstep they take. Through experience, such actions gradually become second nature. With this concept in mind, the team programmed the robots to perform a series of “macro-actions” broken down into multiple steps.

The new approach uses three major components, beginning with a collection of low-level control algorithms—the macro-actions that govern robot behaviors either collectively or individually. The system also uses a set of statistics that cover the algorithms’ execution—decision making—in specific situations. The third part of the system is a method that assigns values to various outcomes (successfully completing a task generates a high positive valuation while excessive energy use would lead to a negative valuation).

To demonstrate their system, CSAIL researchers temporarily converted their lab into a miniature lounge featuring a PR2 robot “bartender” (Figure 3) and a pair of four-wheeled Turtlebot robots that rolled into nearby offices and retrieved drink orders from the human occupants. The Turtlebots then reasoned about which orders were required in the different rooms and whether or not other robots may have already delivered the drinks.

“There are a lot of big challenges with planning any type of communica-



FIGURE 3. MIT CSAIL researchers temporarily converted their lab into a miniature lounge that included a PR2 robot “bartender” and two four-wheeled Turtlebot “waiters.” (Photo courtesy of MIT/CSAIL.)

tion network,” How says. “To be honest, it has been very tough getting information flowing in a reasonable fashion in terms of speed and time.” He notes that, in real-world domains, such as on busy streets, inside buildings, or on treacherous natural landscapes, communication

noise and uncertainty about what is happening makes it hard for robots to make decisions. “Getting clean information from one agent to another, and having them understand what’s received and process it, that is what we are seeking,” How says.

“There’s a bunch of different aspects of signal processing that we’re doing,” How remarks. “A lot of it has to do with a core filtering problem that is going on as each of the agents is measuring the world and extracting information.”

Author

John Edwards (jedwards@edwardsmedia.com) is a technology writer

based in the Phoenix, Arizona, area.

Reference

[1] B. J. Dietvorst, J. P. Simmons, and C. Massey, “Algorithm aversion: People erroneously avoid algorithms after seeing them err,” *J. Exp. Psychol.: Gen.*, vol. 144, no. 1, pp. 114–126, Feb. 2015.

SP

2017-2018

IEEE-USA Government Fellowships



Congressional Fellowships
Seeking U.S. IEEE members interested in spending a year working for a Member of Congress or congressional committee.



Engineering & Diplomacy Fellowship
Seeking U.S. IEEE members interested in spending a year serving as a technical adviser at the U.S. State Department.



USAID Fellowship
Seeking U.S. IEEE members who are interested in serving as advisors to the U.S. government as a USAID Engineering & International Development Fellow.

The application deadline for 2017-2018 Fellowships is 23 December 2016.

For eligibility requirements and application information, go to www.ieeeusa.org/policy/govfel or contact Erica Wissolik by emailing e.wissolik@ieee.org or by calling +1 202 530 8347.




FROM THE EDITORS

Z. Jane Wang, Wade Trappe, and Shuguang Cui

A Feature Article Cluster on Brain Signal Analytics: Analytical Approaches to Enhanced Understanding of Brain Function

Special issues and individual feature articles are two major mechanisms of full-length tutorial surveys of *IEEE Signal Processing Magazine (SPM)*. Special issues are on topics proposed by guest editor teams, who then solicit contributions through an open call for papers to select and assemble about eight to 12 articles complementing each other; the process has a time frame of more than a year and is carried out in a well-structured and coordinated fashion. In contrast, feature articles allow for the contribution of individual articles by authors and are usually handled independent of each other. This past year, the magazine's editorial team explored a mechanism complementing the existing two-leg structure—an article cluster/series that may allow us to embrace important emerging areas more quickly than the current special-issue operation, while providing a more balanced coverage than the current feature article mechanism by one author team. Supported by the senior editorial board, the topic of our first feature article cluster is on a highly interdisciplinary topic of brain signal processing.

Recent advancements in neuroimaging technology have accelerated our understanding of the functioning of human brain, as evidenced by an ever-increasing number of publications associated with brain mapping and reflected

by several concerted international efforts, such as the U.S. BRAIN Initiative, the European Human Brain Project, and the Human Connectome Project, which seek to revolutionize our understanding of the connected brain. The utilization of improved technologies for probing the brain, however, results in vast amounts of data, necessitating concomitant advances in efficient approaches and techniques for data acquisition, modeling, and analytics. It is thus widely envisioned that brain data analytics will play a major role over the next decade in revealing the brain's functional architecture and operational principles.

Traditionally, such brain data has been analyzed from the approach of signal or image processing, but these outcomes likely were published in such journals as *Human Brain Mapping* and *NeuroImage*, outside the IEEE Signal Processing Society (SPS). The goal of this feature article cluster is to bring together a diverse but

complementary set of tutorial/review articles demonstrating the importance of incorporating signal processing strategies into the advances in neuroimaging techniques, data analytics, and modeling for brain function. With this cluster, we would like to showcase the inherently interdisciplinary nature of brain mapping research and the intriguing related signal processing

questions. This will bring engineers, physicists, and neurobiologists together to better understand brain data, encourage future brain data analytics and modeling submissions to IEEE SPS journals, and alert the signal processing community to the exciting potential of brain mapping—an area of growing significance.

In this issue

This cluster comprises five articles from research groups in the United States, Canada, and Europe, includes researchers ranging from engineering to medicine, and provides a broad view of the recent advances regarding two important research directions in brain mapping: 1) brain connectivity and multimodal imaging and 2) multimodal data fusion.

In addition to recent technical advances and developments, all of these articles present interesting medical applications and provide an outlook of future challenges in building a comprehensive understand-

ing of brain mapping.

The first three articles focus on brain connectivity. Brain connectivity is of critical importance for the understanding and assessment of brain functions in normal and nonhealthy states. Brain connectivity can be described at several levels of temporal and spatial scales, and it refers to several different types of

It is thus widely envisioned that brain data analytics will play a major role over the next decade in revealing the brain's functional architecture and operational principles.

connectivity, including anatomical, functional, and effective. For instance, the study of changes in brain networks over time, termed the *chronnectome*, was recently highlighted as one of the “Best of 2014” by the U.S. National Institute of Mental Health. The article by Razi and Friston, “The Connected Brain,” reviews the history of brain mapping and highlights the theoretical ad-

vances made in the popular dynamic causal modeling used for brain connectivity. Li, Shi, and Toga’s article, “Mapping Brain Anatomical Connectivity Using Diffusion Magnetic Resonance Imaging,” focuses on diffusion magnetic resonance imaging and covers key components in the workflow for mapping structural connectivity of the human brain. In “Time-Varying Brain Connectivity in fMRI Data,” Calhoun and Adali focus on time-varying brain connectivity modeling via functional magnetic resonance imaging signals by highlighting some key signal processing aspects of the chronnectomics work and reviewing a family of whole-brain, data-driven blind source separation approaches.

The last two articles focus on the multimodal theme, with one focusing on multimodal imaging techniques and the other on multimodal data fusion. The analysis of multiset data or multimodality imaging data is inherent to many problems in science and engineering, including brain data analytics. Technologies for brain imaging have limitations individually, and, thus, ways to synergistically derive information from complementary modalities have the potential to substantially enhance our understanding of underlying brain activities. The article by Zaidi and Becker, “The Promise of Hybrid PET/MRI,” focuses on the recent advances of the hybrid positron emission tomography/magnetic

We hope that this collection will help introduce more signal processing researchers into this area to accelerate its advancement.

resonance imaging, a major technological breakthrough in neuroimaging, and particularly discusses the important signal processing related topics such as MRI-guided image recon-

struction and correction and presents some potential clinical applications. In the article “Joint Blind Source Separation for Neurophysiological Data Analysis,” Chen, Wang, and McKeown provide an

overview and taxonomy of representative joint blind source separation methods and discuss several real-world neurophysiological applications from both the multiset and multimodal perspectives.

We believe the articles selected for this cluster offer a snapshot of the latest research and showcase the importance of signal processing in the exciting field of brain mapping and brain data analytics. We hope that this collection will help introduce more signal processing researchers into this area to accelerate its advancement.

Acknowledgments

We would like to acknowledge all authors for their outstanding contributions and the reviewers for their thoughtful comments. We are also grateful to Dr. Min Wu, the editor-in-chief of *SPM*, for her gracious support of the cluster, and to Jessica Barragué, managing editor of the magazine, and Rebecca Wollman, SPS publications administrator, for their diligent work and patience in coordinating the cluster.

About the editors



Z. Jane Wang (zjanew@ece.ubc.ca) received the B.Sc. degree from Tsinghua University, Beijing, China, in 1996 and the Ph.D. degree from the University of Connecticut in 2002. Since August 2004,

she has been with the Department of Electrical and Computer Engineering at the University of British Columbia, Canada, where she is currently a professor. Her research interests are in the broad areas of statistical signal processing theory and applications. She is currently a member of the senior editorial board of *IEEE Signal Processing Magazine*.



Wade Trappe (trappe@winlab.rutgers.edu) is a professor in the Electrical and Computer Engineering Department at Rutgers University, and associate director of Rutgers’s WINLAB directing research in wireless security. He received five best papers awards in areas of media security, Internet design, cognitive radio, and mobile computing. He has served as an area editor for *IEEE Signal Processing Magazine (SPM)* and as a regional director of the IEEE Signal Processing Society, and chaired the steering committee of *IEEE Transactions on Mobile Computing*.



Shuguang Cui (cui@tamu.edu) received his Ph.D. degree from Stanford University in 2005. He is currently a professor at Texas A&M University. His current research interest focuses on data-oriented large-scale information analysis and system design. He was selected as the Thomson Reuters Highly Cited Researcher and listed in the “World’s Most Influential Scientific Minds” by ScienceWatch in 2014. He received the IEEE Signal Processing Society 2012 Best Paper Award. He is an IEEE Fellow and IEEE Communications Society Distinguished Lecturer. He is currently the area editor for feature articles of *IEEE Signal Processing Magazine*.

SP

The Connected Brain



IMAGE LICENSED BY INGRAM PUBLISHING

Causality, models, and intrinsic dynamics

Adeel Razi and Karl J. Friston

Recently, there have been several concerted international efforts—the BRAIN Initiative, the European Human Brain Project, and the Human Connectome Project, to name a few—that hope to revolutionize our understanding of the connected brain. During the past two decades, functional neuroimaging has emerged as the predominant technique in systems neuroscience. This is foreshadowed by an ever-increasing number of publications on functional connectivity, causal modeling, connectomics, and multivariate analyses of distributed patterns of brain responses. In this article, we summarize pedagogically the (deep) history of brain mapping. We highlight the theoretical advances made in the (dynamic) causal modeling of brain function, which may have escaped the wider audience of this article, and provide a brief overview of recent developments and interesting clinical applications. We hope that this arti-

Digital Object Identifier 10.1109/MSP.2015.2482121
Date of publication: 27 April 2016

cle engages the signal processing community by showcasing the inherently multidisciplinary nature of this important topic and the intriguing questions that are being addressed.

Introduction

In this article, we use several key dichotomies to describe the evolution and emergence of modeling techniques used to characterize brain connectivity. We provide a historical overview of the brain connectivity literature, starting with the fundamental distinction between functional segregation and integration. In so doing, we introduce a key difference between functional and effective connectivity and emphasize their relationship via underlying models of distributed processing. Next, we consider various causal modeling techniques that are used to infer directed brain connectivity. With the help of a unified framework—based on (neuronal) state-space models—we show how (with a succession of simplifying approximations) standard models of connectivity

can be derived and how various measures of statistical dependencies arise from a generative (state-space) model of neuronal dynamics. Finally, we focus on the application of dynamic causal modeling (DCM) to endogenous neuronal activity and simulations of neuronal fluctuations based on the connectome. We describe a series of recent (and rapid) developments in modeling distributed neuronal fluctuations and how this modeling rests on functional connectivity. We contextualize these developments in terms of some historical distinctions that have shaped our approaches to connectivity in functional neuroimaging.

Notation

We use lowercase italics, x , for scalars and lowercase bold for vectors, \mathbf{x} , and vector functions, $\mathbf{x}(t)$, where each element represents a time-dependent state. Matrices are shown as uppercase bold, \mathbf{X} . In this article, $*$ corresponds to a convolution operator, \dagger denotes the complex conjugate transpose, $\langle \cdot \rangle$ denotes expectation, and \sim denotes discrete time-lagged variables. Fourier transforms of variables are in italic uppercase, such that $\mathbf{FT}(\mathbf{x}(t)) = \mathbf{X}(\omega)$. We use $F(\cdot)$ to denote a variational free-energy functional.

A historical perspective on brain connectivity

The notion of connectivity has a long history in brain imaging that can be traced back to the debates around classicism, modularity, and connectionism. In the recent past, a common notion among neuroscientists was that many functions of the brain were predetermined by its structure and that its structure was programmed by our genes. This view emphasized functional segregation and localizationism, tracing its history back to the days of phrenology (from Gall in the 18th century). Functional localization implies that a function can be localized in a cortical area. This is more general than functional segregation, which suggests that a cortical area is specialized for some aspect of neural processing and that this specialization is anatomically segregated within the cortex. This is similar to an understanding of how computers work, where each part has a preassigned function that cannot be substituted with other parts. However, in past decades, this view has changed, with clear evidence that the neural pathways in our brain are flexible, adaptable, connected, and moldable by changes in our environment or by injury or disease. In short, the brain is quintessentially plastic and can adapt and adopt new functionalities through necessity. This understanding rests on the notion of *connectionism* (a term first coined by Donald Hebb in the 1940s), with the central idea that brain function can be understood as the interaction among simple units, for example, neurons connected by synapses, that give rise to a connected whole that changes over time. Connectionism is closely related to (hierarchical) distributed processing, a perspective that has been substantiated by the work of Hubel and Wiesel (recipients of the Nobel Prize in Physiology or Medicine 1981) on how information

is processed in the visual cortex. They found that the visual system comprises simple and complex cells arranged in a hierarchical fashion. This finding underwrites the focus on neural network implementations based on hierarchical distributed constructs, leading to recent exciting developments in machine learning (e.g., hierarchical Bayesian inference [1] and deep learning algorithms [2]).

These ideas emerged in functional brain imaging as functional segregation and functional integration. Since their inception, there has been a sustained trend to move from functional segregation (and the study of regionally specific brain activation) toward functional integration (and the study of its connectivity). Functional localization implies that a function can be localized to a cortical area, whereas segregation suggests that a

cortical area is specialized for some aspects of perceptual or motor processing and that this specialization is anatomically segregated within the cortex. The cortical infrastructure supporting a single function may then involve many specialized areas whose union is mediated by the functional integration among them. In this view, functional segregation is meaningful only in the context of functional integration and vice versa. There are several descriptions of neuronal processing that accommodate the tendency for brain regions to engage in specialized functions (i.e., segregation) and the tendency to coordinate multiple functions (i.e., integration) through coupling specialized regions. This functional integration is a dynamic self-assembling process, with parts of the brain engaging and disengaging over time, and has been described by appealing to dynamical systems theory, for example, self-organized criticality [3], pattern formation, and metastability [4].

This review pursues another key theme—the distinction between functional and effective connectivity. This dichotomy relies on the definition of connectivity (i.e., functional integration) per se. The former uses a pragmatic definition of connectivity based on (Pearson) correlations and rests on statistical dependencies between remote neurophysiological events. However, this approach is problematic when dealing with distributed neuronal processes in the brain that are mediated by slender (axonal) neuronal connections or wires. A more mechanistic explanation of observed responses comes from the definition of effective connectivity that refers explicitly to the influence that one neural system exerts over another. In [5], it was proposed that “effective connectivity should be understood as the experiment and time-dependent, simplest possible circuit diagram that would replicate the observed timing relationships between the recorded neurons.” This speaks to two important points: effective connectivity is dynamic (activity dependent) and depends on a model of directed interactions or coupling, which we focus on in this review. Given this, an interesting development in functional connectivity now considers temporal dynamics, referred to as *dynamic functional connectivity* [6]. However, these developments fall short of furnishing a causal explanation of the sort provided by (model-based)

Several concerted international efforts hope to revolutionize our understanding of the connected brain.

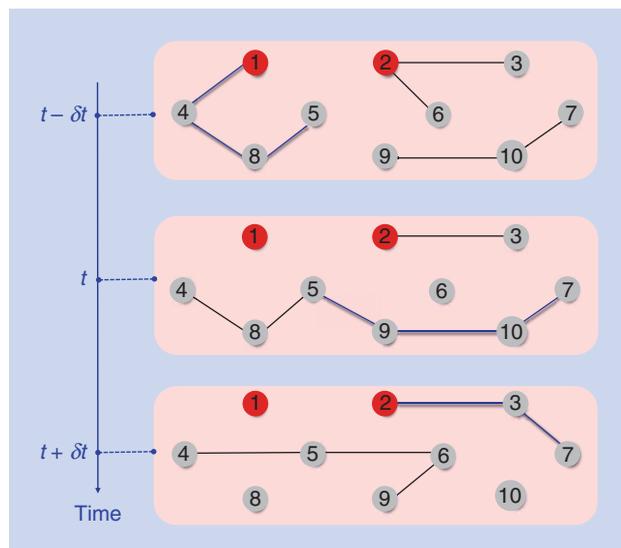


FIGURE 1. This schematic depicts a graph with undirected edges comprising ten nodes, where each node can be considered as a neuronal system. We sketched the evolution of this graph over three time points (under the assumption that each node retains a memory of past influences). Nodes 1 and 2 (shown in red) are the nodes in which a causal relationship is of interest. The key point of this example is that fluctuations in undirected coupling can induce directed dependencies.

effective connectivity. This is because functional connectivity is essentially a description of second-order data features, which precludes a mechanistic explanation of neurophysiological time series. Recent applications of DCM to ongoing (seizure) activity—in epilepsy—rest explicitly on dynamic functional connectivity to estimate the underlying fluctuations in effective connectivity or cortical gain control [7], [8]. In short, the operational distinction between functional and effective connectivity is important because it determines the nature of the inferences made about functional integration and the sorts of questions that can be addressed with careful consideration of the intricate interrelationship between effective and functional connectivity [9], [10].

Put simply, functional connectivity is a measure of statistical dependencies, such as correlations, coherence, or transfer entropy. Conversely, effective connectivity corresponds to the parameter of a model that tries to explain observed dependencies (functional connectivity). In this sense, effective connectivity corresponds to the intuitive notion of directed causal influence. This model-based aspect is crucial because it means that the analysis of effective connectivity can be reduced to model comparison, for example, the comparison of a model with and without a particular connection to infer its contribution to observed functional connectivity. In this sense, the analysis of effective connectivity recapitulates the scientific process because each model corresponds to an alternative hypothesis about how observed data were caused. In our context, these hypotheses pertain to causal models of distributed brain responses. Later, we consider analytical expressions that link effective and functional connectivity and show

that the latter can be derived from the former, whereas the converse is not true.

We have considered the distinction between functional segregation and integration in the brain and how the differences between functional and effective connectivity shape the way we characterize connections and the sorts of questions that are addressed to empirical data. In the next section, we look at the relationship between functional and effective connectivity and expand on the causal aspect of effective connectivity. Interested readers are directed to our previous review [10] for a more detailed discussion on brain connectivity.

Causal analyses of dynamical systems

The brain is a dynamic and self-organizing organ with emergent dynamics. These dynamics can be seen at multiple spatial and temporal scales; for example, there are tens of thousands of synaptic connections to a single neuron, which can fire dozens of times every second. Furthermore, this connectivity itself changes over multiple spatial and temporal scales. The spatial scale we are interested in, as measured by fMRI, is the macroscopic level, where we are interested in distributed processing or connectivity among neural systems and where each neural region or source comprises millions of neurons. As previously noted, the direction of information transfer or directed coupling is important. Figure 1 illustrates the fact that changes in connectivity over time underlie the causal relationship among neuronal systems. In Figure 1, we show a graph with undirected edges among ten nodes, where each node can be regarded as a proxy for a neuronal system (in general, these nodes could also be network devices in a communication network, e.g., exchanging e-mails). Alternatively, if the links represent a distance metric and nodes represent organisms, this could represent a model of how infections are disseminated. In this example, the graph evolves over time. Although the edges of the graph are undirected at each instance, adding a temporal aspect to this evolving graph enables one to infer directed information flow. (Note: Although we have not used the word *information* here in a strictly information-theoretic sense, there is a straightforward analogy between electrical impulses in neural systems and the classic communications theory picture of source, channel, and receiver [11].) For example, if we were interested in causal coupling between nodes 1 and 2 (red in Figure 1), we see that the activity in node 1 affects the activity in node 2, where we assume this influence endures over time. As we can see, node 1 is connected to node 2 via intermediate nodes 4, 8, and 5 (shown as blue edges) at time $t - \delta t$; nodes 9, 10, and 7 at time t ; and node 3 at time $t + \delta t$. This means that node 1 can affect node 2 in the future. However, the converse is not true, in that the activity in node 2 cannot affect the future of node 1. This asymmetry is a signature of causality (i.e., temporal precedence) and rests on accounting for the arrow of time. This is why, as we see in the next section, the statistical models used for characterizing effective connectivity are usually based on differential equations

(or difference equations in discrete time) and therefore explicitly take time into account. This simple example emphasizes the importance of temporal fluctuations in connectivity, even in undirected graphs. However, we do not want to give the impression that temporal precedence is necessary to infer causal relationships. Temporal precedence is an important aspect, and many definitions of causation require cause to precede effect [12], [13], for example, directed functional connectivity measures based on Yule–Walker formulations [vector autoregressive (VAR) models]. However, temporal precedence alone cannot distinguish effective connectivity from spurious dependencies caused by unknown factors. As an example, the barometer falls before the rain, but it does not cause the rain. The type of causality that we are concerned with is based on control theoretic concepts, where the causes (exogenous experimental inputs, endogenous random neural fluctuations, or both) produce effects (neural activity) that are observed empirically through hemodynamics as blood oxygen level-dependent (BOLD) signals. This form of causality is closely related to the probabilistic and graphical framework of causal calculus [14] (see “Simpson–Yule Paradox”), although there is a clear distinction between the two approaches, which we return to later.

We use state-space models to describe the basic concepts here and demonstrate that causality based on temporal precedence can be regarded as a special case of causality based on state-space graphs. In what follows, we look at several measures of causality in functional neuroimaging literature (which refer largely to fMRI but also hold for other modalities such as EEG, MEG, and local field potentials). These measures can be cast in terms of a generalization of state-space models based on stochastic differential equations.

State-space modeling of neuronal dynamics

The most natural framework for modeling distributed and coupled neural activity is to use state-space models. State-space modeling has its origin in control engineering, but the term *state-space* was first used by Kalman [24] and can be traced back to von Bertalanffy, who introduced general systems theory to biology in the 1940s and 1950s. We start with a generic description of coupled neuronal dynamics in terms of differential equations of the form

$$\dot{\mathbf{x}} = f(\mathbf{x}(t), \boldsymbol{\theta}, \mathbf{u}(t)) + \mathbf{w}(t) \quad (\text{state equation}), \quad (1)$$

$$\mathbf{y}(t) = h(\mathbf{x}(t), \boldsymbol{\theta}) + \mathbf{e}(t) \quad (\text{observation equation}), \quad (2)$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ represents a vector of n hidden state variables (where each state could correspond to a vast number of neurons in a cortical area, source, or spatial mode); $\dot{\mathbf{x}}(t)$ represents the change in those state variables; $\boldsymbol{\theta}$ are the underlying (connectivity) parameter that are assumed to be time-invariant; $\mathbf{y}(t)$ is the observed BOLD signal; and $\mathbf{w}(t)$ and $\mathbf{e}(t)$ are state noise (observation or instrument noise,

respectively), which makes this differential equation random. (Note: Strictly speaking, the hidden states include both neuronal and hemodynamic states; however, for simplicity, we ignore hemodynamic states in this article.) The (random) endogenous fluctuations $\mathbf{w}(t)$ on the motion of the hidden neuronal states represent the unknown influences (e.g., spontaneous fluctuations) that can only be modeled probabilistically. (Note: A reviewer of this article rightly pointed out that, in this exposition, we limited ourselves to an additive form of endogenous fluctuations that precludes the more general treatment of state-dependent neuronal fluctuation of the sort $f(\mathbf{x}(t), \boldsymbol{\theta}, \mathbf{u}(t), \mathbf{w}(t))$, which are used in modeling many complex volatile systems [25], including the brain [26].) The neuronal states are hidden because they cannot be measured directly. The function f defines the motion of the coupled dynamical system that is determined by inputs $\mathbf{u}(t)$, which we consider to be deterministic (but could also have stochastic component) and known. Inputs usually pertain to experimentally controlled variables, such as change in stimuli (a visual cue or an auditory signal) or instructions during an fMRI experiment (we see later that this exogenous input is absent in resting-state fMRI). This description of neuronal dynamics provides a convenient model of causal interactions among neuronal populations because it describes when and where exogenous experimental input $\mathbf{u}(t)$ perturbs the system and how (hidden) states influence changes in other states. Note that we have assumed that the form of the system dependencies f (and the connectivity parameters $\boldsymbol{\theta}$) are time-invariant, which means that we are assuming that the structural properties of the system will remain fixed over time (i.e., during the length of data acquisition).

We have not discussed the nature of the state and the observation noise process, which we consider in the section “Dynamic Casual Modeling of Intrinsic Networks.” For now, we assume that they possess usual noise properties, that is, they are independent and identically distributed. We describe a more general framework for analytic (non-Markovian) random fluctuations in the same section. A key simplification in this form of modeling is that we have lumped together many microscopic neuronal states to form hidden states $\mathbf{x}(t)$ that are abstract representations of neuronal activity (cf. a mean field approximation). In reality, the equations of motion—and the observer equation—describe very complicated interactions among millions of neurons. The formulation above corresponds to the amplitude of macroscopic variables or order parameters summarizing the dynamics of large neuronal populations. (Note: In statistical physics, the order parameter is a variable that indicates which phase you are in; for example, in a phase transition between liquid and gas, the order parameter may be the density.) Essentially, this means that the individual neurons become ordered, showing a coordinated dynamic pattern that can be described with the concept of order parameters. This sort of formulation can be motivated by basic

The notion of connectivity has a long history in brain imaging that can be traced back to debates around classicism, modularity, and connectionism.

Simpson–Yule Paradox

The Simpson–Yule paradox, or simply Simpson’s paradox [15]–[17], refers to the disconcerting situation in which statistical relationships between variables (e.g., x and y) are reversed or negated by the inclusion of an additional variable (z); for a more recent discussion, see [18]–[20]. A famous example of this paradox is when the University of California, Berkeley, came under investigation in 1975 for gender bias in graduate admissions. The graduate admissions statistics revealed that men applying were more likely to be admitted than women. However, when data were analyzed for each department separately, the reverse was true: no department was statistically significant in favor of men. The resolution of this paradox turned out to be that women applied for more competitive departments—with low success rates—in relation to men, who applied for fewer competitive majors with greater chances of acceptance. The main point is that conclusions based on data are sensitive to the variables we choose to hold constant, and that is why the “adjustment problem” is so critical in

the analysis of observational studies. Even now, no formal procedure has emerged that tells us whether adjusting for variable z is appropriate for the given study, setting aside intractable criteria [21] based on counterfactuals [22]. However, Simpson’s paradox is easily resolved with causal graphs. A simple graphical procedure provides a general solution to the adjustment problem [23]. This procedure is shown in Figure S1 and summarized as follows:

Objective

Check if z_1 and z_2 are sufficient measurements.

- 1) z_1 and z_2 should not be descendants of x .
- 2) Delete all nonancestors of $\{x, y, z\}$.
- 3) Delete all edges from x .
- 4) Connect any two parents sharing a child.
- 5) Strip arrowheads from all edges.
- 6) Delete z_1 and z_2 . Check if x is disconnected from y in the remaining graph; then z_1 and z_2 are appropriate measurements.

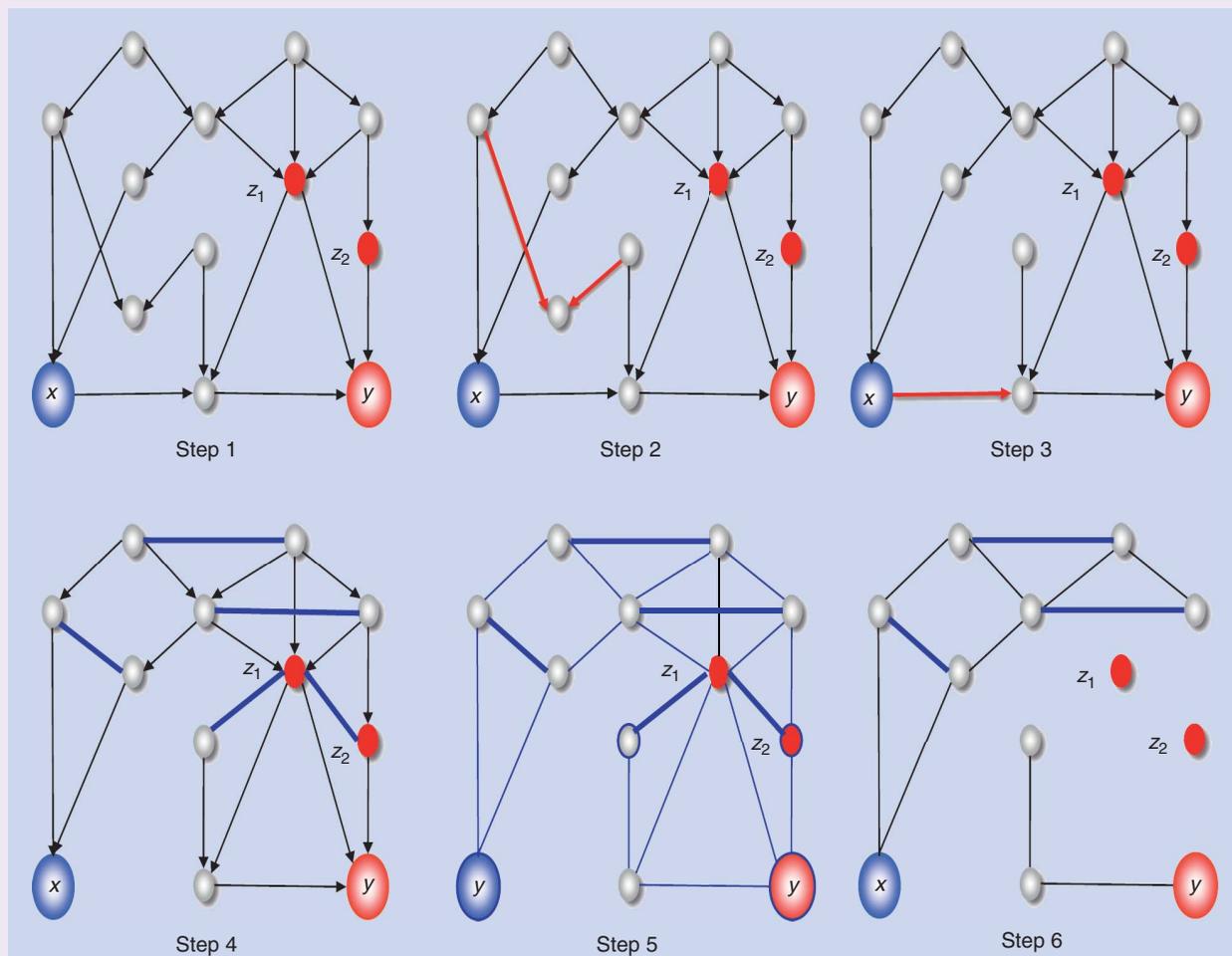


FIGURE S1. A simple graphical procedure provides a general solution to the adjustment problem. (Figure reproduced and redrawn with permission from [23].)

principles [27], for example, the center manifold theorem [28] and the slaving principle [29], [30], that apply generally to coupled dynamical systems.

State-space modeling and effective connectivity

The state and observation equations in (1) and (2) are generic representations; hence, there are several forms that the mappings or functions f and h can take. In turn, these define the sort of inference that can be made and the nature of causal relationships that can be identified from these models. We see in this section that almost all models in neuroimaging can be viewed as special cases of these equations.

Dynamic causal modeling

Although the application of general state-space models in neuroimaging has been around for decades, the explicit use of state-space models based on differential equations can be traced to [31], with the first introduction of a nonlinear neural mass model for EEG data. However, the most widely used and comprehensive framework, which uses Bayesian statistics to make model and parameter inferences, is DCM [32]. When first introduced, DCM used an ordinary differential equation (ODE) but was later extended to state-space models based on stochastic and random differential equations [33], [34]. The most widely used DCM is based on a Taylor expansion of (1) to its bilinear approximations:

$$\dot{\mathbf{x}}(t) = \left(\mathbf{A} + \sum_{j=0}^J \mathbf{B}^j \mathbf{u}_j \right) \mathbf{x}(t) + \mathbf{C} \mathbf{u}(t) + \mathbf{w}(t), \quad (3)$$

where $\mathbf{A} = \partial f / \partial \mathbf{x}$, $\mathbf{B} = \partial^2 f / \partial \mathbf{x} \partial \mathbf{u}$, and $\mathbf{C} = \partial f / \partial \mathbf{u}$ with $\boldsymbol{\theta}_n = \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$. The matrix \mathbf{A} is known as the Jacobian (or Laplace–Beltrami operator) describing the behavior—that is, the effective connectivity—of the system near its fixed point ($f(\mathbf{x}_0) = 0$), in the absence of the fluctuations $\mathbf{w}(t)$ and the modulatory inputs $\mathbf{u}(t)$. The matrices \mathbf{B}^j encode the change in effective connectivity induced by the j th input $\mathbf{u}_j(t)$, and \mathbf{C} embodies the strength of the direct influences of inputs $\mathbf{u}(t)$ on neural activity. In fMRI, the mapping from hidden states to the observed BOLD data $y(t)$ is based on a hemodynamic model that transforms hidden neuronal states of each population or region into predicted BOLD responses using a previously established biophysical model [32], [35], [36]. This hemodynamic model is based on four ODEs and five hemodynamic parameters $\boldsymbol{\theta}_h$, such that $\boldsymbol{\theta} = \{\boldsymbol{\theta}_n, \boldsymbol{\theta}_h\}$. The hemodynamic model describes how neuronal activity engenders vasodilatory signals that lead to increases in blood flow, which in turn changes the blood volume and deoxyhemoglobin content, which subtend the measured signal.

The bilinear approximation to our general state-space model of neurophysiological dynamics furnishes a probabilistic model that specifies the probability of observing any time series given the parameters. This is known as a *likelihood model* and usually assumes that the observed data are a linear mixture of the model predictions and Gaussian observation noise. By combining this likelihood model with prior beliefs (specified in terms of probability distributions), we have

what is called, in Bayesian statistics, a *generative model*. This allows one to use standard (variational) procedures to estimate the posterior beliefs about the parameters and, crucially, the model itself. The real power of DCM lies in the ability to compare different models of the same data. This comparison rests on the model evidence, which is simply the probability of the observed data under the model in question (and given known or designed exogenous inputs). The evidence is also called the *marginal likelihood* because one marginalizes or removes dependencies on the unknown quantities (hidden states and parameters). The model evidence can simply be written as

$$p(\mathbf{y} | m, \mathbf{u}) = \int p(\mathbf{y}, \mathbf{x}, \boldsymbol{\theta} | m, \mathbf{u}) d\mathbf{x} d\boldsymbol{\theta}. \quad (4)$$

Model comparison rests on the evidence for one model relative to another (see [51] for a discussion in the context of fMRI). Model comparison based on the likelihood of different models provides the quantitative basis for all evidence-based hypothesis testing. Usually one selects the best model using Bayesian model comparison, in which different models are specified in terms of priors on the coupling parameters. These are used to switch off parameters by assuming a priori that they are zero (to create a new model). In DCM, priors used are so-called “shrinkage priors” because the posterior estimates shrink toward the prior mean. The size of the prior variance determines the amount of shrinkage. With a null model m_0 and an alternative model m_1 , the Bayesian model comparison rests on computing the logarithm of the evidence ratio

$$\begin{aligned} \ln \left(\frac{p(\mathbf{y} | m_1)}{p(\mathbf{y} | m_0)} \right) &= \ln p(\mathbf{y} | m_1) - \ln p(\mathbf{y} | m_0) \\ &\approx F(\mathbf{y}, \boldsymbol{\mu}_1) - F(\mathbf{y}, \boldsymbol{\mu}_0), \end{aligned} \quad (5)$$

where $F(\cdot)$ is the free energy that provides an (upper bound) approximation to Bayesian model evidence. Note that we have expressed the logarithm of the marginal likelihood ratio as a difference in log evidences. This is the preferred form because model comparison is not limited to two models but can cover a large number of models whose quality can be usefully quantified in terms of their log evidences. A relative log evidence of three corresponds to a marginal likelihood ratio (Bayes factor) of about 20 to one, which is considered strong evidence in favor of one model over another [37]. An important aspect of model evidence is that it includes a complexity cost (which is sensitive not only to the number of parameters but also to their interdependence). This means that a model with redundant parameters would have less evidence, even though it provided a better fit to the data (see [51]). In most current implementations of DCM, the log evidence is approximated with a (variational) free-energy bound that (by construction) is always less than the log evidence. As we see in (5), this bound is a function of the data and (under Gaussian assumptions about the posterior density) some proposed values for the states and parameters. When the free energy is maximized (using gradient ascent) with respect to the proposed values, they become the maximum posterior or conditional estimates, $\boldsymbol{\mu}$, and the free energy, $F(\mathbf{y}, \boldsymbol{\mu}_1) \leq \ln p(\mathbf{y} | m)$, approaches the log evidence.

We return to Bayesian model comparison and the inversion of DCMs in the section “Biophysical Modeling of Neuronal Dynamics.” Next, we consider some alternative models. The first is a discrete time linear approximation to (1), which is the basis of Granger causality.

Vector autoregressive modeling

In contrast to DCM, in which causality is based on control theoretic constructs, (multivariate) autoregressive models [38]–[40] use temporal precedence for inferring causality in BOLD time series [41]. This is known as *directed functional connectivity* in neuroscience. It is straightforward to see that one can convert a state-space model—or DCM—into a VAR model with a few simplifying assumptions. Using a linear approximation to the state-space model of (1) and assuming that we can measure the neuronal states directly [i.e., $\mathbf{y}(t) = \mathbf{x}(t)$], then we can write

$$\mathbf{y}(t) = \tilde{\mathbf{A}}\mathbf{x}(t - \delta) + \mathbf{z}(t), \quad (6)$$

which can be written as

$$\mathbf{Y} = \tilde{\mathbf{Y}}\tilde{\mathbf{A}}^T + \mathbf{Z},$$

where $\tilde{\mathbf{A}} = \exp(\delta\mathbf{A})$ and $\mathbf{z}(t) = \int_0^\delta \exp(\tau\mathbf{A})\mathbf{w}(t - \tau)d\tau$. The second equality expresses the resulting VAR model as a simple general linear model with explanatory variables $\tilde{\mathbf{Y}}$ that correspond to a time-lagged (time \times source) matrix of states. Here, the unknown parameters comprise the autoregression matrix $\tilde{\mathbf{A}}$. Note that the innovations, $\mathbf{z}(t)$, are now a mixture of past fluctuations in $\mathbf{w}(t)$ that are remembered by the system. There is a clear distinction between fluctuations $\mathbf{w}(t)$ that drive the hidden states (1) compared with the innovations $\mathbf{z}(t)$ in (6) that underlie autoregressive dependencies among observation $\mathbf{y}(t)$. There is an important point to note here. Because the reparameterization of the effective connectivity in (3) uses a matrix exponential, the autoregressive coefficients $\tilde{\mathbf{A}}$ in (6) are no longer the parameters of the underlying effective connectivity among neuronal states. This means that any model comparisons—based on classical likelihood ratio tests such as Bayesian information criterion—will be making inferences about the statistical dependencies modeled by the autoregressive process and not about the causal coupling as in DCM. This is why connectivity measures based on autoregressive coefficients, for example, Granger causality [42], are regarded as directed functional connectivity as opposed to effective connectivity. A further distinction is that most Granger causality applications either ignore hemodynamic convolution or assume that hemodynamics are identical and noiseless [147]. An important aspect of Granger causality measures based on autoregressive formulations (we provide analytic links between the two in Figure S2) is that they can become unreliable in the presence of measurement noise and more so when underlying dynamics are dominated by slow (unstable) modes, quantified by the principal Lyapunov exponent [43]. However, there are several recent advances in the Granger causality literature that speak to these limitations [44]–[46].

Structural equation modeling

Structural equation modeling (SEM) [47] is another generic approach developed primarily in economics and social sciences [48], [49] and was used in (structural) neuroimaging for the first time in [50]. We can again see that SEM is a special case of (1) by appealing to the (adiabatic) assumption that neuronal dynamics have reached equilibrium at each point of observation—or, in other words, the dynamics are assumed to occur over a timescale that is short relative to the fMRI sampling interval. In terms of implementation, we can force this condition by having very strong shrinkage priors in DCM. With this assumption, we can reduce the generative model of (3) so that it predicts the observed covariance among regional responses over time instead of predicting the time series itself. Mathematically, this means that we assume $\mathbf{y}(t) = \mathbf{x}(t)$, $\mathbf{u}(t) = 0$, and $\dot{\mathbf{x}}(t) = 0$. This simply means that $\mathbf{x}(t) = \mathbf{y}(t) = -\mathbf{A}^{-1}\mathbf{w}(t)$, which implies that

$$\Sigma_y = \mathbf{A}^{-1}\Sigma_w(\mathbf{A}^{-1})^T, \quad (7)$$

where $\Sigma_y = \langle \mathbf{y}(t)\mathbf{y}(t)^T \rangle$ and $\Sigma_w = \langle \mathbf{w}(t)\mathbf{w}(t)^T \rangle$. Note that we do not have to estimate hidden states because the generative model explains observed covariances in terms of random fluctuations and unknown coupling parameters. The form of (7) has been derived from the generic generative model. In this form, it can be regarded as a Gaussian process model, where the coupling parameters become, effectively, parameters of the covariance among observed signals due to the hidden states. We can also give an alternative formulation of SEM in terms of path coefficients, but we skip this for brevity (for details, see [51]).

Although SEM has been used in fMRI literature, it provides a description of static dependencies; hence, it is not suitable for fMRI (and EEG/MEG) time series, in which the characteristic time constants of the neuronal dynamics and hemodynamics are much larger than the exogenous inputs that drive them. This means that testing for context-sensitive changes in effective connectivity becomes problematic in event-related designs. For example, [52] used simulated fMRI time series from a realistic network model for two task conditions in which the anatomical connectivity was known and could be manipulated. The results suggested that caution is necessary in applying SEM to fMRI data and illustrate that functional interactions among distal network elements can appear abnormal even if only part of a network is damaged.

Another issue when using SEM to infer effective connectivity is that we can only use models of low complexity—usually, (acyclic) models that have no recurrent connections [53]. This is because fitting the sample covariance means that we have to throw away lots of information in the original time series. Heuristically, the ensuing loss of degrees of freedom means that conditional dependencies among the estimates of effective connectivity are less easy to resolve. In machine-learning literature, SEM can be regarded as a generalization of inference on linear Gaussian Bayesian networks that relaxes the acyclic constraint. As such, it is a generalization of structural causal modeling, which deals with directed acyclic graphs (DAGs) (see next

section). This generalization is important in the neurosciences because of the ubiquitous reciprocal connectivity in the brain that renders it cyclic or recursive. Next, we turn to the description of time series based on second-order statistics and show that they can be analytically derived from the state-space model of (1).

Coherence, cross spectra, and correlations

Until now, we have considered only procedures for identifying effective connectivity from fMRI time series. However, the following important question remains: Is there an analytical relationship between functional and effective connectivity? This question is addressed schematically in “Measures of Connectivity” by showing how various measures of statistical dependencies (functional connectivity) are interrelated and how they can be generated from a DCM. This schematic contextualizes different measures of functional connectivity and how they arise from (state-space) models of effective connectivity. In other words, measures that are typically used to characterize observed data can be regarded as samples from a probability distribution over functions whose expectation is known. This means that one can treat normalized measures, such as cross-correlation functions and spectral Granger causality, as explicit functions of the parameters of the underlying generative process.

In “Measures of Connectivity,” we include common (descriptive) measures of functional connectivity that have been used in fMRI, such as the correlation coefficient (the value of the cross-correlation function at zero lag), coherence, and (Geweke) Granger causality [54]. These measures can be regarded as standardized (second-order) statistics based on the cross-covariance function, the cross-spectral density, and the directed transfer functions, respectively. In turn, they are determined by the first-order (Volterra) kernels, their associated transfer functions, and VAR coefficients. For readers not familiar with Volterra kernels, their use provides an alternative to the conventional identification methods by expressing the output signal as high-order nonlinear convolution of the inputs. This can simply be thought of as a functional Taylor expansion and can be regarded as a power series with memory (see [55] for a detailed discussion). All of these representations can be generated from the underlying state-space model used by DCM. Let us examine these relationships further. First, there is a distinction between the state-space model (upper two panels of Figure S2), which refers to hidden or system states, and representations of dependencies among observations (lower panels), which do not. This is important because, although one can generate the dependencies among observations from the state-space model, one cannot do the converse. In other words, it is not possible to derive the parameters of the state-space model (e.g., effective connectivity) from transfer functions or autoregression coefficients. This is why one needs a state-space model to estimate effective connectivity or, equivalently, why effective connectivity is necessarily model-based. Second, we have seen in previous sections that SEM and autoregressive representations can be derived from (bilinear

and stochastic) DCM in a straightforward manner (under certain assumptions). The convolution kernel representation in Figure S2 provides a crucial link between covariance-based second-order measures, such as cross covariance and cross correlation, and their spectral equivalents, such as cross spectra and coherence. Figure S2 also highlights the distinction between second-order statistics (lower two rows) and models of the variables per se (upper three rows). For example, convolution and autoregressive representations can be used to generate time series (or their spectral counterparts), whereas cross-covariance functions and autoregression coefficients describe their second-order behavior. This is important because this second-order behavior can be evaluated directly from observed time series. Indeed, this is the common way of measuring functional connectivity in terms of

(second-order) statistical dependencies. We also highlight the dichotomy between time and frequency representations (measures in the top panel). For example, the (first-order Volterra) kernels in the convolution formulation are the Fourier transform of the transfer functions in frequency space (and vice versa). Similarly, the directed transfer functions of the autoregressive formulation are based on the Fourier transforms of the

autoregression coefficients. Another distinction is between representations that refer explicitly to random (state and observation) noise and autoregressive representations that do not. For example, notice that the cross-covariance functions of the data depend on the cross-covariance functions of state and observation noise. Conversely, the autoregression formulation only invokes (unit normal) innovations (although the autoregression coefficients are an implicit function of both state and observation noise covariance functions). In the current setting, autoregressive representations are not regarded as models but simply as ways of representing dependencies among observations. This is because (hemodynamic) responses do not cause responses—hidden (neuronal) states cause responses.

Crucially, all of the aforementioned formulations of statistical dependencies contain information about temporal lags (in time) or phase delays (in frequency). This means that, in principle, all measures are directed in the sense that the dependencies from one region to another are distinct from the dependencies in the other direction. However, only the autoregressive formulation provides directed measures of dependency—in terms of directed transfer functions or Granger causality. This is because the cross-covariance and cross-spectral density functions between two time series are antisymmetric. The autoregressive formulation can break this (anti)symmetry because it precludes instantaneous dependencies by conditioning the current response on past responses. Note that Granger causality is, in this setting, a measure of directed functional connectivity [56]. This means that Granger causality (or the underlying autoregression coefficients) reflects directed statistical dependencies such that two regions can have strong autoregression coefficients or Granger causality in the absence of a direct effective connection. Finally,

Key modeling assumptions underlying DCM are motivated by neuroanatomical and neurophysiological constraints.

Measures of Connectivity

The upper panel in Figure S2 illustrates the form of a state-space model that comprises differential equations coupling hidden states (first equation) and an observer equation mapping hidden states $\mathbf{x}(t)$ to observed responses $\mathbf{y}(t)$ (second equation). Dynamic causal models are summarized by a Taylor (bilinear) approximation. Crucially, both the motion of hidden states and the responses are subject to random fluctuations, also known as state $\mathbf{w}(t)$ and observation $\mathbf{e}(t)$ noise. The form of these fluctuations is modeled in terms of their cross-covariance functions $\Sigma(t)$ of time t or cross-spectral density functions $\mathbf{g}(t)$ of (radial) frequency ω , as shown in the lower equations. Given this state-space model and its parameters Θ (which include effective connectivity), one can now parameterize a series of representations of statistical dependencies among successive responses as shown in the third row. These include convolution and autoregressive formulations shown on the left and right, respectively, in either time (pink and orange) or frequency (light green) space. The mapping between these representations rests on the Fourier transform, denoted by a dotted line, and its inverse. For example, given the equations of motion and observer function of the state-space model, one can compute the convolution kernels that, when applied to state noise, produce

the response variables. This allows one to express observed responses in terms of a convolution of hidden fluctuations and observation noise. The Fourier transform of these convolution kernels $\kappa(t)$ is called a transfer function $\mathbf{K}(t)$. Note that the transfer function in the convolution formulation maps from fluctuations in hidden states to response variables, whereas the directed transfer function in the autoregressive formulation $\mathbf{S}(t)$ maps directly among different response variables. These representations can be used to generate second-order statistics or measures that summarize the dependencies, as shown in the third row, for example, cross-covariance functions and cross spectra. The normalized or standardized variants of these measures are shown in the lower row and include the cross-correlation function (in time) or coherence (in frequency). The equations show how various representations can be derived from each other. All variables are either vector or matrix functions of time or frequency. For simplicity, the autoregressive formulations are shown in discrete form for the univariate case (the same algebra applies to the multivariate case, but the notation becomes more complicated). Here, $\mathbf{z}(t)$ is a unit normal innovation. Finally, note the Granger causality is only appropriate for a bivariate time series.

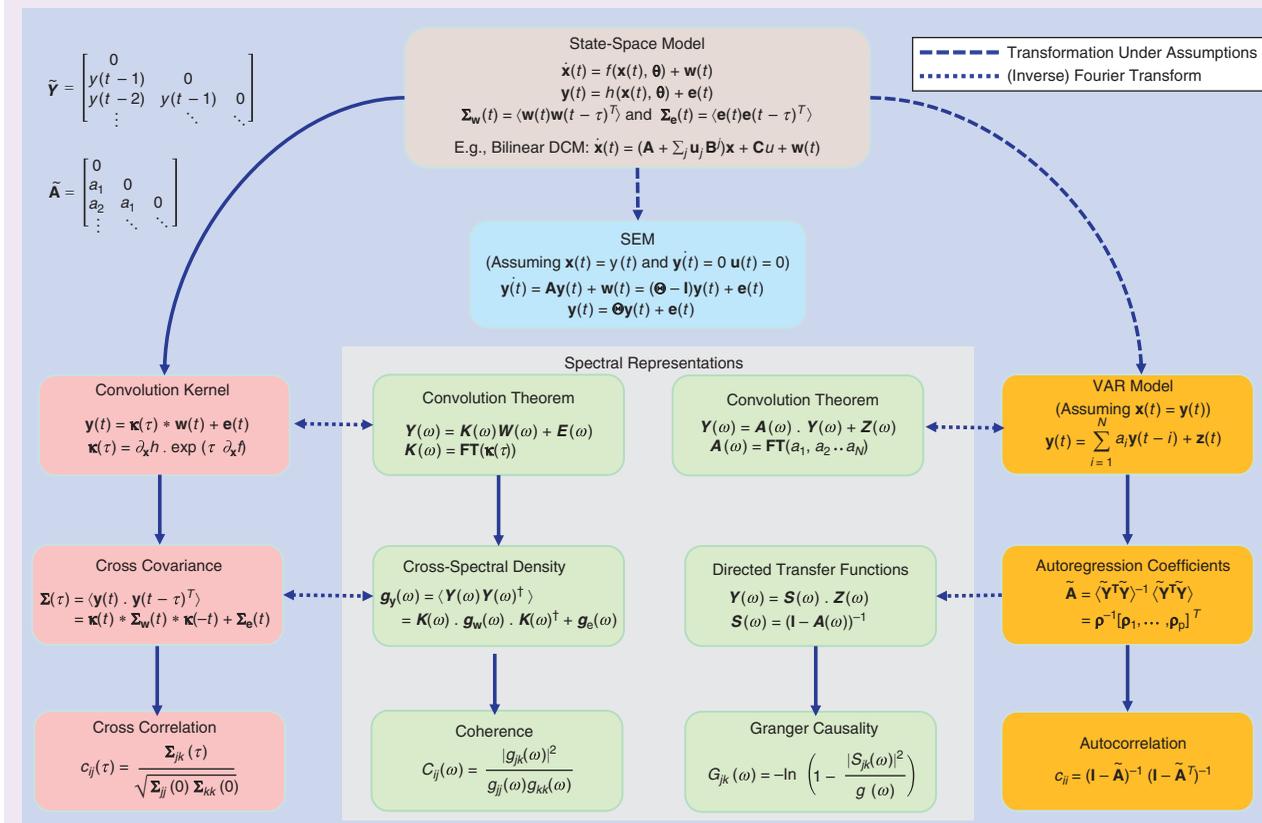


FIGURE S2. The relationship among different formulations of dependencies within a multivariate time series used in fMRI.

there is a distinction between (second-order) effects sizes in the upper row of dependency measures and their standardized equivalents in the lower row (Figure S2). For example, coherence is simply the amplitude of the cross-spectral density normalized by the autospectra of the two regions in question. Similarly, one can think of Granger causality as a standardized measure of the directed transfer function (normalized by the autospectra of the source region).

We also note another widely used measure of functional dependencies known as *mutual information* [57], which quantifies the shared information between two variables and can reflect both linear and nonlinear dependencies. For example, if two time series are independent, there is no shared information, and hence the mutual information is zero. Mutual information can be calculated relatively simply—under the assumption that time series are Gaussian—from coherence in the frequency domain as [58]–[60]

$$\vartheta_{ij} = \frac{1}{2\pi} \int_{\omega_1}^{\omega_2} \log(1 - C_{ij}(\omega)) d\omega, \quad (8)$$

where $C_{ij}(\omega)$ is the coherence (as defined in Figure S2) between the two time series i and j .

In summary, given a state-space model, one can predict or generate the functional connectivity that one would observe in terms of cross-covariance functions, complex cross spectra, or autoregression coefficients (where the latter can be derived in a straightforward way from the former using the Yule–Walker formulation). In principle, this means that one could use either the sampled cross-covariance functions or cross spectra as data features. It would also be possible to use the least-squares estimate of the autoregression coefficients—or, indeed, Granger causality—as data features to estimate the underlying effective connectivity. We describe such schemes in the next section.

Summary

In this section, we have tried to place different analyses of connectivity in relation to each other. The most prevalent approaches to effective connectivity are DCM, SEM, and Granger causality. We highlighted some of the implicit assumptions made when applying SEM and Granger causality to fMRI time series. Next we will focus on generative models of distributed brain responses and consider some of the exciting developments in this field.

Biophysical modeling of neuronal dynamics

Biophysical models of neuronal dynamics are usually used for one of two things: either to understand the emergent properties of neuronal systems or as observation models for measured neuronal responses. We discuss examples of both. In terms of emergent behaviors, we consider dynamics on structure [61]–[69] and how this behavior has been applied to characterizing autonomous or endogenous fluctuations in fMRI [70]–[73].

This section concludes with recent advances in DCM of direct neuronal interactions that support endogenous fluctuations. Some sections below are based on our previous review [10].

Intrinsic dynamics, criticality, and bifurcations

The use of resting-state fMRI [74], [75] or studies based on BOLD signal correlations while the brain is at rest are widespread [76]. These patterns are thought to reflect anatomical connectivity [77] and can be characterized in terms of remarkably reproducible spatial modes (resting-state or intrinsic networks). One of these modes recapitulates the pattern of deactivations observed across a range of activation studies (the default mode [78]). Resting-state fMRI studies show that even at rest, endogenous brain activity is self-organizing and highly

structured. The emerging picture is that endogenous fluctuations are a consequence of dynamics on anatomical connectivity structures with particular scale-invariant characteristics [70], [71], [79], [80]. These are well-studied and universal characteristics of complex systems and suggest that we may be able to understand the brain in terms of universal phenomena [81]. Universality is central to the hypothesis that the cerebral cortex is poised near a critical point where only one variable, a control parameter, determines the macroscopic behavior of the system [82], [83]. This is an important issue because systems near phase transitions show universal phenomena [84]–[88]. Near the critical point, correlations between neurons would occur across all scales, leading to optimized communication [89]. Experimental evidence for this notion has accumulated during the past decades, and power laws and scaling relationships have been found in human neuroimaging time series [90], [91]. However, it should be noted that with more attention on this new direction, there are a variety of distributions (e.g., stretched exponential, Rayleigh, double exponential, and lognormal) that are found in neurophysiological time series [26], [92], [93]. Hence, there may be a need to carefully disambiguate the causes of these heavy-tailed distributions found in the brain and behavior. From the dynamical system perspective, endogenous dynamics are thought to be generated by the dynamic instabilities that occur near bifurcations, that is, dynamics that accompany a loss of stability when certain control parameter(s) reach a critical value [26], [94]–[96]. The eigenmodes of neuronal (effective) connectivity that define the stability of the resting state give rise to scale-free fluctuations that emerge from the superposition of the few modes that decay slowly. These slowly fluctuating (unstable) modes have Lyapunov exponents that are close to zero. This occurs when systems approach transcritical bifurcations (or stochastic Hopf bifurcations when the eigenvalues are complex [97], [98] and show critical slowing [93]). Put simply, this means that the ensuing networks are defined by trajectories that have fixed points close to instability and that the neuronal fluctuations persist over longer timescales to generate the patterns responsible for the emergence of intrinsic brain networks. The

Spectral and stochastic DCMs furnish estimates of the effective connectivity that underlies intrinsic brain networks.

amplitudes of these eigenmodes or patterns correspond to the order parameters described in the “State-Space Modeling and Effective Connectivity” section. The (negative) inverse of the Lyapunov exponent corresponds to the characteristic time constant of each mode, where each mode with a small exponent (large time constant) corresponds to an intrinsic brain network or resting-state network.

Causal modeling of neuronal dynamics

The past decade has seen the introduction of graph theory to brain imaging. Graph theory provides an important formulation for understanding dynamics on structure. Developments in this area have progressed on two fronts: understanding connections between graphs and probability calculus and the use of probabilistic graphs to resolve causal interactions. The probabilistic graph framework goes beyond classical constructs by providing powerful symbolic machinery and notational convenience (e.g., the use of dependency graphs to resolve Simpson’s paradox; see “Simpson–Yule Paradox”).

Within this enterprise, one can differentiate at least two streams of work: one based on Bayesian dependency graphs or graphical models called structural causal modeling [99] and the other based on causal influences over time, which we consider under DCM. Structural causal modeling originated with SEM [47] and uses graphical models (Bayesian dependency graphs or Bayes nets) in which direct causal links are encoded by directed edges. These tools have been largely developed by Pearl [22] and are closely related to the ideas in [100]–[102]. An essential part of network discovery in structural causal modeling is the concept of intervention: eliminating connections in the graph and setting certain nodes to given values. Structural causal modeling lends a powerful and easy-to-use graphical method to show that a particular model specification identifies a causal effect of interest. Moreover, the results derived from structural causal modeling do not require specific distributional or functional assumptions, such as multivariate normality, linear relationships, and so on. However, it is not the most suitable framework to understand coupled dynamical systems because it is limited in certain respects. Crucially, it deals only with conditional independencies in DAGs. This is problematic because brains perform computations on a directed and cyclic graph. Every brain region is connected reciprocally (at least polysynaptically), and every computational theory of brain function rests on some form of reciprocal or reentrant message passing. Another drawback is that the causal calculus of structural causal modeling ignores time. Pearl argued that a causal model should rest on functional relationships between variables. However, these functional relationships cannot deal with (cyclic) feedback loops. Pearl [14] argued for DCMs when attempting to identify hysteresis effects, where causal influences depend on the history of the system. Interestingly, the DAG restriction can be finessed by considering dynamics and temporal precedence within

Neural pathways are flexible, adaptable, connected, and moldable by changes in our environment or by injury or disease.

structural causal modeling. This is because the arrow of time can be used to convert a directed cyclic graph into an acyclic graph when the nodes are deployed over successive time points. This leads to SEM with time-lagged data and related autoregression models, such as those employed by Granger causality described previously. As established in the previous section, these can be regarded as discrete time formulations of DCMs in continuous time.

Structural and dynamic causal modeling

As already established, in relation to the modeling of fMRI time series, DCM refers to the (Bayesian) inversion and comparison of models that cause observed data. These models are usually state-space models expressed as (ordinary, stochastic, or random) differential equations that govern the motion of hidden neurophysiological states. These models are generally equipped with an observer function that maps from hidden states to observed signals [see (1)]. The basic idea behind DCM is to formulate one or more models of how data are caused in terms of a network of distributed sources. These sources talk to each other through parameterized connections and influence the dynamics of hidden states that are intrinsic to each source. Model inversion provides estimates of their parameters and the model evidence.

We have introduced DCM for fMRI using a simple state-space model based on a bilinear approximation (extensions to, for example, nonlinear [103] and two-state [104] DCM, among others, are also available and are in use) to the underlying equations of motion that couple neuronal states in different brain regions [32]. Most DCMs consider point sources for both fMRI and EEG/MEG data (cf. equivalent current dipoles) and are formally equivalent to the graphical models used in structural causal modeling. However, in DCM, they are used as explicit generative models of observed responses. Inference on the coupling within and between nodes (brain regions) is generally based on perturbing the system and trying to explain the observed responses by inverting the model. This inversion furnishes posterior or conditional probability distributions over unknown parameters (e.g., effective connectivity) and the model evidence for model comparison [105]. The power of the Bayesian model comparison in the context of DCM has become increasingly evident. This now represents one of the most important applications of DCM and allows different hypotheses to be tested, where each DCM corresponds to a specific hypothesis about functional brain architectures [106]–[112]. DCM has been used mostly for (task-based) fMRI and electrophysiological dynamics (EEG/MEG/LFPs), but the most recent advances have focused on the modeling of intrinsic brain networks in the absence of exogenous influence, known as resting-state fMRI [74]. In the remainder of this section, we briefly review these developments and discuss these new mathematical models. We also showcase some of their clinical applications to neurodegenerative diseases, such as Parkinson’s disease.

Dynamic casual modeling of intrinsic networks

There has been an explosion of research examining spontaneous fluctuations in fMRI signals (Figure 2). These fluctuations can be attributed to spontaneous neuronal activity, which is usually ignored in deterministic models of responses to (designed) experimental inputs. Deterministic DCMs are cast as multiple-input, multiple-output systems, in which exogenous inputs perturb the brain to produce an observed BOLD response. In the absence of external inputs, as in the case of resting-state fMRI, neuronal networks are driven by activity that is internal to the system [113]. The generative model for resting-state fMRI time series has the same form as (3) but discounts exogenous modulatory input. It should be noted that we can still include exogenous (or experimental) inputs, $u(t)$, in our model. These inputs drive the hidden states and are usually set to zero in resting-state models. It is perfectly possible to have external (nonmodulatory) stimuli, as in the case of conventional functional neuroimaging studies. For example, in [114] we used an attention-to-visual-motion paradigm to illustrate this point. Figure 3 provides a schematic of the resulting stochastic DCM. In contrast to the previous section, we adopt a generalized framework in which state noise $w(t)$ and observation noise $e(t)$ are analytic (i.e., non-Markovian). This simply means that generalized motion of the state noise $\mathbf{w}(t) = [\mathbf{w}(t), \dot{\mathbf{w}}(t), \ddot{\mathbf{w}}(t) \dots]$ is well defined in terms of its higher-order statistics. Similarly, the observation noise $\tilde{\mathbf{e}}(t) = [\mathbf{e}(t), \dot{\mathbf{e}}(t), \ddot{\mathbf{e}}(t) \dots]$ has a well-defined covariance (for a more detailed discussion, see [115]). Consequently, the stochastic part of the generative model in (1) can be conveniently parameterized in terms of its precision (inverse covariance). This allows us to cast (1) as a random differential equation instead of a stochastic differential equation, hence eschewing Itô calculus [34], [116]. Interested readers will find a theoretical motivation for using analytic state noise in [34]. Under linearity assumptions, (1) can be written compactly in generalized coordinates of motion as

$$\begin{aligned} \mathbf{D}\tilde{\mathbf{x}}(t) &= \tilde{\mathbf{f}}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, \boldsymbol{\theta}) + \tilde{\mathbf{w}}(t) \\ \tilde{\mathbf{y}}(t) &= \tilde{\mathbf{h}}(\tilde{\mathbf{x}}, \boldsymbol{\theta}) + \tilde{\mathbf{e}}(t), \end{aligned} \quad (9)$$

where \mathbf{D} is the block diagonal temporal derivative operator, such that the vectors of generalized coordinates of motion are shifted as we move from lower orders of motion to higher orders [115]. For resting-state activity, (9) takes a very simple linear form:

$$\mathbf{D}\tilde{\mathbf{x}}(t) = \mathbf{A}\tilde{\mathbf{x}}(t) + \mathbf{C}\tilde{\mathbf{u}}(t) + \tilde{\mathbf{v}}(t). \quad (10)$$

This is an instance of a linear dynamical system with quasideterministic behavior [117], [118]. Put simply, the linear dynamical system described by (10) is insensitive to the initial conditions. For this reason, it can exhibit only a limited repertoire of behavior: linear systems can contain closed orbits, but they will not be isolated; no limit cycles—either stable or unstable—can exist, which precludes chaotic behavior. Technically speaking, if $\boldsymbol{\lambda}$ represents the eigenvalues of the Jacobian $\partial_{\tilde{\mathbf{x}}} \mathbf{f} = \mathbf{A}$, that is, $\boldsymbol{\lambda} = \mathbf{v}^\dagger \mathbf{A} \mathbf{v}$, where \dagger denotes the

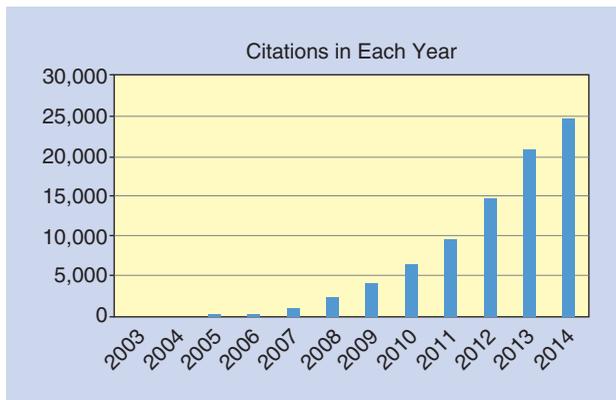


FIGURE 2. Citation rates for resting-state fMRI studies. These citations were identified by searching for “fMRI*” and “resting state.” (Source: Web of Science.)

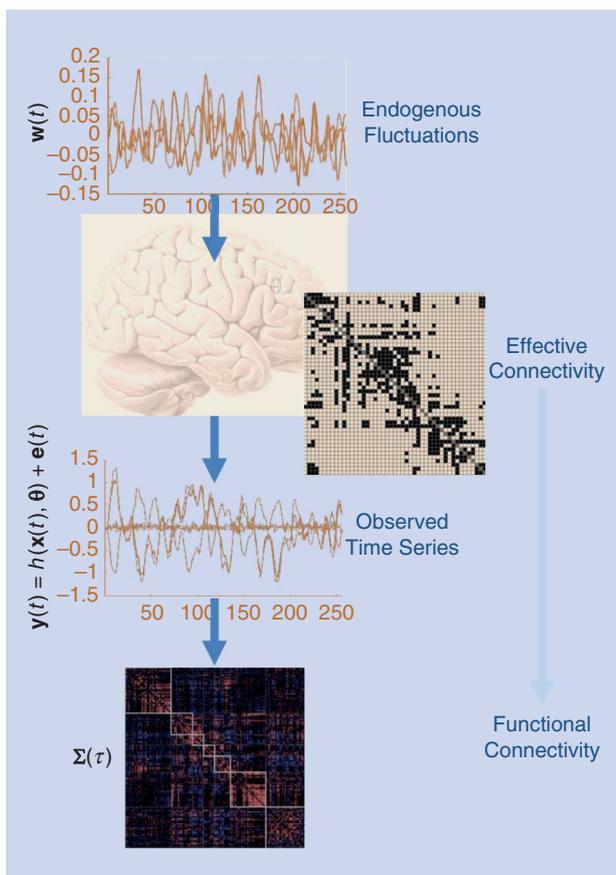


FIGURE 3. This schematic illustrates the forward (dynamic causal) model for modeling intrinsic or endogenous fluctuations. The endogenous fluctuations (state noise) are the driving input to the state-space model of effective connectivity, which is a function of the current neuronal states $\mathbf{x}(t)$ and the connectivity parameters $\boldsymbol{\theta}$ that define the underlying structure or functional architecture of the model and the random fluctuations $\mathbf{w}(t)$. The driving fluctuations cause change in neural activity that can, in turn, be observed using the observer function h after addition of observation noise $\mathbf{e}(t)$. The associated functional connectivity (e.g., cross-covariance function) can be calculated easily from this forward or generative model (see Figure S2 in “Measures of Connectivity”) for any given parameters. Note that the effective connectivity matrix shown is actually a structural connectivity matrix of the famous macaque/CoCoMac. We use it here as a schematic for effective connectivity.

generalized inverse, then the Lyapunov exponents $\Re(\lambda)$ of this linear dynamical system will always be negative. In general, the Jacobian is not symmetrical (causal effects are asymmetric); the modes and eigenvalues take complex values. See [119] for a detailed treatment of the special case of symmetrical connectivity, in which the eigenmodes of functional and effective connectivity become the same. It is worth noting that these eigenmodes are also closely related to (group) independent component analysis (ICA) except with a rotation based on higher-order statistics (for details, see [120]).

There are currently two schemes to invert models of the form (9). They differ in what data features they use for the parameter estimation. The first inverts data in the time domain, and the model is used to predict the time series per se.

This is referred to as *stochastic DCM* [116]. The second approach makes predictions in the frequency domain and is based on fitting second-order data features, such as cross spectra. This is referred to as *spectral DCM* [114], [121]. We briefly review both schemes and illustrate their clinical applications. For a schematic illustration of DCM of intrinsic dynamics, see Figure 4. Figure 5 presents a comparison of the two schemes.

Stochastic dynamic causal models

Stochastic DCM entails inverting a model of the form given by (10) in the time domain, which includes state noise. This requires estimation of not only the model parameters (and any hyperparameters that parameterize the precision of generalized

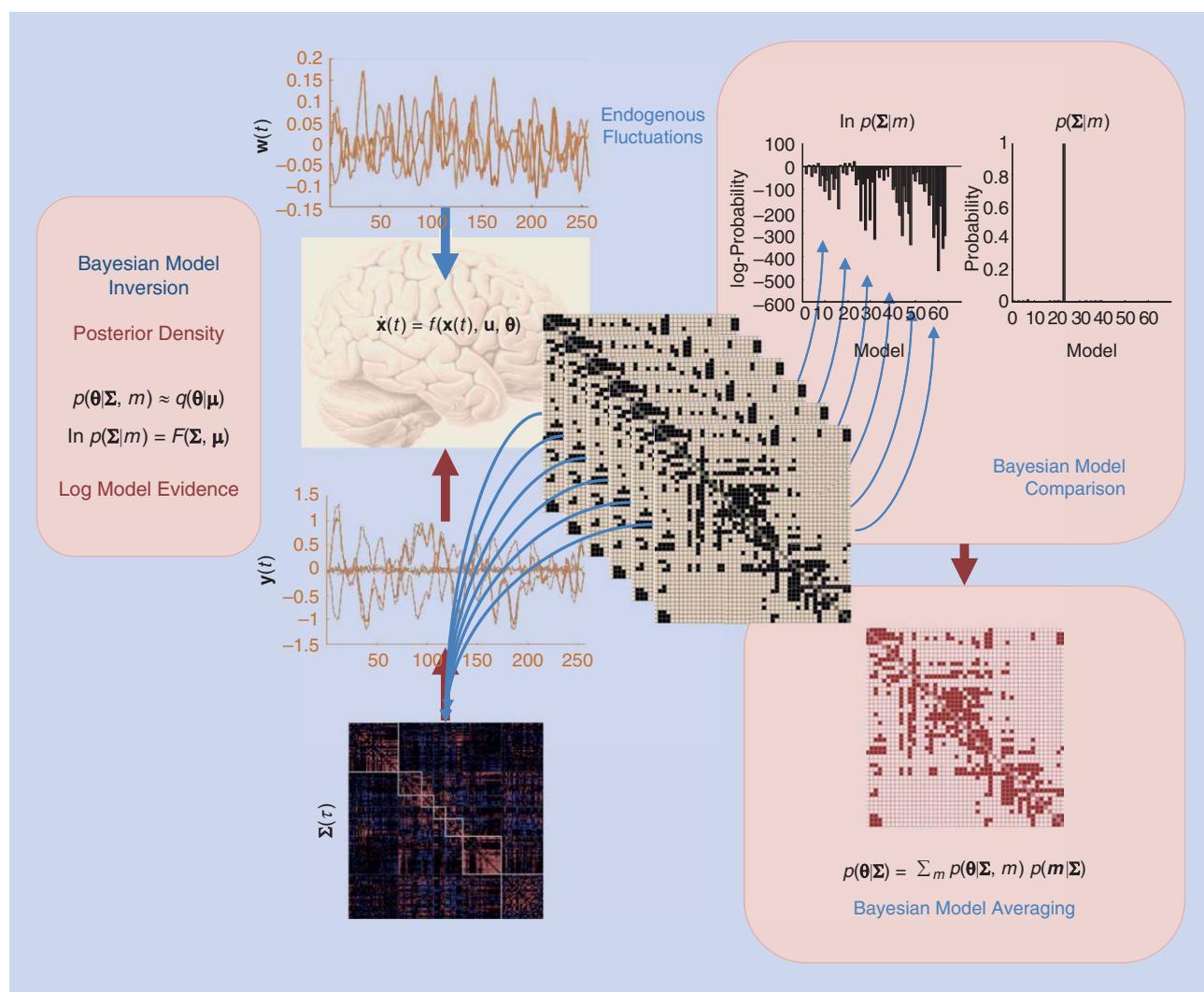


FIGURE 4. This schematic shows a DCM that embodies the best effective connectivity—identified using Bayesian model inversion (top left panel)—among hidden neuronal states that explains the observed functional connectivity, $\Sigma(t)$, among hemodynamic responses. This explanation is possible because the cross spectra contain all the information about (second-order) statistical dependencies among regional dynamics. Bayesian model inversion furnishes posterior estimates for the parameters of each model and provides the associated log model evidence in terms of a variational free-energy bound. Because the mapping from functional connectivity to effective connectivity is not objective (there may be many combinations of effective connectivity parameters that induce the same functional connectivity), one can use a Bayesian model comparison (top right panel) to score competing models. The model with the highest model evidence can then be selected. Alternatively, one can use Bayesian model averaging to average all possible models (bottom right panel).

random fluctuations) but also the hidden states, which become random (probabilistic) variables. The unknown quantities to be estimated under a stochastic DCM are $\boldsymbol{\psi} = \{\tilde{\mathbf{x}}(t), \boldsymbol{\theta}, \boldsymbol{\sigma}\}$, where $\boldsymbol{\sigma}$ refers to any hyperparameters describing random fluctuations. In terms of temporal characteristics, the hidden states are time-variant, whereas the model parameters (and hyperparameters) are time-invariant.

There are various variational schemes in the literature that can invert such models, for example, dynamic expectation maximization (DEM) [122] and generalized filtering (GF) [34]. There is a subtle but important distinction between DEM and GF. DEM calls on the mean field approximation described above, that is, it assumes $q(\boldsymbol{\psi}) = q(\tilde{\mathbf{x}}(t))q(\boldsymbol{\theta})q(\boldsymbol{\sigma})$, whereas GF, as the name suggests, is more general in that it does not make this assumption. However, both schemes assume a fixed-form Gaussian distribution for the approximate conditional

posterior densities (the Laplace approximation). GF considers all unknown quantities to be conditionally dependent variables, that is, $q(\boldsymbol{\psi}) = q(\tilde{\mathbf{x}}, \boldsymbol{\theta}, \boldsymbol{\sigma})$, and produces time-dependent conditional densities for all unknown quantities. The time-invariant parameters and hyperparameters are cast as time-variant with the prior constraint that their temporal variation is small. In brief, this online scheme assimilates log evidence at each time point in the form of variational free energy and provides time-dependent conditional densities for all unknown quantities. This is in contrast to schemes such as DEM (or deterministic model inversion using variational Laplace) with mean field approximations, which assimilates all the data before computing the free energy.

“Effective Connectivity in Parkinson’s Disease” shows an exemplar data analysis reported in [123] that used stochastic DCM to quantify effective connectivity changes in

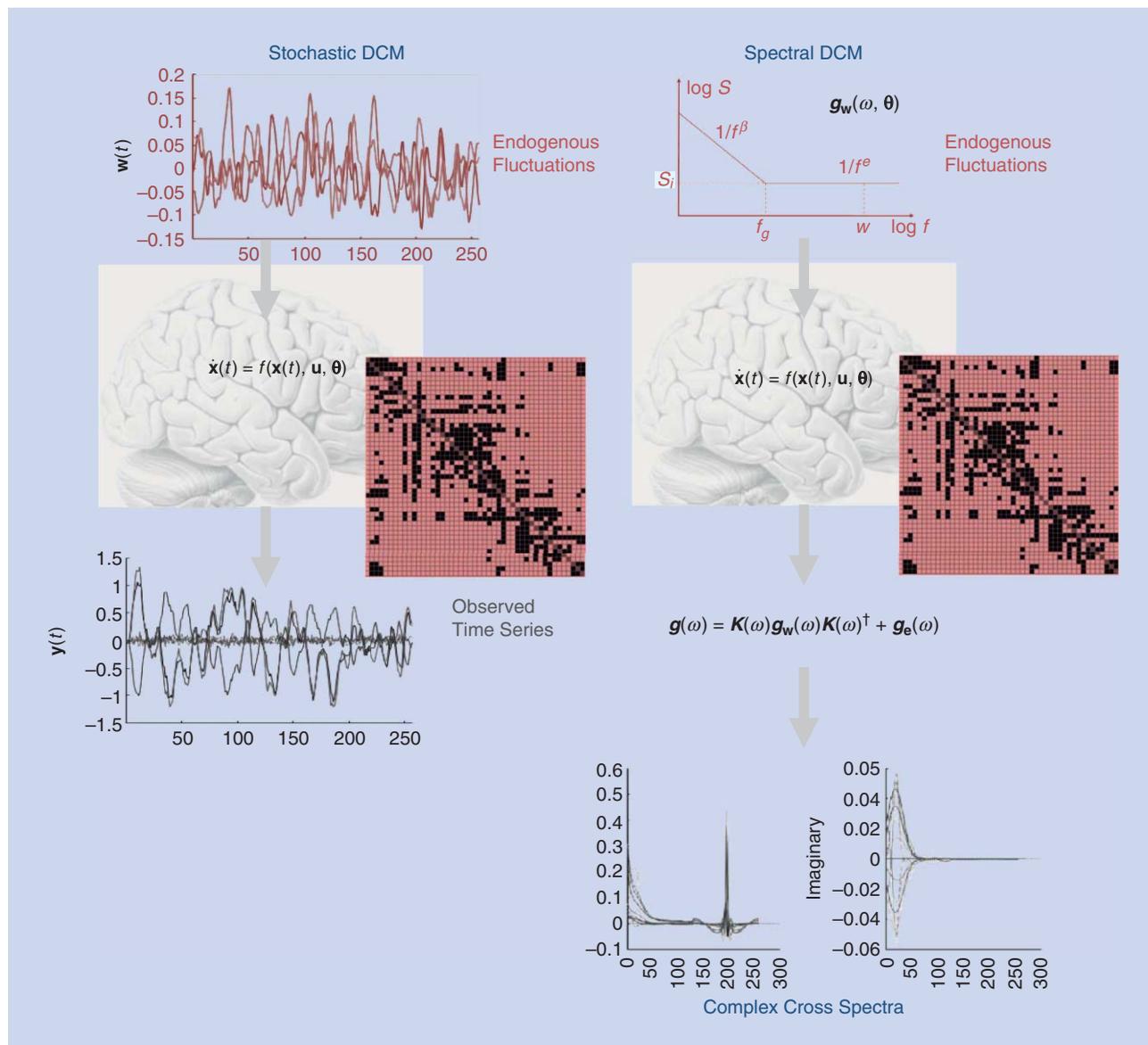


FIGURE 5. A schematic illustrating the distinction between stochastic and spectral DCM. See the “Biophysical Modeling of Neuronal Dynamics” section for a detailed description of how these schemes are used to model intrinsic network dynamics.

Effective Connectivity in Parkinson's Disease

Given the marked clinical effect of subthalamic nucleus (STN) deep brain stimulation (DBS) in patients with Parkinson's disease, Kahan et al. [123] used stochastic dynamic causal modeling (DCM) to estimate the coupling between key nodes of the basal ganglia network and to study whether this coupling was changed by DBS. In Figure S3(a), a network was specified based on human and animal literature, and priors were placed on the nature of the coupling (excitatory or inhibitory) based on the neurochemical systems known to mediate neuronal connections. The literature-based anatomical model of the motor cortico-striato-thalamic loop was further simplified by removing the pallidal nodes and summarizing polysynaptic connections [thick arrows joining the putamen (Put), STN, and thalamus (Tha)]. Red arrows indicate excitatory coupling, and blue arrows indicate inhibitory coupling.

Placing priors on the direction of coupling was enabled using the two-state DCM (left). In (b), it is shown that model inversion yielded coupling parameters on and off DBS, demonstrating significant DBS-related changes in extrinsic (between-node) coupling throughout the network. Paired t-tests revealed significant differences between extrinsic coupling on and off stimulation. Corticostriatal, direct pathway, and thalamocortical connections were potentiated by DBS, whereas STN afferents (lower panel) and efferents (upper panel) were attenuated. Note the difference in scale between the upper and lower panels. This is because the STN was modeled as a hidden node that was not measured with fMRI. Using a series of regression models, (c) shows the modulatory effects of DBS on connectivity to predict the clinical improvements seen in the patient cohort. (See [123] for more details.)

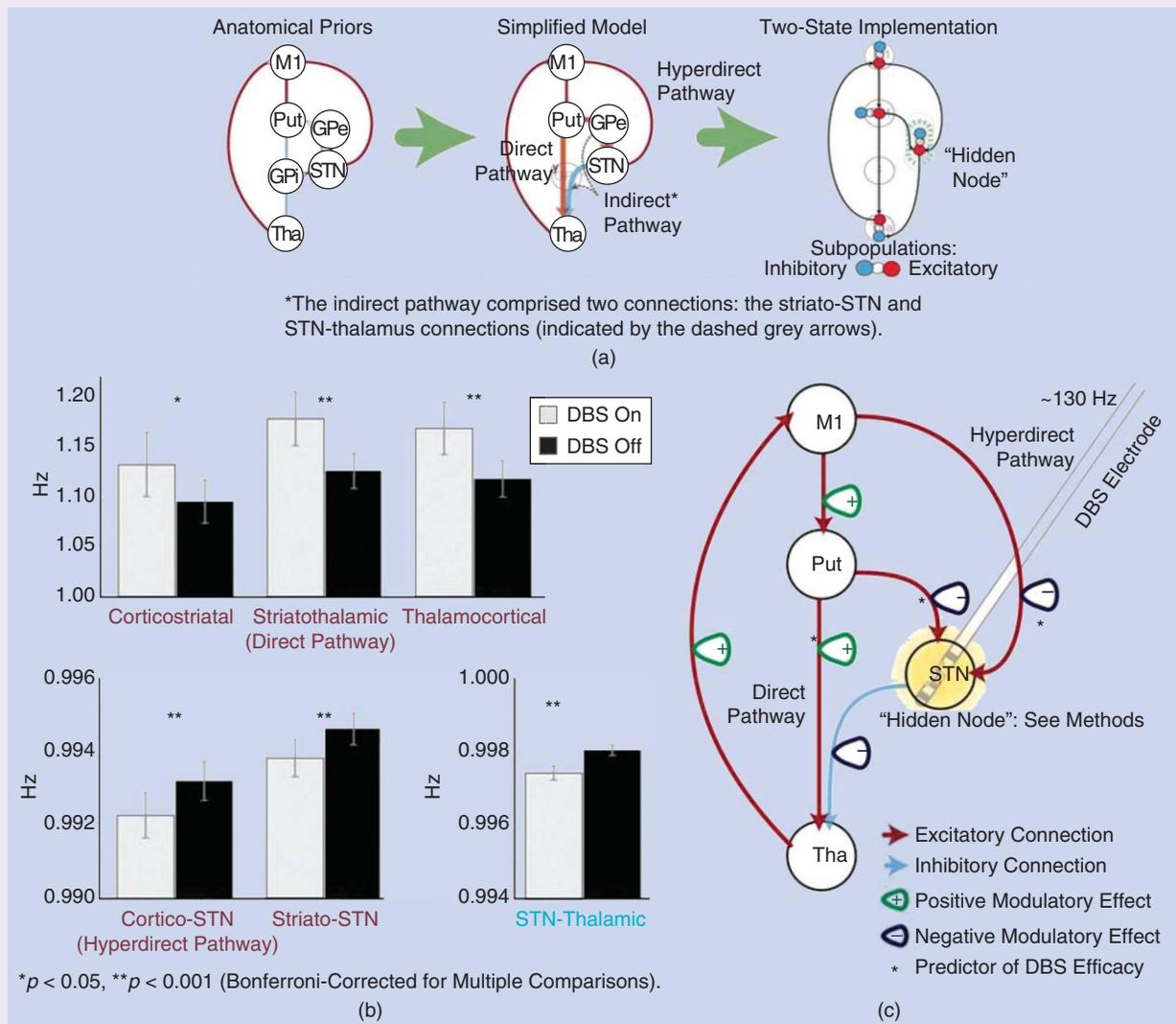


FIGURE S3. The summary of data analysis reported in [123] using stochastic DCM in Parkinson's disease under DBS.

Parkinson's disease. Depleted of dopamine, the dynamics of the Parkinsonian brain impact on both action and resting motor activity. Deep brain stimulation (DBS) has become an established means of managing these symptoms, although its mechanisms of action remain unclear. Using stochastic DCM, Kahan et al. [123] modeled the effective connectivity underlying low-frequency BOLD fluctuations in the resting Parkinsonian motor network. They were particularly interested in the distributed effects of DBS on cortico-subcortical connections. Specifically, they showed (see Figure S3 in "Effective Connectivity in Parkinson's Disease") that subthalamic nucleus (STN) DBS modulates all major components of the motor cortico-striato-thalamo-cortical loop, including the corticostriatal, thalamocortical, direct, and indirect basal ganglia pathways and the hyperdirect STN projections. The strength of effective STN afferents and efferents was reduced by stimulation, whereas corticostriatal, thalamocortical, and direct pathways were strengthened. Remarkably, regression analysis revealed that the hyperdirect, direct, and basal ganglia afferents to the STN predicted clinical status and therapeutic response to DBS; however, suppression of the sensitivity of the STN to its hyperdirect afferents by DBS may subvert the clinical efficacy of DBS. These findings highlight the distributed effects of stimulation on the resting motor network and provide a framework for analyzing effective connectivity in resting-state functional MRI with strong a priori hypotheses.

Spectral dynamic causal models

Although the stochastic models in (10) and their inversion in the time domain provide a useful means to estimate effective connectivity, they also entail the estimation of hidden states. This poses a difficult inverse problem that is computationally demanding, especially when the number of hidden states becomes large. To finesse this problem, a DCM based on a deterministic model that generates predicted cross spectra was explored [114], [121]. This scheme provides a constrained inversion of the stochastic model by parameterizing the spectral density neuronal fluctuations. This parameterization also provides an opportunity to compare parameters encoding neuronal fluctuations among groups. The parameterization of endogenous fluctuations means that the states are no longer probabilistic; hence, the inversion scheme is significantly simpler, requiring estimation of only the parameters (and hyperparameters) of the model. The ensuing model inversion in the spectral domain is similar in spirit to previous approaches described in [26], [98], and [124]. Put simply, although GF estimates time-dependent fluctuations in neuronal states producing observed data, spectral DCM simply estimates the time-invariant parameters of their cross spectra. Effectively, this is achieved by replacing the original time series with their second-order statistics (i.e., cross spectra). This means that instead of

estimating time-varying hidden states, we are estimating their covariance. In turn, this means that we need to estimate the covariance of the random fluctuations using a scale-free (power law) form for the state noise (resp. observation noise) that can be motivated from previous work on neuronal activity [125]–[127]:

$$\begin{aligned} \mathbf{g}_w(\omega, \boldsymbol{\theta}) &= \boldsymbol{\alpha}_w \omega^{-\beta_w} \\ \mathbf{g}_e(\omega, \boldsymbol{\theta}) &= \boldsymbol{\alpha}_e \omega^{-\beta_e}. \end{aligned} \quad (11)$$

Here, $\mathbf{g}_x(\omega) = X(\omega)X(\omega)^\dagger$ represents the complex cross spectra, where $X(\omega)$ is the Fourier transform of $\mathbf{x}(t)$, $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\} \subset \boldsymbol{\theta}$ are the parameters controlling the amplitudes and exponents of the spectral density of the neural fluctuations, and $\omega = 2\pi f$ is the angular frequency. This models neuronal noise with generic $1/f^\gamma$ spectra that characterize fluctuations in systems that are at nonequilibrium steady state. A linear scaling regime of the spectral density in double logarithmic coordinates—implicit in (11)—is not by itself indicative of a scale-free critical process unless γ is less than 1.5 Hz (and the regime scales over

several orders of magnitude). For the human EEG, this is generally not the case: above 10 Hz, $\gamma = 2.5$, and above 70 Hz, γ is usually greater than 3.5, which is consistent with a Poisson process (see [128] and [129]). However, at low frequencies (less than 1.5 Hz), the slope is more shallow, and it is likely that the amplitude or power envelopes of faster frequencies are scale-free [130, 131] or another heavy-tailed distribution [132]. Using the model parameters $\boldsymbol{\theta} \supseteq \{\mathbf{A}, \mathbf{C}, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$, one can simply generate the expected cross spectra as follows:

$$\begin{aligned} \mathbf{y}(t) &= \boldsymbol{\kappa}(t) \otimes \mathbf{w}(t) + \mathbf{e}(t) \\ \boldsymbol{\kappa}(t) &= \partial_{\mathbf{x}} g \exp(t \partial_{\mathbf{x}} f) \\ \mathbf{g}_y(\omega, \boldsymbol{\theta}) &= |\mathbf{K}(\omega)|^2 \mathbf{g}_w(\omega, \boldsymbol{\theta}) + \mathbf{g}_e(\omega, \boldsymbol{\theta}), \end{aligned} \quad (12)$$

where $\mathbf{K}(\omega)$ is the Fourier transform of the system's (first-order) Volterra kernels $\boldsymbol{\kappa}(t)$, which are a function of the Jacobian or effective connectivity (see Figure S2). The unknown quantities $\boldsymbol{\psi} = \{\boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\sigma}\}$ of this deterministic model can now be estimated using standard variational Laplace [133]. The resulting inversion provides the free energy bound on the log-evidence $\log p(\mathbf{g}_y(\omega)|m)$ and approximate conditional densities $q(\boldsymbol{\psi}) \approx p(\boldsymbol{\psi}|\mathbf{g}(\omega), m)$. Here $\mathbf{g}_y(\omega)$ represents the predicted cross spectra that can be estimated, for example, using an autoregressive model.

An example from aging

Finally, in "Aging and Spectral DCM," we show an example from recent work on aging [134] that used spectral DCM. Well-being across the lifespan depends on the preservation of cognitive function. It was hypothesized that successful cognitive aging is determined by the connectivity within and between large-scale brain networks at rest. Spectral DCM

Spectral and stochastic DCMs furnish estimates of the effective connectivity that underlies intrinsic brain networks.

Aging and Spectral DCM

In Figure S4(a), spatial distribution of three independent components using group independent component analysis (ICA) ($n = 602$) are identified as the default-mode network (DMN) (in blue), the dorsal attention network (DAN) (in red), the salience network (SN) (in yellow), and the peaks of their corresponding nodes (green circles). Temporal correlation is between the first eigenvariates of the ensuing time series across all nodes and networks. Coefficients for how well effective connectivity (white), neuronal (green), and hemodynamic (red) parameters predict age are shown in (b), and dynamic causal modeling parameters with bars (95% confidence intervals) that exclude zero are considered as significant predictors. A between-network canonical variate analysis is shown in (c). More specifically, shown is a heliograph of variate

loadings for the first canonical variate, where the relative sizes of correlations are indicated by the relative length of the bars (the dark is positive, and the white is negative). These reflect the statistical relationship between variables of effective connectivity (connectivity profile) and cognitive performance (cognitive profile) ($r = 0.440$, $p < 0.001$). Variables with low contribution ($r < 0.3$) are shown as bars with a dashed outline. Half-maximum strength of a correlation is indicated by dashed rings (outer, $r = +0.5$; inner, $r = -0.5$). The corresponding bivariate canonical correlations for three age groups are shown in (d). The relationships between connectivity and cognitive profiles are more pronounced for older patients, suggesting that performance in older adults reflects a preserved connectivity.

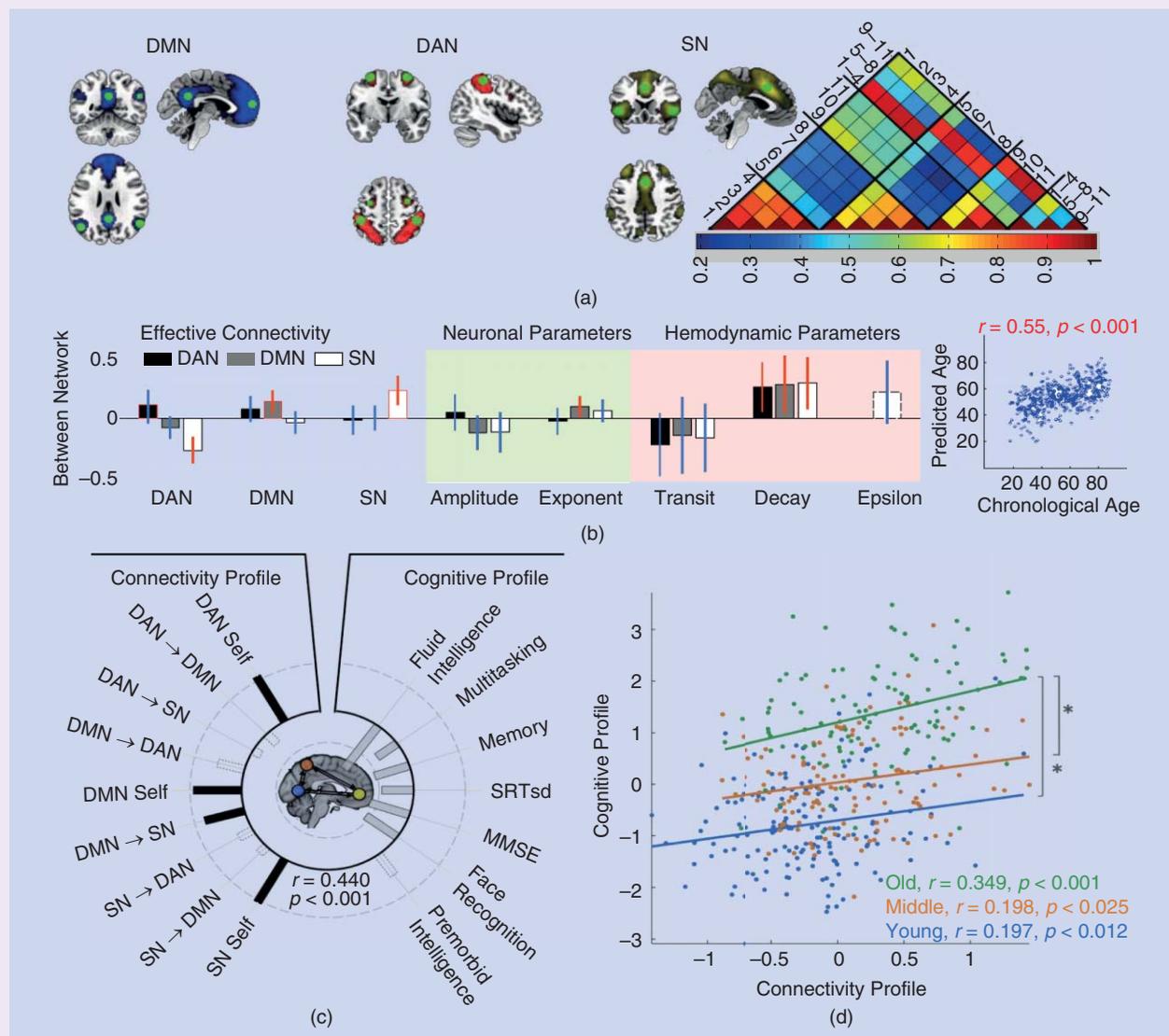


FIGURE S4. The summary of between-network connectivity changes over the adult lifespan.

was used to explain the spectral characteristics of resting-state fMRI data from 602 healthy adults in a cohort across ages 18–88 (www.cam-can.org). The location of the key cortical regions in each network was identified by spatial ICA using group ICA [120] to extract 20 low-dimensional components. The three well-established functional networks, the salience network (SN), dorsal attention network (DAN), and default-mode network (DMN), were then identified by spatial matching to preexisting templates [135]. Effective connectivity was assessed within and between these three key large-scale networks although, for brevity, we have included more interesting results only for the between-network connectivity in this review. In brief, a two-step process is used in which ICA identifies linearly coherent networks, and the (potentially nonlinear) relationship among these networks is then tested within a causal modeling framework using spectral DCM. This approach has been used several times in both task-based and resting-state fMRI data [136]–[138].

Using multiple linear regression, it was found that about 30% of age variance can be predicted ($r = 0.544$, $p < 0.001$) by 1) increased inhibitory self-connections in SN and DMN, 2) decreased effective connectivity from DAN to SN, and 3) increased hemodynamic decay times for all networks [Figure S4(b)]. Subsequently, a classical multivariate test (canonical variate analysis) was used to determine to what degree the DCM parameters predict cognitive performance, shown in Figure S4(c). For between-network analysis, the corresponding canonical vector suggested that high performance across a range of cognitive tasks [high scores of general intelligence (Cattell), face processing (Benton Faces), memory (story recall), multitasking (Hotel), and response consistency (inverse of response variability on simple motor task)] was associated with less self-inhibition of the networks and a smaller influence of the DMN on SN ($r = 0.447$, $p < 0.001$). In other words, about 20% of the variance in performance—across a range of cognitive tasks studied—could be predicted from changes in effective connectivity between networks. To further investigate whether the relationship between cognitive performance and connectivity was age-dependent, moderation analysis was used. It was found that the interaction between age and connectivity values (age \times connectivity profile) predicted a significant proportion of variance in cognitive performance ($T(398) = 3.115$, p (one-tailed) < 0.001). The direction of the interaction was such that increasing age strengthened the relationship between cognitive and connectivity profiles. This is shown in Figure S4(d), where the relationship between cognitive performance and connectivity profile becomes stronger for older age groups. This is an interesting study because it used spectral DCM to dissociate neuronal from vascular components of the fMRI signal to find age-dependent and behaviorally relevant differences in resting-state effective connectivity between large-scale brain networks. Taken

Resting-state fMRI studies show that intrinsic brain activity is self-organizing and highly structured.

together, the results suggest that maintaining a healthy resting-state connectivity becomes increasingly important for older adults to maintain high levels of domain-general cognitive function and may play a critical role in the mechanisms of healthy cognitive aging.

Summary

Both spectral and stochastic DCMs furnish estimates of the effective connectivity that underlies intrinsic brain networks.

These estimates are based on BOLD data acquired at rest using different inversion schemes. We suppose that these resting-state networks emerge from the dynamic instabilities and critical slowing near transcritical bifurcations. In this setting, neuronal activity is modeled with random differential equations, which can be estimated using stochastic inversion schemes

(such as GF in stochastic DCM) or by deterministic schemes modeling observed functional connectivity (specifically, the cross-spectral densities modeled by spectral DCM).

Discussion

The limitations and challenges of DCM and the implicit scoring of large numbers of models have been addressed in a number of critical reviews (e.g., [139] and [140]). Their key conclusions highlight several issues. First, although the modeling assumptions underlying DCM are motivated by neuroanatomical and neurophysiological constraints, their plausibility is difficult to fully establish. For example, in DCM for fMRI, physiological details of the neurovascular coupling are potentially important. Many DCMs neglect the potential influence of inhibitory activity on the hemodynamic response and call on a simplistic account of the metabolic cascade that relates synaptic activity and neuronal metabolism to the vasodilatation. In principle, these are issues that can be resolved using Bayesian model comparison. In other words, if a more complex and complete model is supported by the data, one can always optimize the DCM. Examples of this include recent trends toward more detailed physiological modeling. For example, several extensions are proposed in [141], such as an adaptive two-state neuronal model that accounts for a wide range of neuronal time courses during stimulation and post-stimulus deactivation, a neurovascular coupling model that links neuronal activity to blood flow in a strictly feedforward fashion, and a balloon model that can account for a vascular uncoupling between blood flow and blood volume due to viscoelastic properties of venous blood vessels.

There are also questions about the robustness of the statistical (approximate Bayesian) inference techniques employed in DCM. For example, it has been argued that 1) the number of parameters and the complexity of the models preclude robust parameter estimation [140], [142]; 2) Bayesian model comparison cannot compare DCMs in the sense that it cannot falsify them; and 3) selecting a model based on the model evidence does not ensure that it will generalize. All of these concerns

stem from frequentist thinking and are dissolved within a Bayesian framework (see [139] for a detailed discussion). There are also several well-founded technical concerns about the variational Bayes (VB) schemes employed in DCM. For example, the objective function based on the free-energy functional is prone to local maxima that can result in inconsistent parameter estimations and model comparisons (e.g., across trials or subjects). There are several experimental studies (e.g., [143]–[147]) that have addressed the reproducibility of DCM and provide reassuring experimental validation. There is an issue of overconfidence usually associated with VB schemes due to the potentially biased inference that results from mean field and Laplace approximations to the posterior density. This issue has been addressed by simulation studies that compare the results of VB with standard (e.g., Gibb's) sampling methods. The failures of approximate Bayesian inference are usually mitigated by formulating the inversion problem in a way that eschews brittle nonlinearities.

Given these issues, one obvious alternative is to use either exact inference schemes, such as Markov chain Monte Carlo (MCMC) or nonparametric methods based on Gaussian processes. Both have recently been explored for inverting Bayesian hierarchical models. For example, Gaussian process optimization was used for model inversion in [148], several gradient-free MCMC schemes (e.g., for random walk-based Hasting's sampling, adaptive MCMC sampling, and population-based MCMC sampling) were explored in [149], and more robust gradient-based MCMC schemes (e.g., for Hamiltonian and Langevin MCMC sampling) were extensively studied in [150]. However, these alternative and promising inference methods are still in an early phase of development and validation phase and will require exhaustive experimental studies to establish validity.

Clearly, most of these issues transcend DCM per se and speak to the challenges facing any modeling initiative that has to contend with big data and a large model or hypothesis space. These challenges have focused recent research on contextualizing the inversion of models of single subjects using (empirical or hierarchical) Bayesian models that distinguish between within- and between-subject effects on one hand and the scoring of large model spaces with techniques such as Bayesian model reduction on the other. This is an active research field with developments nearly every month.

In conclusion, we have used several distinctions to review the history and modeling of macroscopic brain connectivity. We started with the distinction between functional segregation and integration. In functional integration, we considered the key distinction between functional and effective connectivity and their relationship to underlying models of distributed processing. In effective connectivity, we looked at structural and dynamic causal modeling while highlighting recent advances in the DCM of resting-state fMRI data.

We close with a few words on recent large-scale projects in neurosciences, for example, the American BRAIN Initiative and the European Human Brain Project. These initiatives reflect an increasing appreciation of the impor-

tance of neuroscience and the challenges of understanding how brains work. Furthermore, they represent initiatives that exploit remarkable advances in computer science and neuroimaging at many different scales (from the molecular to multisubject) and the modeling (and mining) of the resulting data. The experience of the systems neuroscience community, with the big data obtained from neuroimaging, is reflected in this review. This experience highlights the importance of formal models of how data are generated and the computational schemes used to evaluate and invert these models. We are just embarking on a difficult journey to uncover the governing principles of how brains work and their functional (computational) architectures. Perhaps it is fitting to end with an encouraging quote from Abdus Salam (recipient of the Nobel Prize in Physics 1979): "Nature is not economical of structures—only of principles."

Acknowledgments

This work was funded by the Wellcome Trust and the CHDI Foundation. We thank our colleagues at the Wellcome Trust Centre for Neuroimaging, on whose work this article is based. We acknowledge very useful comments from reviewers that helped improve the readability of this article.

Authors

Adeel Razi (adeel.razi@ieee.org) received his B.E. degree in electrical engineering from the NED University of Engineering and Technology, Karachi, Pakistan; his M.S. degree in communications engineering from the University of Technology Aachen (RWTH), Germany; and his Ph.D. degree in electrical engineering from the University of New South Wales, Sydney, Australia, in 2012. Upon the completion of his undergraduate studies, he joined the Department of Electronic Engineering at the NED University of Engineering and Technology, where he is an associate professor. He is currently on leave from NED University to work at the Wellcome Trust Centre for Neuroimaging, University College London, United Kingdom. His general research interest is in demystifying (complex) networks. He currently works on modeling the changes in connectivity patterns of (large) human brain networks.

Karl J. Friston (k.friston@ucl.ac.uk) is a theoretical neuroscientist and authority on brain imaging. He invented statistical parametric mapping, voxel-based morphometry, and dynamic causal modeling. He received the first Young Investigators Award in Human Brain Mapping (1996) and was elected a fellow of the Academy of Medical Sciences (1999). In 2000, he was president of the international Organization of Human Brain Mapping. He was awarded the Minerva Golden Brain Award in 2003 and was elected a fellow of the Royal Society in 2006. In 2008, he received a medal from the Collège de France, and, in 2011, he received an honorary doctorate from the University of York. He became a fellow of the Society of Biology in 2012, received the Weldon Memorial Prize and Medal in 2013, and was elected as a member of Excellence in the Life Sciences in 2014.

References

- [1] D. George and J. Hawkins, "Towards a mathematical theory of cortical micro-circuits," *PLoS Comput. Biol.*, vol. 5, no. 10, e1000532, Oct. 2009.
- [2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [3] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: An explanation of 1/f noise," *Phys. Rev. Lett.*, vol. 59, no. 4, pp. 381–384, July 1987.
- [4] E. Tognoli and J. A. S. Kelso, "The metastable brain," *Neuron*, vol. 81, no. 1, pp. 35–48, Jan. 2014.
- [5] A. Aertsen and F. Preissl, "Dynamics of activity and connectivity in physiological neuronal networks," in *Nonlinear Dynamics and Neuronal Networks*, H. Schuster, Ed. Weinheim, Germany: VCH Verlag, 1991, pp. 281–301.
- [6] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, et al., "Dynamic functional connectivity: Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360–378, Oct. 2013.
- [7] G. K. Cooray, B. Sengupta, P. Douglas, and K. Friston, "Dynamic causal modelling of electrographic seizure activity using Bayesian belief updating," *NeuroImage*, vol. 125, pp. 1142–1154, Jan. 2016.
- [8] G. K. Cooray, B. Sengupta, P. Douglas, M. Englund, R. Wickstrom, and K. Friston, "Characterising seizures in anti-NMDA-receptor encephalitis with dynamic causal modelling," *NeuroImage*, vol. 118, pp. 508–519, May 2015.
- [9] P. A. Robinson, "Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 85, no. 1, 011912, Jan. 2012.
- [10] K. J. Friston, "Functional and effective connectivity: A review," *Brain Connect.*, vol. 1, no. 1, pp. 13–36, 2011.
- [11] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, 1948.
- [12] Y. Shoham, *Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence*. Cambridge, MA: MIT Press, 1988.
- [13] P. Suppes, *A Probabilistic Theory of Causality*. Amsterdam: North-Holland Publishing, 1970.
- [14] J. Pearl, *Causality: Models, Reasoning and Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [15] G. U. Yule, "Notes on the theory of association of attributes in statistics," *Biometrika*, vol. 2, no. 2, pp. 121–134, 1903.
- [16] K. Pearson, A. Lee, and L. Bramley-Moore, "Genetic (reproductive) selection: Inheritance of fertility of man," *Philos. Trans. R. Soc. London A, Math. Phys. Sci.*, vol. 73, pp. 534–539, Jan. 1899.
- [17] E. H. Simpson, "The interpretation of interaction in contingency tables," *J. R. Stat. Soc. Series B Stat. Method.*, vol. 13, no. 2, pp. 238–241, 1951.
- [18] T. W. Armistead, "Resurrecting the third variable: A critique of Pearl's causal analysis of Simpson's paradox," *Am. Stat.*, vol. 68, no. 1, pp. 1–7, Feb. 2014.
- [19] J. Pearl, "Comment: Understanding Simpson's paradox," *Am. Stat.*, vol. 68, no. 1, pp. 8–13, Feb. 2014.
- [20] R. Christensen, "Comment," *Am. Stat.*, vol. 68, no. 1, pp. 13–17, Jan. 2014.
- [21] P. R. Rosenbaum and D. B. Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika*, vol. 70, no. 1, pp. 41–55, 1983.
- [22] J. Pearl, "Causal diagrams for empirical research," *Biometrika*, vol. 82, no. 4, pp. 669–710, Dec. 1995.
- [23] J. Pearl, "The new challenge: From a century of statistics to the age of causation," *Comput. Sci. Stat.*, vol. 29, no. 2, pp. 415–423, 1997.
- [24] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
- [25] C. Anteneodo and R. Riera, "Additive-multiplicative stochastic models of financial mean-reverting processes," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 72, no. 2, 026106, Aug. 2005.
- [26] F. Freyer, J. A. Roberts, P. Ritter, and M. Breakspear, "A canonical model of multistability and scale-invariance in biological systems," *PLoS Comput. Biol.*, vol. 8, no. 8, e1002634, 2012.
- [27] K. J. Friston, B. Li, J. Daunizeau, and K. Stephan, "Network discovery with DCM," *NeuroImage*, vol. 56, no. 3, pp. 1202–1221, June 2011.
- [28] J. Carr, *Applications of Centre Manifold Theory*. Berlin: Springer-Verlag, 1981.
- [29] H. Haken, *Synergetics: An Introduction. Non-Equilibrium Phase Transition and Self-Organisation in Physics, Chemistry and Biology*. Berlin: Springer-Verlag, 1983.
- [30] V. L. Ginzburg and L. D. Landau, "On the theory of superconductivity," *Zh. Eksp. Teor. Fiz.*, vol. 20, pp. 1064–1082, 1950.
- [31] P. A. Valdes, J. C. Jimenez, J. Riera, R. Biscay, and T. Ozaki, "Nonlinear EEG analysis based on a neural mass model," *Biol. Cybern.*, vol. 81, nos. 5–6, pp. 415–424, Nov. 1999.
- [32] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003.
- [33] K. J. Friston, "Variational filtering," *NeuroImage*, vol. 41, no. 3, pp. 747–766, July 2008.
- [34] K. J. Friston, K. Stephan, B. J. Li, and J. Daunizeau, "Generalised filtering," *Math. Probl. Eng.*, vol. 2010, 621670, 2010.
- [35] R. B. Buxton, E. C. Wong, and L. R. Frank, "Dynamics of blood flow and oxygenation changes during brain activation: The balloon model," *Magn. Reson. Med.*, vol. 39, no. 6, pp. 855–864, June 1998.
- [36] K. E. Stephan, N. Weiskopf, P. M. Drysdale, P. A. Robinson, and K. J. Friston, "Comparing hemodynamic models with DCM," *NeuroImage*, vol. 38, no. 3, pp. 387–401, Nov. 2007.
- [37] R. E. Kass and A. E. Raftery, "Bayes factors," *J. Amer. Statist. Assoc.*, vol. 90, no. 430, pp. 773–795, June 1995.
- [38] R. Goebel, A. Roebroeck, D. S. Kim, and E. Formisano, "Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping," *Magn. Reson. Imaging*, vol. 21, no. 10, pp. 1251–1261, Dec. 2003.
- [39] L. Harrison, W. D. Penny, and K. Friston, "Multivariate autoregressive modeling of fMRI time series," *NeuroImage*, vol. 19, no. 4, pp. 1477–1491, Aug. 2003.
- [40] P. A. Valdes-Sosa, A. Roebroeck, J. Daunizeau, and K. Friston, "Effective connectivity: Influence, causality and biophysical modeling," *NeuroImage*, vol. 58, no. 2, pp. 339–361, Sept. 2011.
- [41] S. L. Bressler and A. K. Seth, "Wiener-Granger causality: A well established methodology," *NeuroImage*, vol. 58, no. 2, pp. 323–329, Sept. 2011.
- [42] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, Aug. 1969.
- [43] K. J. Friston, A. M. Bastos, A. Oswal, B. van Wijk, C. Richter, and V. Litvak, "Granger causality revisited," *NeuroImage*, vol. 101, pp. 796–808, Nov. 2014.
- [44] G. Deshpande, K. Sathian, and X. Hu, "Assessing and compensating for zero-lag correlation effects in time-lagged Granger causality analysis of fMRI," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1446–1456, June 2010.
- [45] M. Havlicek, J. Jan, M. Brazdil, and V. D. Calhoun, "Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data," *NeuroImage*, vol. 53, no. 1, pp. 65–77, Oct. 2010.
- [46] L. Barnett and A. K. Seth, "Granger causality for state-space models," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 91, no. 4, Apr. 2015.
- [47] S. Wright, "Correlation and causation," *J. Agric. Res.*, vol. 20, no. 7, pp. 557–585, Jan. 1921.
- [48] O. Duncan, *Introduction to Structural Equation Models*. New York: Academic Press, 1975.
- [49] A. S. Goldberger, "Structural equation methods in social-sciences," *Econometrica*, vol. 40, no. 6, pp. 979–1001, Nov. 1972.
- [50] A. R. McIntosh and F. Gonzalez-Lima, "Structural modeling of functional neural pathways mapped with 2-deoxyglucose: Effects of acoustic startle habituation on the auditory system," *Brain Res.*, vol. 547, no. 2, pp. 295–302, May 1991.
- [51] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston, "Modelling functional integration: A comparison of structural equation and dynamic causal models," *NeuroImage*, vol. 23, no. S1, pp. S264–S274, 2004.
- [52] J. Kim and B. Horwitz, "How well does structural equation modeling reveal abnormal brain anatomical connections? An fMRI simulation study," *NeuroImage*, vol. 45, no. 4, pp. 1190–1198, May 2009.
- [53] K. A. Bollen, *Structural Equations with Latent Variables*. Hoboken, NJ: Wiley, 1989.
- [54] J. Geweke, "Measurement of linear dependence and feedback between multiple time series," *J. Amer. Statist. Assoc.*, vol. 77, no. 378, pp. 304–313, 1982.

- [55] J. S. Bendat, *Nonlinear System Analysis and Identification from Random Data*. Hoboken, NJ: Wiley, 1990.
- [56] K. Friston, R. Moran, and A. K. Seth, "Analysing connectivity with Granger causality and dynamic causal modelling," *Curr. Opin. Neurobiol.*, vol. 23, no. 2, pp. 172–178, Apr. 2013.
- [57] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
- [58] D. Zhou, W. K. Thompson, and G. Siegle, "MATLAB toolbox for functional connectivity," *NeuroImage*, vol. 47, no. 4, pp. 1590–1607, Oct. 2009.
- [59] R. Salvador, J. Suckling, C. Schwarzbauer, and E. Bullmore, "Undirected graphs of frequency-dependent functional connectivity in whole brain networks," *Philos. Trans. R. Soc. London B, Biol. Sci.*, vol. 360, no. 1457, pp. 937–946, May 2005.
- [60] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, M. W. Woolrich, "Network modelling methods for fMRI," *NeuroImage*, vol. 54, no. 2, pp. 875–891, Jan. 2011.
- [61] W. J. Freeman, "Characterization of state transitions in spatially distributed, chaotic, nonlinear, dynamical systems in cerebral cortex," *Integr. Physiol. Behav. Sci.*, vol. 29, no. 3, pp. 294–306, July–Sept. 1994.
- [62] S. Coombes and S. H. Doole, "Neuronal populations with reciprocal inhibition and rebound currents: Effects of synaptic and threshold noise," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*, vol. 54, no. 4, pp. 4054–4065, Oct. 1996.
- [63] P. A. Robinson, C. J. Rennie, and J. J. Wright, "Propagation and stability of waves of electrical activity in the cerebral cortex," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 56, no. 1, pp. 826–840, July 1997.
- [64] I. Tsuda, "Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems," *Behav. Brain Sci.*, vol. 24, no. 5, pp. 793–810, Oct. 2001.
- [65] W. J. Freeman, "A field-theoretic approach to understanding scale-free neocortical dynamics," *Biol. Cybern.*, vol. 92, no. 6, pp. 350–359, June 2005.
- [66] S. L. Bressler and E. Tognoli, "Operational principles of neurocognitive networks," *Int. J. Psychophysiol.*, vol. 60, no. 2, pp. 139–148, May 2006.
- [67] B. Kriener, T. Tetzlaff, A. Aertsen, M. Diesmann, and S. Rotter, "Correlations and population dynamics in cortical networks," *Neural Comput.*, vol. 20, no. 9, pp. 2185–2226, Sept. 2008.
- [68] M. Rubinov, O. Sporns, C. van Leeuwen, and M. Breakspear, "Symbiotic relationship between brain structure and dynamics," *BMC Neurosci.*, vol. 10, p. 55, June 2009.
- [69] M. A. Buice and J. D. Cowan, "Statistical mechanics of the neocortex," *Prog. Biophys. Mol. Biol.*, vol. 99, nos. 2–3, pp. 53–86, Feb.–May 2009.
- [70] C. J. Honey, R. Kotter, M. Breakspear, and O. Sporns, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 104, no. 24, pp. 10240–10245, June 2007.
- [71] G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, and R. Kotter, "Key role of coupling, delay, and noise in resting brain fluctuations," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 25, pp. 10302–10307, June 2009.
- [72] C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. P. Thiran, R. Meuli, and P. Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 6, pp. 2035–2040, Feb. 2009.
- [73] G. Deco, V. K. Jirsa, and A. R. McIntosh, "Emerging concepts for the dynamical organization of resting-state activity in the brain," *Nat. Rev. Neurosci.*, vol. 12, no. 1, pp. 43–56, Jan. 2011.
- [74] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," *Magn. Reson. Med.*, vol. 34, no. 4, pp. 537–541, Oct. 1995.
- [75] B. B. Biswal, J. Van Kylen, and J. S. Hyde, "Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps," *NMR Biomed.*, vol. 10, nos. 4–5, pp. 165–170, June–Aug. 1997.
- [76] J. S. Damoiseaux and M. D. Greicius, "Greater than the sum of its parts: A review of studies combining structural connectivity and resting-state functional connectivity," *Brain Struct. Funct.*, vol. 213, no. 6, pp. 525–533, Oct. 2009.
- [77] M. D. Greicius, K. Supekar, V. Menon, and R. F. Dougherty, "Resting-state functional connectivity reflects structural connectivity in the default mode network," *Cereb. Cortex*, vol. 19, no. 1, pp. 72–78, Jan. 2009.
- [78] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 98, no. 2, pp. 676–682, Jan. 2001.
- [79] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *J. Neurosci.*, vol. 26, no. 1, pp. 63–72, Jan. 2006.
- [80] D. S. Bassett, A. Meyer-Lindenberg, S. Achard, T. Duke, and E. Bullmore, "Adaptive reconfiguration of fractal small-world human brain functional networks," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 51, pp. 19518–19523, Dec. 2006.
- [81] O. Sporns, *Networks of the Brain*. Cambridge, MA: MIT Press, 2010.
- [82] J. Beggs, "Editorial: Can there be a physics of the brain?" *Phys. Rev. Lett.*, vol. 114, no. 22, 220001, June 2015.
- [83] J. M. Beggs, "The criticality hypothesis: How local cortical networks might optimize information processing," *Philos. Trans. A, Math. Phys. Eng. Sci.*, vol. 366, no. 1864, pp. 329–343, Feb. 2008.
- [84] V. K. Jirsa, R. Friedrich, H. Haken, and J. A. Kelso, "A theoretical model of phase transitions in the human brain," *Biol. Cybern.*, vol. 71, no. 1, pp. 27–35, 1994.
- [85] V. K. Jirsa and H. Haken, "Field theory of electromagnetic brain activity," *Phys. Rev. Lett.*, vol. 77, no. 5, pp. 960–963, July 1996.
- [86] V. K. Jirsa and J. A. Kelso, "Spatiotemporal pattern formation in neural systems with heterogeneous connection topologies," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*, vol. 62, no. 6, pp. 8462–8465, Dec. 2000.
- [87] E. Tognoli and J. A. Kelso, "Brain coordination dynamics: True and false faces of phase synchrony and metastability," *Prog. Neurobiol.*, vol. 87, no. 1, pp. 31–40, Jan. 2009.
- [88] W. Tschacher and H. Haken, "Intentionality in non-equilibrium systems? The functional aspects of self-organized pattern formation," *New Ideas Psychol.*, vol. 25, no. 1, pp. 1–15, Apr. 2007.
- [89] W. L. Shew and D. Plenz, "The functional benefits of criticality in the cortex," *Neuroscientist*, vol. 19, no. 1, pp. 88–100, Feb. 2013.
- [90] O. Shriki, J. Alstott, F. Carver, T. Holroyd, R. N. Henson, M. L. Smith, R. Coppola, E. Bullmore, et al., "Neuronal avalanches in the resting MEG of the human brain," *J. Neurosci.*, vol. 33, no. 16, pp. 7079–7090, Apr. 2013.
- [91] A. Haimovici, E. Tagliazucchi, P. Balenzuela, and D. R. Chialvo, "Brain organization into resting state networks emerges at criticality on a model of the human connectome," *Phys. Rev. Lett.*, vol. 110, no. 17, 178101, Apr. 2013.
- [92] J. A. Roberts, T. W. Boonstra, and M. Breakspear, "The heavy tail of the human brain," *Curr. Opin. Neurobiol.*, vol. 31, pp. 164–172, Apr. 2015.
- [93] J. A. Kelso, "Instabilities and phase transitions in human brain and behavior," *Front. Hum. Neurosci.*, vol. 4, p. 23, 2010.
- [94] G. Deco and V. K. Jirsa, "Ongoing cortical activity at rest: Criticality, multi-stability, and ghost attractors," *J. Neurosci.*, vol. 32, no. 10, pp. 3366–3375, Mar. 2012.
- [95] F. Freyer, J. A. Roberts, R. Becker, P. A. Robinson, P. Ritter, and M. Breakspear, "Biophysical mechanisms of multistability in resting-state cortical rhythms," *J. Neurosci.*, vol. 31, no. 17, pp. 6353–6361, Apr. 2011.
- [96] P. A. Robinson, C. J. Rennie, and D. L. Rowe, "Dynamics of large-scale brain activity in normal arousal states and epileptic seizures," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 65, no. 4, 041924, Apr. 2002.
- [97] M. J. Aburn, C. A. Holmes, J. A. Roberts, T. W. Boonstra, and M. Breakspear, "Critical fluctuations in cortical models near instability," *Front. Physiol.*, vol. 3, p. 331, Aug. 2012.
- [98] P. A. Robinson, "Neurophysical theory of coherence and correlations of electroencephalographic and electrocorticographic signals," *J. Theor. Biol.*, vol. 222, no. 2, pp. 163–175, May 2003.
- [99] H. White and X. Lu, "Granger causality and dynamic structural systems," *J. Financ. Economet.*, vol. 8, no. 2, pp. 193–243, 2010.
- [100] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*. New York: Springer, 1993.
- [101] S. L. Lauritzen, *Graphical models*. Gloucestershire, U.K.: Clarendon Press, 1996.
- [102] J. Robins, "A new approach to causal inference in mortality studies with a sustained exposure period—Application to control of the healthy worker survivor effect," *Math. Model.*, vol. 7, nos. 9–12, pp. 1393–1512, 1986.
- [103] K. E. Stephan, L. Kasper, L. M. Harrison, J. Daunizeau, H. E. den Ouden, M. Breakspear, K. J. Friston, "Nonlinear dynamic causal models for fMRI," *NeuroImage*, vol. 42, no. 2, pp. 649–662, Aug. 2008.

- [104] A. C. Marreiros, S. J. Kiebel, and K. J. Friston, "Dynamic causal modelling for fMRI: A two-state model," *NeuroImage*, vol. 39, no. 1, pp. 269–278, Jan. 2008.
- [105] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston, "Comparing dynamic causal models," *NeuroImage*, vol. 22, no. 3, pp. 1157–1172, July 2004.
- [106] F. Acs and M. W. Greenlee, "Connectivity modulation of early visual processing areas during covert and overt tracking tasks," *NeuroImage*, vol. 41, no. 2, pp. 380–388, June 2008.
- [107] P. Allen, A. Mechelli, K. E. Stephan, F. Day, J. Dalton, S. Williams, and P. K. McGuire, "Fronto-temporal interactions during overt verbal initiation and suppression," *J. Cogn. Neurosci.*, vol. 20, no. 9, pp. 1656–1669, Sept. 2008.
- [108] M. Boly, M. I. Garrido, O. Gosseries, M. A. Bruno, P. Boveroux, C. Schnakers, M. Massimini, V. Litvak, et al., "Preserved feedforward but impaired top-down processes in the vegetative state," *Science*, vol. 332, no. 6031, pp. 858–862, May 2011.
- [109] M. J. Grol, J. Majdandzic, K. E. Stephan, L. Verhagen, H. C. Dijkerman, H. Bekkering, F. A. J. Verstraten, and I. Toni, "Parieto-frontal connectivity during visually guided grasping," *J. Neurosci.*, vol. 27, no. 44, pp. 11877–11887, Oct. 2007.
- [110] S. Heim, S. B. Eickhoff, A. K. Ischebeck, A. D. Friederici, K. E. Stephan, and K. Amunts, "Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM," *Hum. Brain Mapp.*, vol. 30, no. 2, pp. 392–402, Feb. 2009.
- [111] A. P. Smith, K. E. Stephan, M. D. Rugg, and R. J. Dolan, "Task and content modulate amygdala-hippocampal connectivity in emotional retrieval," *Neuron*, vol. 49, no. 4, pp. 631–638, Feb. 2006.
- [112] C. Summerfield and E. Koehlin, "A neural representation of prior information during perceptual inference," *Neuron*, vol. 59, no. 2, pp. 336–347, July 2008.
- [113] A. A. Faisal, L. P. Selen, and D. M. Wolpert, "Noise in the nervous system," *Nat. Rev. Neurosci.*, vol. 9, no. 4, pp. 292–303, Apr. 2008.
- [114] K. J. Friston, J. Kahan, B. Biswal, and A. Razi, "A DCM for resting state fMRI," *NeuroImage*, vol. 94, no. 100, pp. 396–407, July 2014.
- [115] K. Friston, "Hierarchical models in the brain," *PLoS Comput. Biol.*, vol. 4, no. 11, e1000211, Nov. 2008.
- [116] B. Li, J. Daunizeau, K. E. Stephan, W. Penny, D. Hu, and K. Friston, "Generalised filtering and stochastic DCM for fMRI," *NeuroImage*, vol. 58, no. 2, pp. 442–457, Sept. 2011.
- [117] J. Daunizeau, K. E. Stephan, and K. J. Friston, "Stochastic dynamic causal modelling of fMRI data: Should we care about neural noise?" *NeuroImage*, vol. 62, no. 1, pp. 464–481, Aug. 2012.
- [118] M. M. Tropper, "Ergodic and quasi-deterministic properties of finite-dimensional stochastic systems," *J. Stat. Phys.*, vol. 17, no. 6, pp. 491–509, 1977.
- [119] K. J. Friston, J. Kahan, A. Razi, K. E. Stephan, and O. Sporns, "On nodes and modes in resting state fMRI," *NeuroImage*, vol. 99, pp. 533–547, May 2014.
- [120] V. D. Calhoun, T. Adali, G. D. Pearson, and J. J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis," *Hum. Brain Mapp.*, vol. 14, no. 3, pp. 140–151, Nov. 2001.
- [121] A. Razi, J. Kahan, G. Rees, and K. J. Friston, "Construct validation of a DCM for resting state fMRI," *NeuroImage*, vol. 106, pp. 1–14, Feb. 2015.
- [122] K. J. Friston, N. Trujillo-Barreto, and J. Daunizeau, "DEM: A variational treatment of dynamic systems," *NeuroImage*, vol. 41, no. 3, pp. 849–885, July 2008.
- [123] J. Kahan, M. Urner, R. Moran, G. Flandin, A. Marreiros, L. Mancini, M. White, J. Thornton, et al., "Resting state functional MRI in Parkinson's disease: The impact of deep brain stimulation on 'effective' connectivity," *Brain*, vol. 137, no. 4, pp. 1130–1144, 2014.
- [124] P. A. Robinson, C. J. Rennie, D. L. Rowe, and S. C. O'Connor, "Estimation of multiscale neurophysiologic parameters by electroencephalographic means," *Hum. Brain Mapp.*, vol. 23, no. 1, pp. 53–72, Sept. 2004.
- [125] C. J. Stam and E. A. de Bruin, "Scale-free dynamics of global functional connectivity in the human brain," *Hum. Brain Mapp.*, vol. 22, no. 2, pp. 97–109, June 2004.
- [126] C. W. Shin and S. Kim, "Self-organized criticality and scale-free properties in emergent functional neural networks," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 74, no. 4, 045101, Oct. 2006.
- [127] J. M. Beggs and D. Plenz, "Neuronal avalanches in neocortical circuits," *J. Neurosci.*, vol. 23, no. 35, pp. 11167–11177, Dec. 2003.
- [128] C. Bedard, H. Kroger, and A. Destexhe, "Does the 1/f frequency scaling of brain signals reflect self-organized critical states?" *Phys. Rev. Lett.*, vol. 97, no. 11, 118102, Sept. 2006.
- [129] K. J. Miller, L. B. Sorensen, J. G. Ojemann, and M. den Nijs, "Power-law scaling in the brain surface electric potential," *PLoS Comput. Biol.*, vol. 5, no. 12, e1000609, Dec. 2009.
- [130] K. Linkenkaer-Hansen, V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi, "Long-range temporal correlations and scaling behavior in human brain oscillations," *J. Neurosci.*, vol. 21, no. 4, pp. 1370–1377, Feb. 2001.
- [131] M. G. Kitzbichler, M. L. Smith, S. R. Christensen, and E. Bullmore, "Broad-band criticality of human brain network synchronization," *PLoS Comput. Biol.*, vol. 5, no. 3, e1000314, Mar. 2009.
- [132] F. Freyer, K. Aquino, P. A. Robinson, P. Ritter, and M. Breakspear, "Bistability and non-Gaussian fluctuations in spontaneous cortical activity," *J. Neurosci.*, vol. 29, no. 26, pp. 8512–8524, July 2009.
- [133] K. Friston, J. Mattout, N. Trujillo-Barreto, J. Ashburner, and W. Penny, "Variational free energy and the Laplace approximation," *NeuroImage*, vol. 34, no. 1, pp. 220–234, Jan. 2007.
- [134] K. A. Tsvetanov, R. N. A. Henson, L. K. Tyler, A. Razi, L. Geerligs, T. Ham, and J. Rowe, "Higher extrinsic and intrinsic brain network connectivity maintains cognition across the lifespan despite age-related decay of regional neuronal activation," *J. Neurosci.*, to be published.
- [135] W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, and M. D. Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cereb. Cortex*, vol. 22, no. 1, pp. 158–165, Jan. 2012.
- [136] P. L. St. Jacques, P. A. Kragel, and D. C. Rubin, "Dynamic neural networks supporting memory retrieval," *NeuroImage*, vol. 57, no. 2, pp. 608–616, July 2011.
- [137] M. C. Stevens, K. A. Kiehl, G. D. Pearson, and V. D. Calhoun, "Functional neural networks underlying response inhibition in adolescents and adults," *Behav. Brain Res.*, vol. 181, no. 1, pp. 12–22, July 2007.
- [138] X. Di and B. B. Biswal, "Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging," *NeuroImage*, vol. 86, pp. 53–59, Feb. 2014.
- [139] J. Daunizeau, O. David, and K. E. Stephan, "Dynamic causal modelling: A critical review of the biophysical and statistical foundations," *NeuroImage*, vol. 58, no. 2, pp. 312–322, Sept. 2011.
- [140] G. Lohmann, K. Erfurth, K. Muller, and R. Turner, "Critical comments on dynamic causal modelling," *NeuroImage*, vol. 59, no. 3, pp. 2322–2329, Feb. 2012.
- [141] M. Havlicek, A. Roebroeck, K. Friston, A. Gardumi, D. Ivanov, and K. Uludag, "Physiologically informed dynamic causal modeling of fMRI data," *NeuroImage*, vol. 122, pp. 355–372, Nov. 2015.
- [142] K. Friston, J. Daunizeau, and K. E. Stephan, "Model selection and gobbledygook: Response to Lohmann et al," *NeuroImage*, vol. 75, pp. 275–278, discussion 279–281, July 2013.
- [143] S. Frassle, K. E. Stephan, K. J. Friston, M. Steup, S. Krach, F. M. Paulus, and A. Jansen, "Test-retest reliability of dynamic causal modeling for fMRI," *NeuroImage*, vol. 117, pp. 56–66, Aug. 2015.
- [144] J. B. Rowe, L. E. Hughes, R. A. Barker, and A. M. Owen, "Dynamic causal modelling of effective connectivity from fMRI: Are results reproducible and sensitive to Parkinson's disease and its treatment?" *NeuroImage*, vol. 52, no. 3, pp. 1015–1026, Sept. 2010.
- [145] B. Schuyler, J. M. Ollinger, T. R. Oakes, T. Johnstone, and R. J. Davidson, "Dynamic causal modeling applied to fMRI data shows high reliability," *NeuroImage*, vol. 49, no. 1, pp. 603–611, Jan. 2010.
- [146] M. I. Garrido, J. M. Kilner, S. J. Kiebel, K. E. Stephan, and K. J. Friston, "Dynamic causal modelling of evoked potentials: A reproducibility study," *NeuroImage*, vol. 36, no. 3, pp. 571–580, July 2007.
- [147] O. David, I. Guillemain, S. Saitlet, S. Reynt, C. Deransart, C. Segebarth, and A. Depaulis, "Identifying neural drivers with functional MRI: An electrophysiological validation," *PLoS Biol.*, vol. 6, no. 12, pp. 2683–2697, Dec. 2008.
- [148] E. I. Lomakina, S. Paliwal, A. O. Diaconescu, K. H. Brodersen, E. A. Aponte, J. M. Buhmann, and K. E. Stephan, "Inversion of hierarchical Bayesian models using Gaussian processes," *NeuroImage*, vol. 118, pp. 133–145, Sept. 2015.
- [149] B. Sengupta, K. J. Friston, and W. D. Penny, "Gradient-free MCMC methods for dynamic causal modelling," *NeuroImage*, vol. 112, pp. 375–381, May 2015.
- [150] B. Sengupta, K. J. Friston, and W. D. Penny, "Gradient-based MCMC samplers for dynamic causal modelling," *NeuroImage*, vol. 125, pp. 1107–1118, Jan. 2016.

Mapping Brain Anatomical Connectivity Using Diffusion Magnetic Resonance Imaging



IMAGE LICENSED BY INGRAM PUBLISHING

Structural connectivity of the human brain

Junning Li, Yonggang Shi, and Arthur W. Toga

In 2009, the National Institutes of Health ambitiously launched the Human Connectome Project [1] to promote engineering capabilities for imaging and analyzing brain connections. One of the primarily promoted technologies is diffusion magnetic resonance (dMR) imaging, which noninvasively maps brain connectivity at a macroscopic scale by measuring water molecules' anisotropic diffusion constrained

by neural fibers. Following years of steady advancement, the dMR imaging technique has reached unprecedented spatial and angular resolution, and its computational analysis methods, stimulated by growing research needs, have also blossomed. This has been achieved by joint contributions from various areas, such as signal processing, applied mathematics, network analysis, and so on. In this article, we outline the milestones on this exciting path of interdisciplinary technology development with the aim of bringing these advancements to engineers outside the medical imaging community.

Digital Object Identifier 10.1109/MSP.2015.2510024
Date of publication: 27 April 2016

Introduction

This article covers key components in the workflow to map structural connectivity of the human brain, including data acquisition, neural fiber orientation modeling, image processing, tractography, and applications to brain studies. For each topic, basic theories are reviewed, and major breakthroughs and state-of-the-art technologies are discussed. In addition to an overview of existing technologies, we also attempt to provide an outlook of future challenges in building a comprehensive connectivity map that integrates genetic and functional information.

dMR imaging

At a microscopic scale, water molecules in an isotropic medium move freely in all directions in a jittery and erratic fashion. This random walk was first noticed in 1827 by Scottish botanist Robert Brown, explained physically by Albert Einstein in 1905, and later rigorously modeled as a mathematical stochastic process by Norbert Wiener. Movement of these tiny molecules inspired mathematicians to develop elegant theories now widely used not only by financial analysts but also by medical engineers to “look” through the human brain in vivo at its sophisticated neural network. In an anisotropic medium, such as brain tissues containing neuronal fibers,

water molecules move faster along, rather than across, structural constraints. By measuring their anisotropic diffusion, we can infer the underlying structure of neuronal fibers. Figure 1 shows a typical information flow of studying brain connectivity using dMR imaging: neural fibers introduce anisotropic diffusion (a) and influence magnetic resonance (MR) signals generated from spinning protons (b); dMR signals are collected along many diffusion directions (c) and then reconstructed at each voxel location as fiber orientation distributions (FODs) (d); local fiber structures are assembled and “weaved” together as long fiber tracks (e) to build a network (g) connecting brain cortex regions (f). These networks are further investigated for their relationship with the brain’s function and development. We elaborate this procedure step by step in the following sections.

We first look at how dMR signals are generated. Understanding the generation procedure is important for the processing and interpretation of dMR images. It also sheds light on the potential limitations and caveats of using dMR images. We start with the excitation of protons in a magnetic field and then explain how diffusion affects signal resonance, with formulations in the so-called k -space and q -space. The relationship between diffusion and magnetic resonance is the key to dMR imaging and its data analysis.

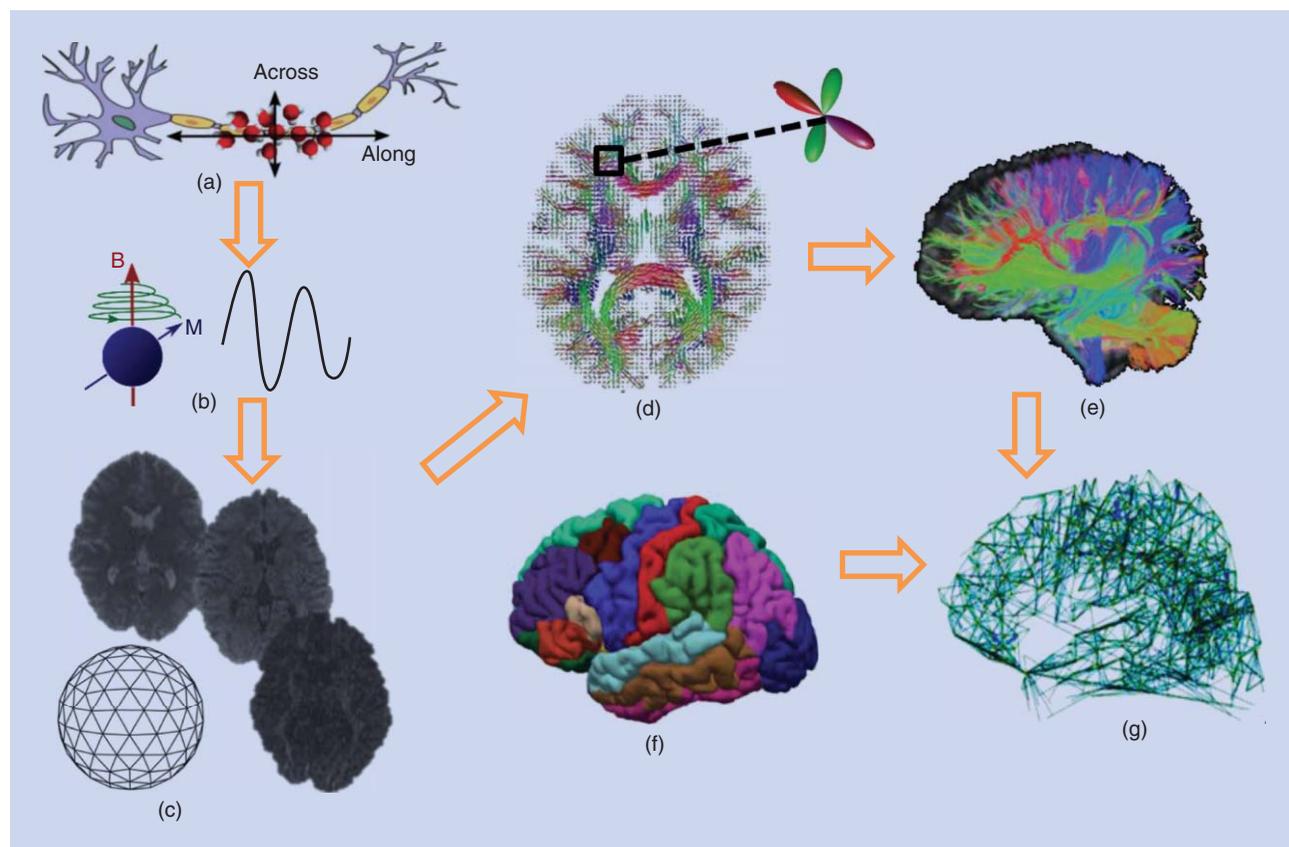


FIGURE 1. The information flow of structural connectivity analysis. (a) Water molecules move faster along, rather than across, neuronal fibers. (b) Diffusion affects the electromagnetic waves radiated by precessing protons. (c) dMR imaging captures diffusion signals along different directions and forms images. (d) FODs are reconstructed from diffusion images. (e) Fiber tracks are simulated from FOD images. (f) The brain cortex is segmented into many regions using structural MR images, for example, with FreeSurfer software [2]. (g) Connectivity networks between cortex regions are constructed from fiber tracks.

Nuclear magnetic resonance

A proton/hydrogen, spinning with its positive electrical charge, forms a tiny magnetic moment along its spin axis. Under normal circumstances, protons spin randomly in all directions [Figure 2(a)]. In the presence of an external magnetic field, they align with the field, getting polarized [Figure 2(b)]. In this case, if the spin axis is perturbed away from the field direction, it will rotate in a spiral path to gradually realign with the magnetic field [Figure 2(c)]. This is analogous to the motion of a rotating gyroscope hung on a rope. Perturbation from the equilibrium is called *excitation*, and restoration is called *relaxation*.

The spiral return to equilibrium consists of three different motions. The first motion is precession [Figure 2(d)] around the field direction at the Larmor frequency (named after Sir Joseph Larmor, 1875–1942) $\omega = \gamma B$, where γ is the gyromagnetic ratio [3]. This Larmor precession is essential in MR imaging. We can transmit radio waves at this frequency to excite protons, and the excited protons will in return radiate electromagnetic waves at the same frequency in their precession. To maximize the radiated energy, protons are flipped by 90° in excitation, perpendicular to the field direction. This resonated radio wave is collected as a function of time, and its Fourier transform shows a peak near the Larmor frequency, reflecting protons' spin densities. If a magnetic field with a linear gradient is applied, then protons' spatial locations will be "encoded" into their precession frequencies. In this way, we can recover the spatial density of spinning protons in the Fourier domain, as elaborated later in the "k-Space" and "q-Space" sections. The other two motions are the precession plane's movement toward the equilibrium and shrinkage of the precession radius. Usually, they are exploited to produce an image contrast ratio [3], but we do not go into their details here.

k-Space

If a gradient g is added to a uniform magnetic field B , then protons at location r precess at their local Larmor frequency

$\gamma(B + g \cdot r)$. The radio wave they radiate, as a function of time, is $s(r)e^{i\gamma Bt}e^{i\gamma(g \cdot r)t}$, where the magnitude $s(r)$ reflects the amount of local magnification and is the interest of imaging. Because the carrier wave $e^{i\gamma Bt}$ can be removed with a heterodyne mixer in the receiver [3], it effectively reduces to $s(r)e^{i\gamma(g \cdot r)t}$. (A heterodyne mixer shifts signal frequency by multiplying it with another wave. If two signals at frequencies f_1 and f_2 , respectively, are mixed with multiplication, the result is a mixture of frequencies $f_1 + f_2$ and $|f_1 - f_2|$. After removing the unwanted frequency with a high- or low-pass filter, a signal with shifted frequency is produced.) Instead of naturally referencing with time t , Sir Peter Mansfield and Peter Grannell in 1973 introduced a vector $k = \gamma g t$ [4]. In k -space [5], the total signal collected from all protons in volume V turns out to be

$$S(k) = \int_V s(r)e^{ik \cdot r} dr. \tag{1}$$

It is clearly shown in (1) that local signal $s(r)$ can be reconstructed from $S(k)$ with the Fourier transform [4]. This relationship lies in the heart of spatial reconstruction, which Mansfield used to develop the imaging techniques that earned him a Nobel Prize.

q-Space

In a homogeneous field, the exchange of protons between different locations does not affect local signals because they all precess at the same Larmor frequency in the same phase after their initial excitation. When protons precess at different phases spatially, then their movement in and out of a location will mix phases and reduce signal magnitude. If we can encode a proton's displacement into its phase shift, we will be able to quantitatively deduce its displacement from the signals reduced by phase mixture.

In 1965, Stejskal and Tanner [6] invented a widely used scheme that makes such a deduction of diffusivity from MR possible, as shown in Figure 3. After excitation, a strong gradient pulse g is applied for a short duration δ to establish a

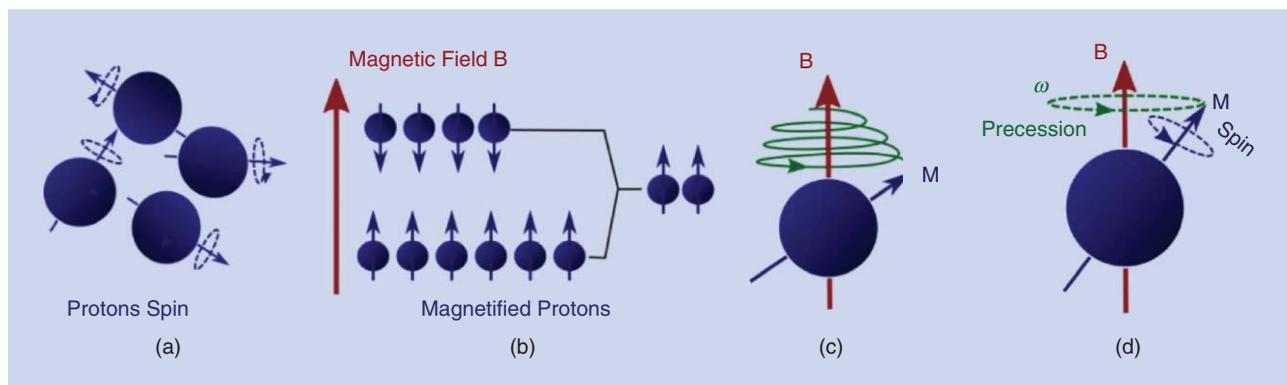


FIGURE 2. Nuclear MR. (a) Protons spin randomly, each forming a tiny magnetic moment. (b) In the presence of magnetic field B , protons align with the field. Some spin "down" against the field, and more spin "up" along the field. The net magnification is along the field. (c) If a proton is perturbed from the field direction, it will return to its equilibrium in a spiral path, similar to a rotating gyroscope hung on a rope. It radiates electromagnetic waves during this process. (d) The spiral path has a precession component, and the frequency is determined by the Larmor equation $\omega = \gamma B$, where γ is the gyromagnetic ratio [3]. In the human body, there are numerous protons that can be used to generate MR signals.

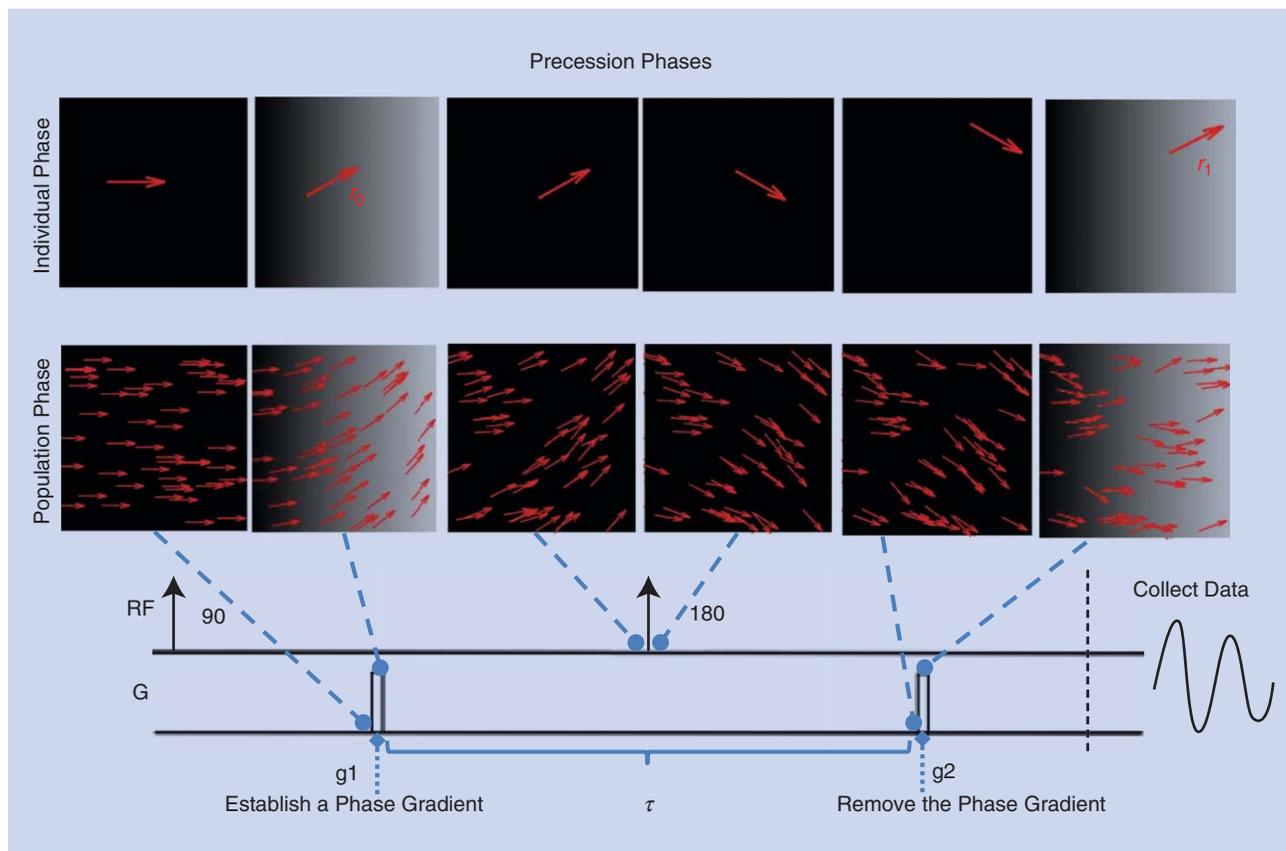


FIGURE 3. The paradigm of Stejskal and Tanner’s [6] diffusion MR imaging scheme. The first row shows the phase changes of a single spinning proton. Before application of the first gradient pulse g_1 , its phase is zero. During g_1 , it receives a phase shift $q \cdot r_0$ at its location r_0 . The 180° pulse reverses the phase gradient and negates the initially received phase tag to $-q \cdot r_0$. Immediately before application of the second gradient pulse, the proton has moved to a new location r_1 , and the second gradient pulse exerts a phase shift $q \cdot r_1$ on the proton. The net phase shift the proton receives is $q \cdot (r_1 - r_0)$. If we define displacement vector $v \equiv r_1 - r_0$, then the phase shift is $q \cdot v$, independent of the starting location r_0 . The second row shows how diffusion causes phase incoherence among protons at the same snapshot time points of the first row. Due to displacement, protons cannot restore their initial phases, with some shifted forward and some backward, so their phases become incoherent, and the signal is reduced.

phase gradient q . If δ is so short that protons have hardly displaced before the establishment of q , protons at location r_0 will almost instantly receive a phase shift of $q \cdot r_0$. Later, a radio frequency pulse is transmitted to flip the spinning protons by 180° . This flipping technique, invented by Hahn in 1950 [7], reverses the phase gradient spatially and negates the initially received phase shift to $-q \cdot r_0$. After time τ of the first gradient pulse, a second gradient pulse is applied to remove the reversed phase gradient. After migrating to location r_1 , protons will receive a phase shift $q \cdot r_1$. In this way, the displacement from r_0 to r_1 is translated to a net phase shift $q(r_1 - r_0)$. If we denote the displacement vector as $v \equiv r_1 - r_0$, then the net phase shift is solely determined by the displacement vector, independent of the initial location. Diffusion duration is controlled by the interval τ between the two gradient pulses.

The total signal $s_\tau(q)$ in the Stejskal–Tanner experiment is integration of all possible displacement signals in terms of $e^{iq \cdot v}$ and weighted with the average displacement probability $P_\tau(v) = \int P_\tau(r \rightarrow r+v)p(r)dr$, where $p(r)$ is the spin density at location r , and $P_\tau(r \rightarrow r+v)$ is the probability of

moving from r to $r+v$ after time τ . $P_\tau(v)$ averages displacement probability over start locations and is called the *ensemble average propagator*. If s_0 is the signal without the two gradient pulses, then $s_\tau(q)$ relates to s_0 through the superposition of all phase shifts as

$$s_\tau(q) = s_0 \int P_\tau(v) e^{iq \cdot v} dv. \tag{2}$$

There is a close resemblance between (2) and (1), and similarly we call the space formed by a phase gradient q -space. The ensemble average diffusion propagator $P_\tau(v)$ can be obtained from $s_\tau(q)$ with the Fourier transform. Note that (2) assumes that the two gradient pulses are so short that their phase shifts have perfectly tagged the protons’ locations before the protons have hardly displaced. However, this is not true in practice, so (2) is inaccurate in most cases.

The diffusion propagator of isotropic media is a Gaussian distribution with variance $2D\tau$, where D is the diffusion coefficient. Because the Fourier transform of a Gaussian distribution is still a Gaussian, the diffusion signal in this case also has a Gaussian format: $s_\tau(q) = s_0 e^{-|q|^2 D \tau}$. As mentioned

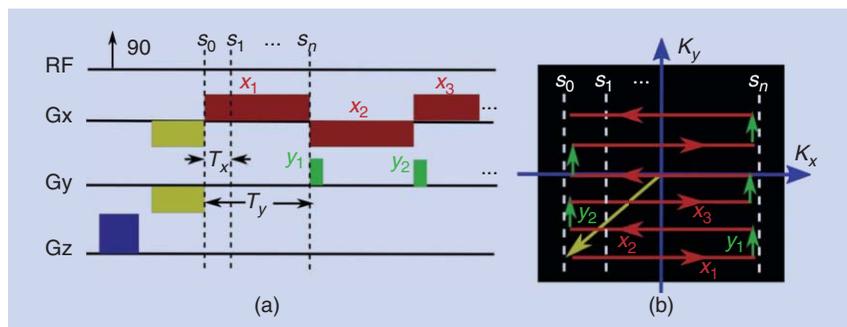


FIGURE 4. EPI: (a) gradient pulses and (b) the corresponding path in k -space. In the company of a gradient pulse in the z direction [(a), blue], a radio frequency selectively excites a slice of an image. Negative gradient pulses [(a), yellow] along the x and y directions then initialize the scan at the lower left corner of k -space, as shown by the yellow arrow in (b). As the x gradient switches [(a), red], it travels back and forth in k -space. Whenever the x gradient changes direction, a short gradient pulse in the y direction [(a), green] is applied to move the scan in the x direction to a new y coordinate in k -space. If data are sampled at s_0, s_1, \dots, s_n in each scan of an x -direction line, then it takes time T_x to move one step in the x direction and T_y in the y direction. T_y is much longer than T_x .

previously, this equation is inaccurate because of the unneglectable gradient pulse duration δ . Adjusted for rectangle gradient pulses, $s_\tau(q)$ becomes $s_0 e^{-|q|^2 D(\tau - \delta/3)}$. The $|q|^2(\tau - \delta/3)$ term determines how sensitive $s_\tau(q)$ is to the diffusion coefficient D , and Bassler, Mattiello, and Le Bihan [8] coined the name b -value for it. The direction of a q vector can be assigned to the b -value derived from it. In this way, we get a b vector from a q vector.

There are several schemes to sample b vectors. Sampling b vectors of the same magnitude is called *single-shell acquisition*. Sampling b vectors of multiple magnitudes is called *multishell acquisition*. Sampling b vectors placed on a three-dimensional lattice is called *diffusion spectrum imaging*.

Data acquisition

Innovative engineering ideas are often needed to turn elegant theories into practical, efficient, and affordable technologies, and such is the case of dMR image acquisition. One of the major difficulties encountered with dMR imaging is its acquisition time because it essentially collects data for a six-dimensional space: three dimensions for k -space and the other three dimensions for q -space. In the early days of MR imaging, it typically took 10–20 minutes to scan a nondiffusion structural image of 64×64 pixels [9]. Currently, the protocol developed in the Human Connectome Project takes about 55 minutes to twice scan a complete brain at the 1.25-mm resolution (about $150 \times 150 \times 150$ pixels) in 270 different diffusion directions [10]. This amazing speed was achieved with a series of breakthroughs.

Echo-planar imaging

Echo-planar imaging (EPI) was invented by Sir Peter Mansfield in 1977 [11]. This technique made it possible to obtain an individual MR slice in the time frame of 50–100 milliseconds. After selectively exciting a slice of an image with the radio frequency determined by a gradient pulse along

the z direction, EPI efficiently traverses k -space by modulating the field gradient in the x and y directions, as illustrated in Figure 4. It first sends negative gradient pulses in both the x and y directions to initialize the scan at the left lower corner in k -space. It then alternatively switches between positive and negative gradients along the x direction, iterating forth and back in k -space. Meanwhile, whenever it reverses the x -gradient direction, it sends a short positive gradient pulse along the y direction to start the scan with a new y coordinate in k -space. In this way, k -space is iterated in a zig-zag manner, and a slice of the image can be scanned in one excitation. Because gradient pulses along the x direction control

protons' precession frequencies and those along the y direction shift their phases each time, they are called the frequency- and phase-encoding directions, respectively. Note that it takes a much longer time for EPI to move one step along the phase-encoding direction than the frequency-encoding direction. This will make artifacts more pronounced along the former than the latter, as discussed in the section "Artifacts of dMR Imaging."

Parallel imaging

Parallel imaging exploits receivers' localized spatial sensitivities to reconstruct an image with reduced sampling in k -space. Ordinarily, if k -space is undersampled, the reconstructed image in a reduced field of view shows an overlapping effect, that is, aliasing. In the aliased view, a pixel value is the linear combination of its aliasing pixels in the full view. If multiple receiver coils are placed at different locations, each of their reconstructed images takes a different linear combination because of their different spatial sensitivities. With many different linear combinations, it is possible to separate all the aliasing pixels provided that the number of receiver coils is at least the undersample rate. The separation can be achieved in two ways: either in the image domain, as with sensitivity encoding (Figure 5) [12], or in the frequency domain, as with the generalized autocalibrating partially parallel acquisitions [13].

Multiband multislice imaging

Multiband multislice imaging excites multiple image slices by transmitting multiband radio frequencies and reconstructs the excited slices with signals collected by multiple receivers [14]. Because the signals received by each coil are a unique linear combination of the excited slices, they can be separated as long as the number of receivers is not fewer than the number of excited slices. The multiband multislice technique can be used together with parallel imaging to achieve two-way acceleration [15].

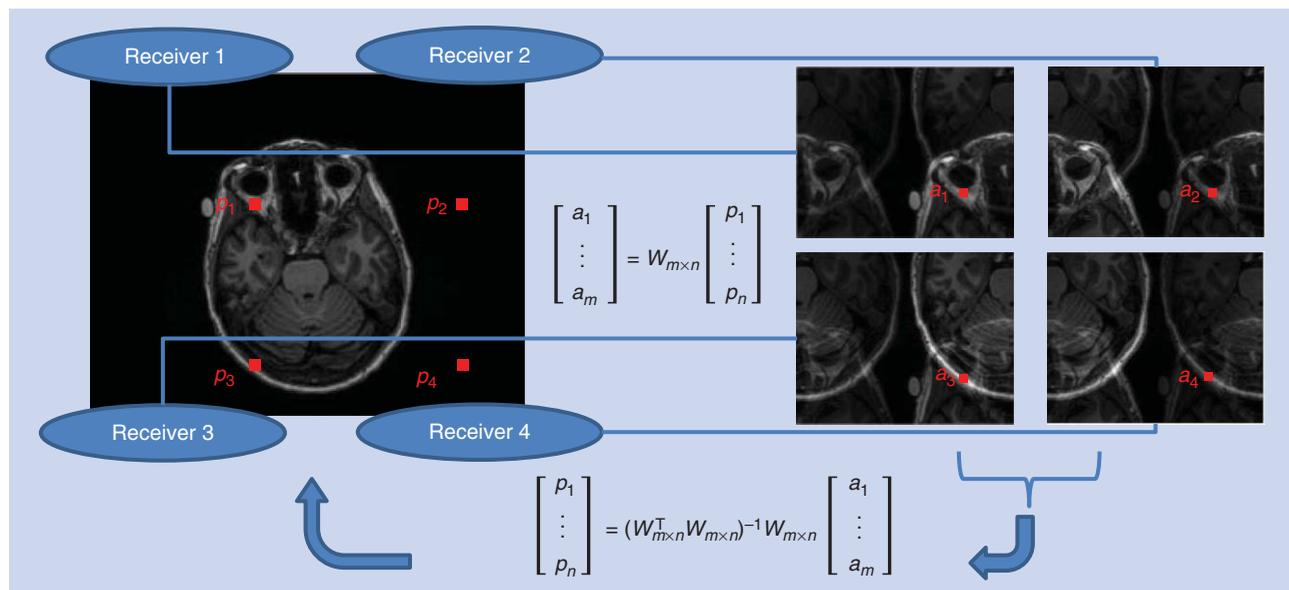


FIGURE 5. Parallel imaging with sensitivity encoding [12]. Suppose that we have $m = 4$ receivers and an image acquisition undersample k -space by two in both the x and y axes. Each pixel in the reduced field of view has $n = 4$ aliasing pixels. The n aliasing pixels p_1, \dots, p_n are linearly combined with a sensitivity matrix $W_{m \times n}$ in each of the reconstructed images, resulting in a_1, \dots, a_2 in the reduced field of view. The full-view image can be recovered by solving a linear problem for each of the aliased pixels.

Compressed sensing

In its theory, compressed sensing ensures that data sparse in a domain can be reconstructed without much information loss by sporadically sampling them in another almost random domain. Because there is only a limited number of crossing fibers at one voxel, it is possible to reconstruct diffusion displacement distribution in a compressed manner. To further accelerate dMR imaging with compressed sensing, various sparse bases have been proposed, including wavelets [16], spherical ridgelets [17], and adaptive dictionaries [18].

Artifacts of dMR imaging

As shown in the previous sections, dMR imaging encodes protons' locations by their precession phase and recovers spatial information, such as voxel location and displacement, by transforming signals from the frequency domain, such as

k -space or q -space, to the spatial domain. This is the essential rule of MR imaging. When assumptions establishing the phase–location relationship are violated, then artifacts come in, and some might be very insidious. Here, we briefly go through some of them without diving into details of their related MR pulse sequences but explain from k -space and q -space perspectives.

Gibbs ringing

Gibbs ringing, also known as *truncation* or *spectral leakage*, appears as spurious ripples around sharp edges. It is a consequence of reconstructing images from MR signals with the Fourier transform. With an infinitely wide frequency band, any signal can be almost perfectly represented. However, in MR imaging practice, we can sample only a finite number of frequencies. As a result, the high-frequency oscillation

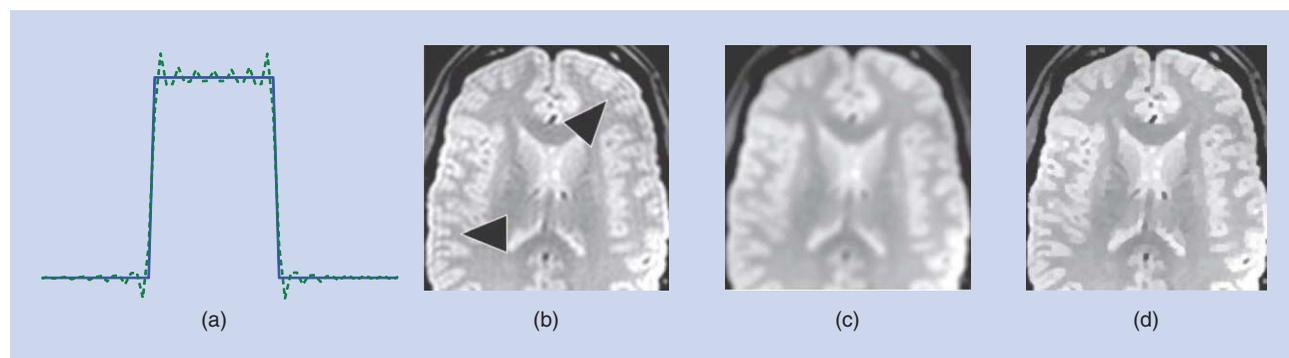


FIGURE 6. Gibbs ringing. (a) When sharp edges are approximated with a truncated Fourier series, the missing high-frequency components produce ripples around the edges. (b) An MR image with Gibbs ringing. (c) The same image low-pass filtered. (d) The same image corrected with a total-variation method. [(b)–(d) are reprinted from [20] with permission.]

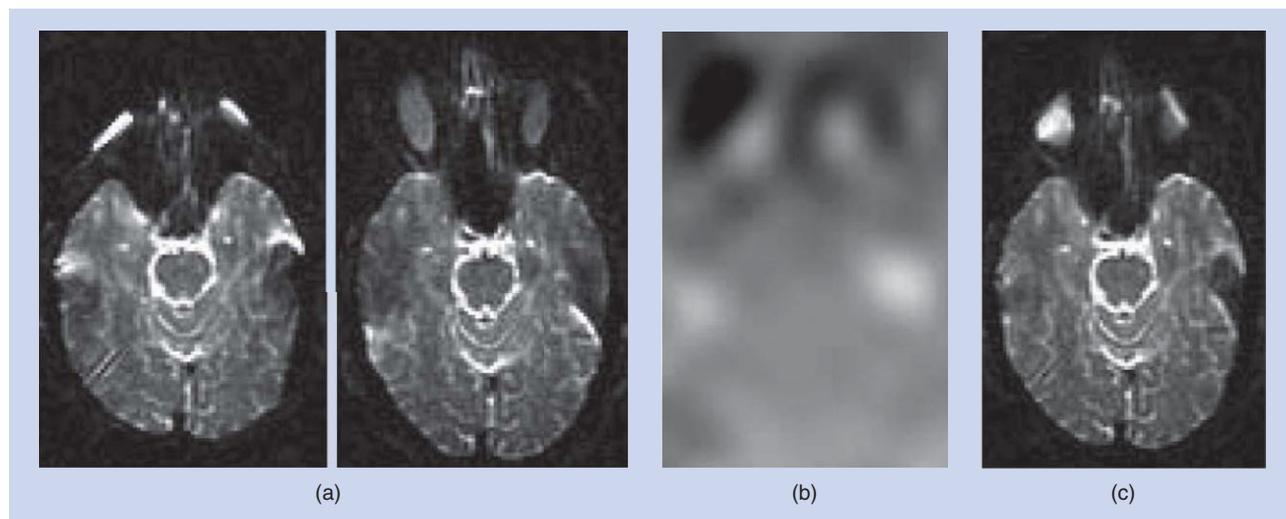


FIGURE 7. An example of geometric distortion. (a) Two diffusion MR images collected with opposite phase-encoding directions. (b) The estimated field map from (a) by nonlinear registration. (c) The corrected image. (Figure reprinted from [22] with permission.)

required to represent sharp edges is truncated, and such truncation shows in the reconstructed images as ripples near high-contrast edges, as shown in Figure 6.

Because fewer samples are usually taken in the phase-encoding direction, the Gibbs artifact is most prominent along this direction. One remedy is to increase the number of phase-encoding steps or reduce the field of view. However, it can never be eliminated because it fundamentally lies in the Fourier reconstruction rule of MR imaging. Low-pass filtering in k -space [19] is a straightforward postprocessing method to minimize Gibbs ringing, but this may blur the image and lower its quality. More complicated methods, such as total variation [20], have also been developed.

Geometric distortion

As EPI accelerates imaging speed, it also brings a very insidious artifact: geometric distortion. With a homogenous gradient field, the protons' spatial locations can be linearly mapped to their precession phase or frequency. What will happen if the gradient varies spatially, or equivalently, and the magnetic field is nonlinear? The result is not additive noise overlaid on a "noise-free" image, but spatial distortion because the mapping between the spatial location and signal phase becomes nonlinear. Even with a scanner producing a perfect linear magnetic field, tissues with different magnetic susceptibility still introduce local nonlinear gradients.

Geometric distortion is far more obvious in the phase-encoding direction than in the frequency-encoding direction. If there is a nonlinear background gradient other than the controlled one, it shifts the protons' precessions during the image acquisition procedure, and this effect accumulates. For the frequency-encoding direction, the accumulation time for each sample step is just the interval between two sampled points. For the phase-encoding direction, the accumulation time for each scan line is the interval between the start time points of two lines. This is much longer than that for the frequency-

encoding direction and consequently accumulates much more distortion effect. Distortion along the frequency-encoding direction usually is at the subvoxel level, whereas along the phase-encoding direction, it could be three to ten voxels.

Geometric distortion can be corrected by unwarping with an estimated background field map [21]. Distortion stretches regions where the unwanted background gradient is along the encoding direction and squeezes where it is against. If the same image is scanned twice, with opposite encoding directions, then it will experience two opposite distortions. With the two distorted images, it is possible to estimate the distortion field [22], as shown in Figure 7. Stretched regions can be better restored than squeezed regions because they still hold all intensity information.

Eddy currents

A time-varying magnetic field generates an electrical field, and, in turn, a time-varying electrical field generates a magnetic field. When strong magnetic gradient pulses are switched on and off rapidly, which is common in dMR imaging, they induct electrical currents on the conductive surface of MR scanners. The currents, called *eddy currents*, may persist and produce a magnetic gradient other than the controlled one. As explained previously, such an extra gradient will lead to geometric distortion.

It is better to suppress the effect of eddy currents during acquisition rather than during postprocessing. It can be reduced first at the source with shielded gradient coils and further by using a special twice-refocused spin echo [23]. It can also be corrected during the postprocessing stage with an estimated field map.

Fat shift

So far, we have assumed that all protons have the same gyromagnetic ratio γ , so in a magnetic field with a homogeneous gradient, their Larmor frequency should be

linearly associated with their location. However, we also have a considerable amount of fat in the scalp whose protons precess with a much higher gyromagnetic ratio, in long chains of carbons with one to three hydrogen atoms. Because spatial location in MR imaging is encoded with phase or equivalently frequencies, fat signals in the reconstructed image will not appear additively at the location of scalp but, instead, map to locations where water protons precess with the same Larmor frequency as they do. In a magnetic field of 3 T, fat protons precess with a frequency 400 Hz higher than water protons. The phase-encoding dimension of a typical EPI has around 30 Hz/pixel, so the 400-Hz difference will show as a shift of a dozen pixels in the phase-encoding direction.

There is another reason that fat shift is more harming to dMR imaging than other modalities. The diffusion coefficient of fat is much lower than that of water, so its signal attenuation, according to the equation $S = S_0 e^{-bD}$, is stronger than that of water protons. Therefore, fat shift appears as bright curved lines in dMR images or as dark lines in derived diffusion coefficient images, as shown in Figure 8.

This fat-shift effect can be suppressed by various methods, and there is no simple answer to which is best. The most widely used is to first excite fat protons at their frequency and remove their phase coherence with a dephasing gradient pulse before imaging. Although they still spin and precess, as the fat protons are dephased, their MR signals become very weak in comparison with those of water protons.

Reconstruction of FODs

Diffusion signals captured by dMR imaging distinguish from fibers' spatial distribution in the following senses. First, they reflect the Brownian motion restricted by neuronal fibers rather than neuronal fibers themselves. Second, they are the average of diffusion signals within voxels, not a detailed microscopic image. The voxel resolution of dMR images is usually 2 mm, and that which was acquired in the Human Connectome Project is 1.25 mm [25]. On the

other hand, the diameter of the axon is at the micrometer level [26]. Therefore, it is impossible to reconstruct the exact fiber structure with dMR images. However, it is possible to estimate statistical properties of neuronal fibers from dMR images by modeling diffusion properties of brain tissues. As an inverse problem, such estimation typically involves a representation of FODs and a forward model to relate FODs to diffusion signals. After briefly introducing two popular representations of FODs, diffusion tensors and spherical harmonics, we discuss the essential part of reconstruction: signal generation models.

Diffusion tensors

In the early 1990s, it was feasible to scan the brain only in a few directions. The limited angular resolution did not support complicated models, so FODs were depicted with the most concise anisotropic model: Gaussian distributions determined by symmetric, positive definite matrices, which are called *diffusion tensors* [8]. The principal eigenvector of a diffusion tensor reflects the dominant fiber direction, and its eigenvalues characterize rotation-invariant properties. The tensor model is unable to effectively account for crossing fibers, as shown in Figure 9. In the presence of crossing fibers, it usually reduces to a nearly isotropic diffusion “ball” or a thin and round “plate.” Such a side effect will lead fiber track simulation to propagate in the wrong directions when fibers actually cross each other, which is not rare in the brain. To solve this problem, high-order diffusion tensors have been proposed [27].

Spherical harmonics

Breakthroughs in dMR imaging have made it practical to scan the brain in a hundred or more directions in a reasonable time, resolving the problem of crossing fibers [10]. To fully utilize such high angular resolution in data acquisition, a probabilistic distribution F defined on a unit sphere S^2 has been employed, replacing diffusion tensors. Similarly, every smooth function in a linear space can be represented with

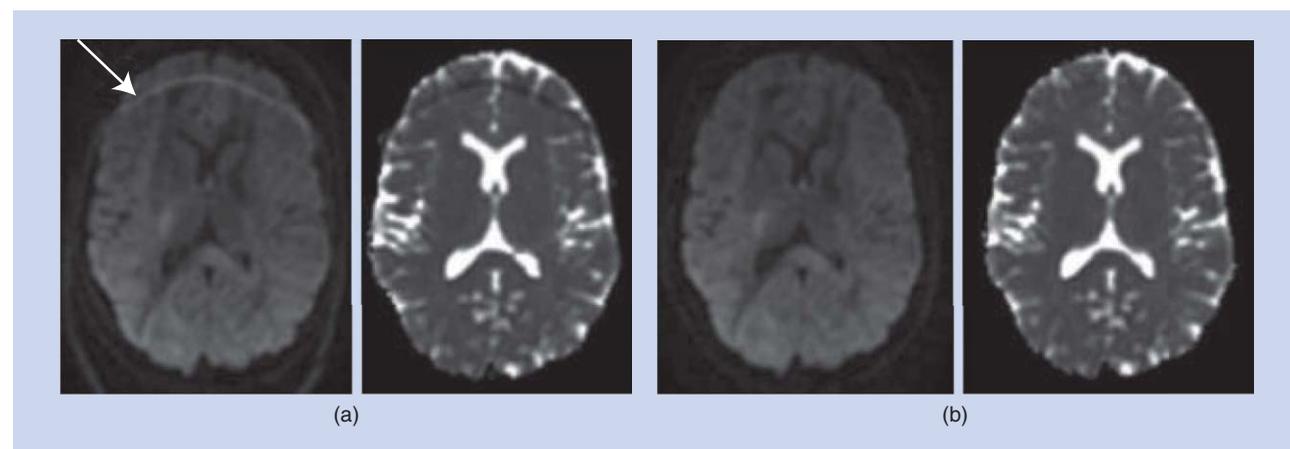


FIGURE 8. An example of fat shift. (a) An image with fat shift (left) and the derived apparent diffusion coefficient (ADC) image (right). (b) An image of the same subject with fat shift suppressed (left) and the derived ADC image. (Figure reprinted from [24] with permission.)

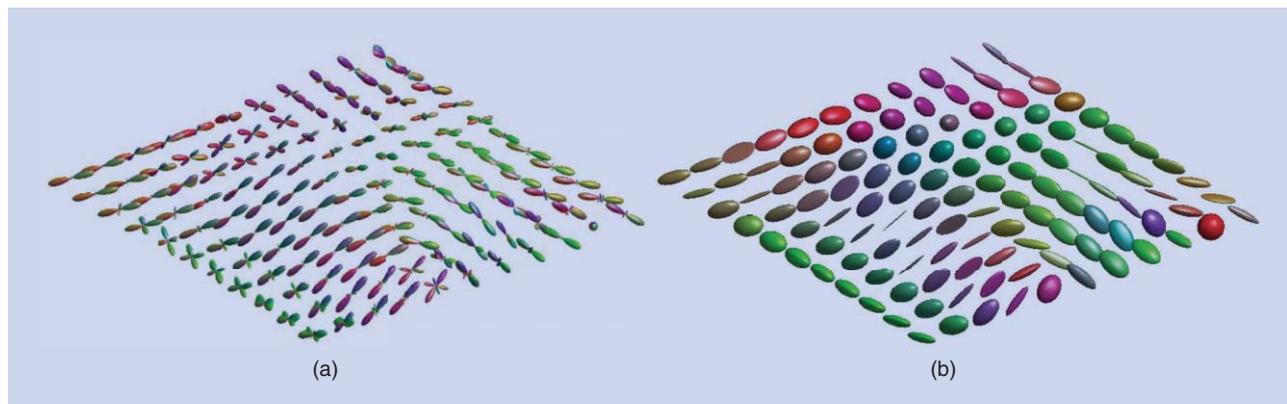


FIGURE 9. A patch of an FOD image represented with (a) eighth-order spherical harmonics and (b) second-order diffusion tensor. For the spherical harmonic representation, FOD values in different directions are shown as radius length, and each direction is color coded by linearly combining the red, green, and blue colors according to the x , y , and z components of unit directional vectors. For the diffusion tensor representation, FODs are rendered as ellipsoids and color coded according to their principal eigenvector. Diffusion tensors cannot faithfully represent crossing fibers.

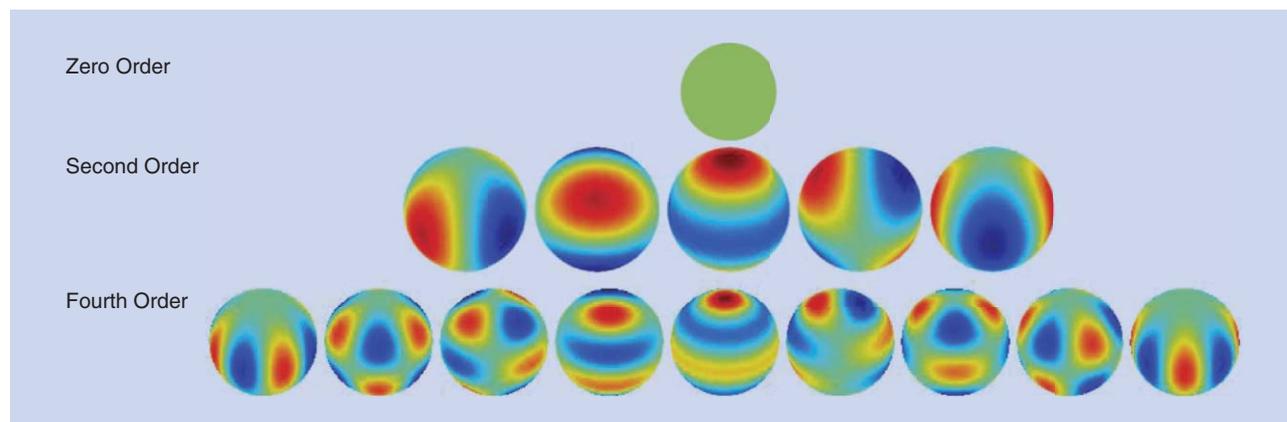


FIGURE 10. Spherical harmonics at the zero, second, and fourth orders. Function values are color coded from blue (low) to red (high).

a set of sine and cosine functions, and a smooth spherical function can also be represented with a set of orthonormal functions oscillating on the sphere at different frequencies, that is, spherical harmonics [28], as shown in Figure 10. The more frequency bands are employed, the more details can be represented. Because spherical harmonic functions are essentially polynomials, they are interchangeable with high-order tensors [27].

The purpose of dMR imaging is not only to estimate diffusion coefficients but also to relate with underlying FODs and tissue composition. Because it is an ill-posed inverse problem, such estimation must be done with highly simplified signal generation models. As an emerging field under active exploration and also due to a lack of validation with microscopic images, researchers have not reached agreement, and many issues are still controversial.

Given an FOD F and a diffusion model H for fibers, the diffusion signal is the convolution of F and H on a unit sphere: $F *_{s^2} H$ (where $*_{s^2}$ denotes spherical convolution) plus contributions from other tissues. The estimation of F relies on diffusion models for neuronal fibers and other restrictive biological structures. A comprehensive analysis [29] has proposed the

following three types of components to characterize dMR signals: intra-axonal compartments, extra-axonal compartments, and highly restricted compartments. The total diffusion signal is approximated by the sum of the three components. It should be noted that such models are just an abstraction and simplification of the underlying complicated microscopic structure. There are also different options. For example, another study [30] showed that highly restricted compartments are unnecessary and contribute little to dMR signals.

The intra-axonal compartment represents contributions from neuronal fibers. A neuronal fiber, a thin and long geometric object, can be abstracted as a stick [Figure 11(a)] that allows only water molecules to move along its longitudinal direction [31] or as a slim cylinder [Figure 11(b)] that allows limited transverse motion [32]. For the cylinder model, fiber diameters can be further weighted with a random distribution, such as the gamma distribution [33], to handle their inhomogeneity. However, the estimation of axon diameters is a difficult issue and possible only with very high gradients [34].

The extra-axonal compartment accounts for hindered diffusion not directly related to neuronal fibers. Because they should

not have a dominant direction, they could be modeled as unbounded homogeneous material [Figure 11(c)] with unknown diffusivity d [31] or an unknown diffusion tensor D [Figure 11(d)] [32].

The highly restricted component is for water molecules trapped in such tiny structures that a magnetic field gradient can hardly affect their diffusion signals. Their diffusion profile can be treated as a small ball [Figure 11(e)] [29] or, more restrictively, a “dot” [Figure 11(f)] that completely prevents water molecules from moving [29], [35].

Figure 12 shows that diffusion models affect reconstruction results significantly. In this example, the stick–ball–dot model solved with convex optimization [35] yields sharper and more stable results.

FOD image processing

Reconstruction of FODs ends at the stage of extracting voxel-wise fiber information from dMR images. Afterward, processing will not directly handle dMR images but, in most cases, directly work with FOD images. Tasks applicable to ordinary scalar-valued images, such as enhancement, segmentation, and registration, can also be performed with FOD images. In this article, we do not discuss these image processing tasks themselves, but focus on a fundamental level: manifold structures of FODs. Manifold structures not only are the driving

force behind many image processing tasks but also need special treatment for FODs.

Many image processing methods are built on interaction between voxel values. For example, in image smoothing, neighboring voxels exert forces on the central one to make its value gradually coherent with theirs. By defining different driving forces, various effects can be achieved [38]. In image registration, voxel-interaction forces between two images drive them to spatially deform and gradually become similar to each other. Even interpolation involves this kind of voxel interaction. The interpolated value can be regarded as the one in balance with the weighted forces from its neighbors. In general, many image processing methods can be formulated as the minimization of an energy function involving interaction forces between voxels.

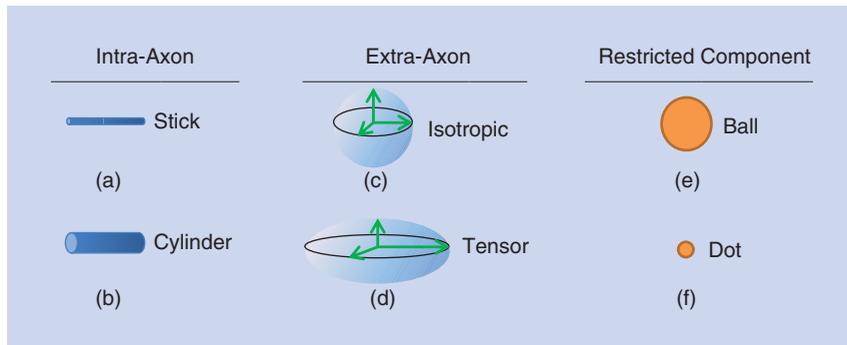


FIGURE 11. (a)–(f) Diffusion models for reconstructing FODs.

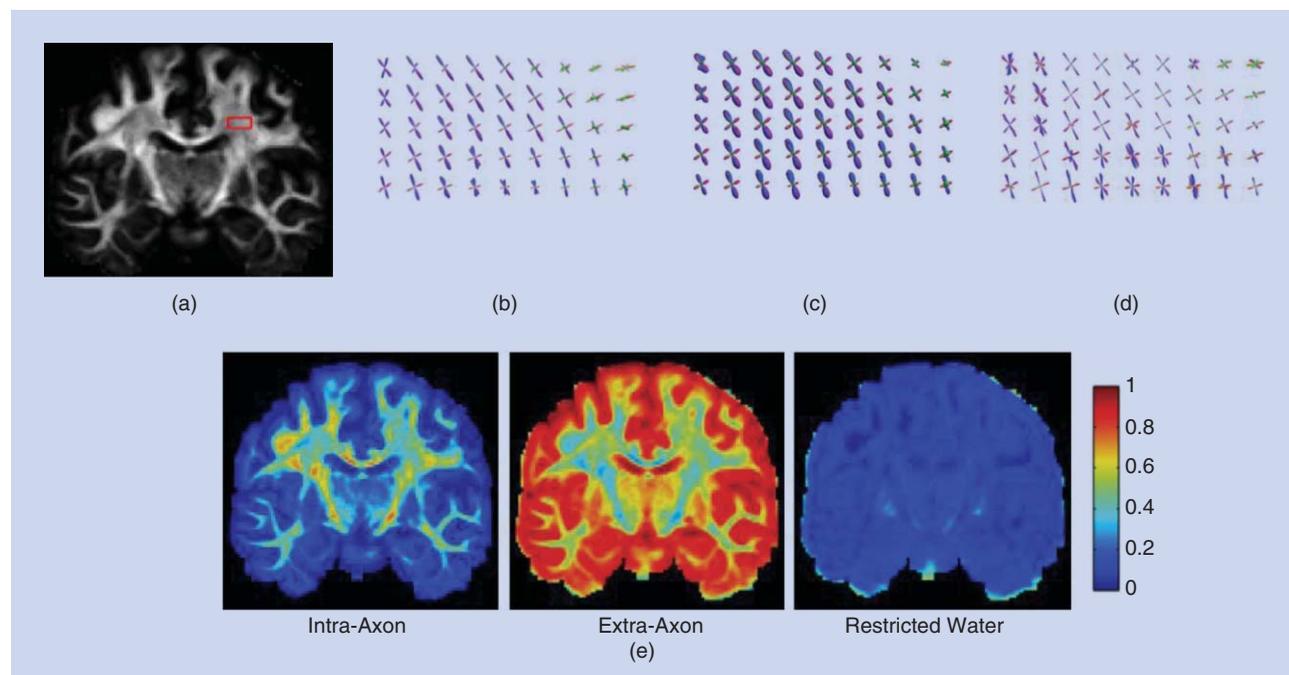


FIGURE 12. Reconstruction of FODs for a small brain region. (a) The region of interest is in the red box. FODs are shown reconstructed with (b) a stick–ball–dot model and convex optimization [35], (c) spherical deconvolution of diffusion distributions [36], and (d) a stick–ball model and Markov chain Monte Carlo [37]. Diffusion models significantly affect reconstruction results. In this example, the model-based and regularized method (b) [35] yields sharper and more stable results. (e) Estimated fraction of the intra-axon, extra-axon, and restricted water components. (Figure reprinted from [35] with permission.)

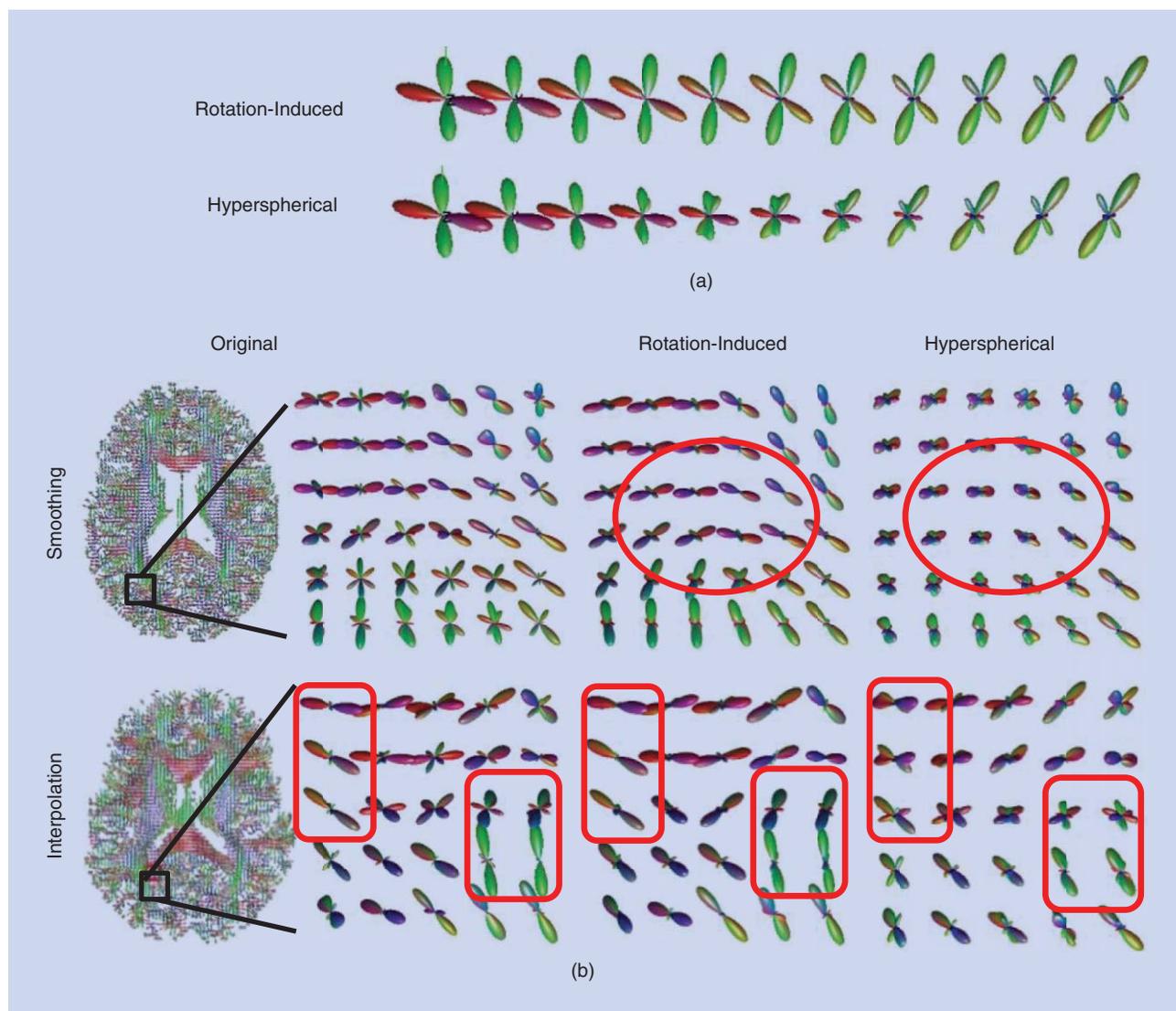


FIGURE 13. (a) Trajectories connecting two FODs on the rotation-induced and hyperspherical manifolds. FOD values in different directions are shown as radius length, and each direction is color coded by linearly combining the red, green, and blue colors according to the x, y, and z components of unit directional vectors. (b) Examples of smoothing (upper row) and interpolating (lower row) FODs. The right column shows the original image and highlighted regions. The middle and left columns show the results of the rotation-induced and hyperspherical manifolds, respectively. The red boxes show that the hyperspherical manifold mixes curving fibers instead of aligning them, introducing the “swelling” effect. (Figure reprinted from [40] with permission.)

Processing FOD images is challenging because each voxel is a spherical function whose mathematical properties are much more complicated than a simple intensity scalar. Even fundamental tasks such as interpolation, smoothing, segmentation, and registration must be redesigned. To use well-developed image processing frameworks, it is crucial to develop suitable manifold structures for FODs.

Log-Euclidean manifold for diffusion tensors

The simplest manifold structure for diffusion tensors is linear. Given two diffusion tensors D_1 and D_2 , represented as symmetric positive definite matrices, their distance is $\|D_1 - D_2\|_2$. However, this structure has a noticeable defect: the linear average of two diffusion tensors may produce a tensor whose

determinant is greater than both of them. The determinant, in the scenario of diffusion, has solid physical meaning: its square root is proportional to the width of the diffusion region. It is undesirable to have the region expanded. As a remedy, in 2006 Arsigny et al. [39] introduced the log-Euclidean manifold [39]: the distance between two diffusion tensors is $\|\log(D_1) - \log(D_2)\|_2$. This manifold essentially maps a diffusion tensor D to its matrix logarithm $\log(D)$. Because D is symmetric and positive definite, $\log(D)$ is unique.

Hyperspherical manifold

The full description of FODs is a nonnegative spherical function $F: S^2 \rightarrow R_{>0}$ whose integral on the sphere equals one. In resemblance to the problem with diffusion tensors, a linear

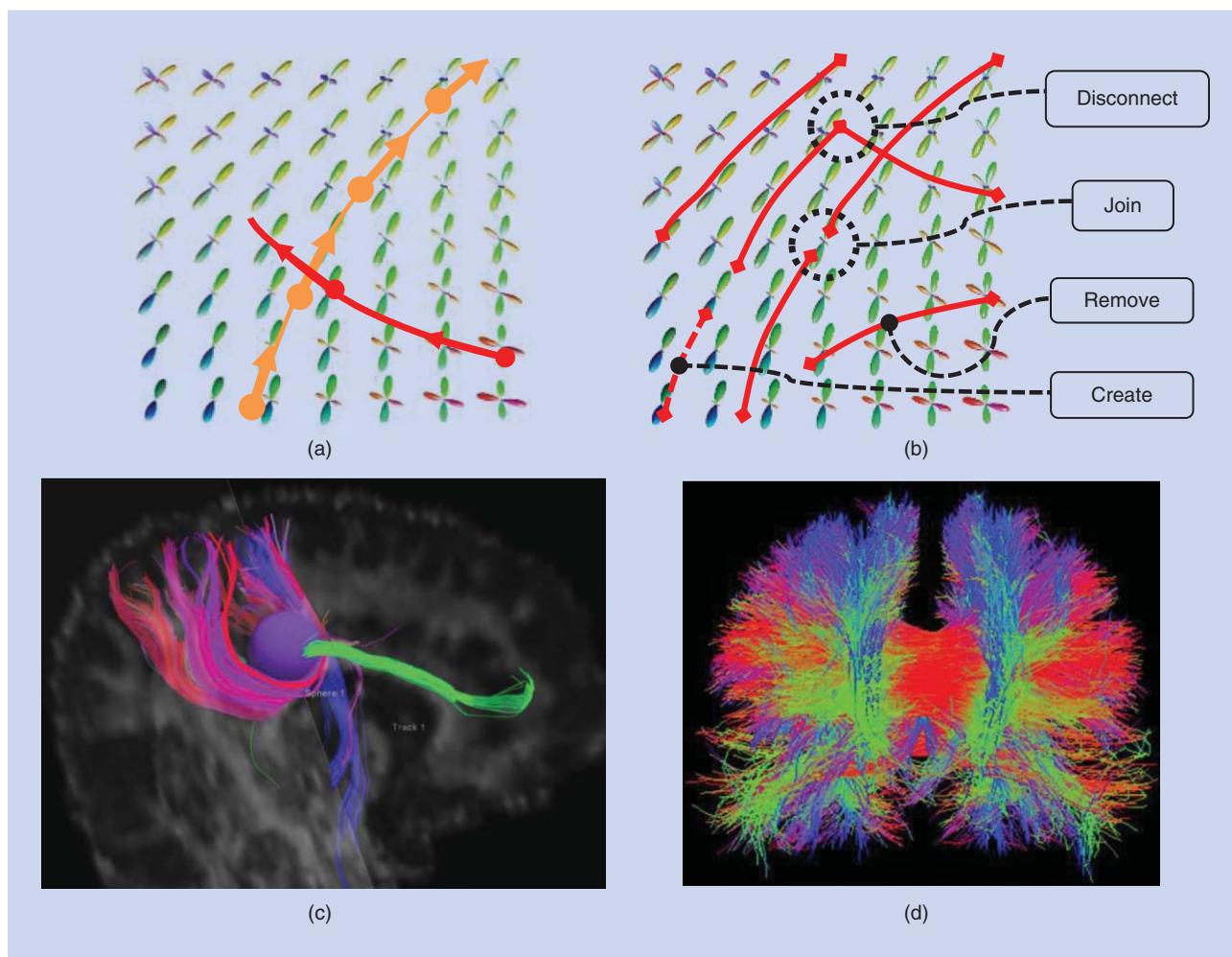


FIGURE 14. Two tractography methods. (a) The local method grows a fiber track step by step, according to local fiber information. (b) The global method improves a set of fiber tracks by disconnecting, joining, removing, and creating fibers with the aim of minimizing an energy function that matches fiber tracks and the FOD image through a generative model. (c) Fiber bundles simulated with a local tractography method seeded in a ball region. (d) An example of fiber tracks simulated by Daducci et al.'s [47] global tractography method. In (c) and (d), fiber tracks are color coded with the directional vector from their start point to their end point, with the red, green, and blue colors for the left–right, back–front, and up–down directions, respectively. [(d) reprinted from [47] with permission.]

combination of spherical functions disperses probability density on the sphere and hence reduces its angular focus. Cheng et al. [41] proposed a hyperspherical manifold of FODs. Because $\int_{u \in S^2} F(u) du = 1$, its square-rooted function $f = \sqrt{F} \geq 0$ satisfies $\int_{u \in S^2} f^2(u) du = 1$, or equivalently, the norm of f is one. This implies that f resides on a unit sphere in a hyperspace, a well-studied manifold. On this manifold, the distance between two FODs is the length of the great arc connecting them, and the interaction force is in the tangent plane of the hypersphere.

Rotation-induced spherical manifold

Although the hyperspherical manifold reduces the angular dispersion of FODs, it does not directly handle the main cause of dispersion: the FODs' difference in orientation. As neuronal fibers extend in the brain, they do not take straight paths, but turn gradually to make curves. Li et al. [40] proposed a rotation-induced manifold to directly handle this ori-

entational difference. The rotation-induced manifold treats tangent vectors on the hyperspherical surface differently, that is, whichever can be realized by rotation is associated with a parameter λ_{rot} , and whichever purely reflects shape differences is associated with a constant one. This unequal treatment does not change the topology, but defines a new Riemannian metric. By setting λ_{rot} smaller than one, the FODs' orientational difference is underweighted, and consequently FODs with similar shapes but different orientations are squeezed closer. If λ_{rot} equals zero, then the manifold reduces to a quotient space completely discarding differences caused by rotation. Figure 13 compares the rotation-induced and hyperspherical manifolds, showing that manifold structures can lead to very different processing effects. Reisert and Kiselev [42] have provided a framework for embedding fiber continuity into FOD reconstruction. It is interesting to utilize these FOD manifolds in their framework.

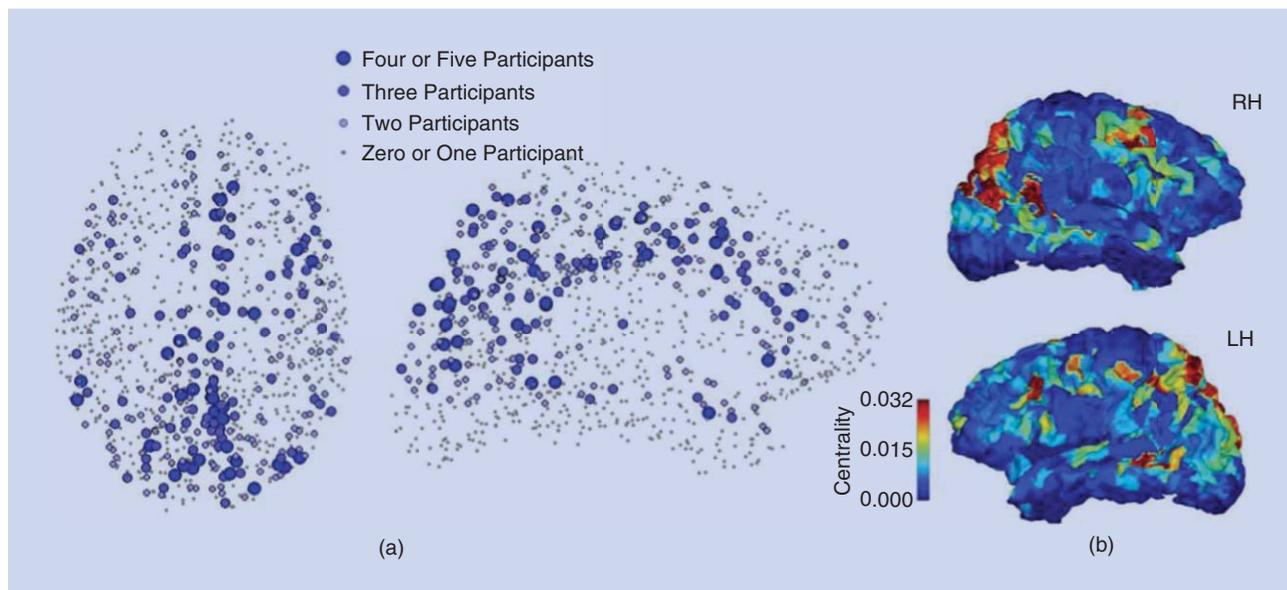


FIGURE 15. The connectivity centrality of brain regions. (a) Region centrality obtained from high-resolution connection matrices of five participants. The plot shows how consistently region centrality ranked in the top 20% among the 998 regions. (b) Lateral views of the right (RH) and left (LH) cerebral hemispheres showing connectivity centrality averaged across all five participants. (Figure reprinted from [48] with permission.)

Simulation of fiber tracks

Based on local fiber orientation, tractography simulates fiber paths stretching from one brain region to another to statistically model geometric and connectional properties of the fiber network. Millions of tracks can be simulated, grouped as bundles, and later analyzed with graph theory. It should be noted that simulated fiber tracks do not represent real neuronal fibers, but statistically reflect their spatial and orientational distributions. A single simulated fiber track does not have much realistic meaning, but statistics derived from a huge number of tracks are useful.

Tractography methods can be classified into two broad categories: local methods and global methods (Figure 14). Local methods “grow” a neuronal fiber from a seed step by step with the guidance of local fiber orientation. Usually, it becomes the simulation of an ordinary differential equation (ODE): $dr/dt = G(r)$, where r is the current location of a fiber point, $G(r)$ is the fiber direction at this point, and dr/dt is the growing velocity of the fiber. $G(r)$ can be either deterministic or stochastic [37], [43]. If it is deterministic, it usually follows the FOD’s peak closest to the previous track direction. In this case, a fiber track is determined by its initial seed location and direction. If it is stochastic, the direction is randomly sampled according to the FOD. Probabilistic tractography explores more possibilities than deterministic ones. ODEs of higher orders or more sophisticated models can be used. To reduce errors, some methods [44] employ the fourth-order Runge–Kutta method instead of the first-order explicit Euler method. Local tractography algorithms usually run quite fast, but they share a common problem: local errors accumulate and propagate. To regularize the problem, most algorithms

impose an upper bound on the curvature of their fiber tracks to avoid sharp turns.

Another approach treats tractography as an inverse problem: finding a set of fiber tracks that generate signals to match the dMR image [45]. The advantage of this method is obvious: to optimize an objective function, fiber tracks interact with each other, which may lead to more stable results and may also avoid repetitively sampling similar fibers. However, it is challenging to jointly solve such a problem with a huge number of fiber tracks. Therefore, it usually takes much longer computation time than local methods. In 2011, Reisert et al. [46] reduced the computation time to a practical range of several hours for about 10^5 fiber tracks. They formulated an energy function

$$E = E(\text{Tracks}) + E(\text{Tracks}, \text{Data}),$$

where $E(\text{Tracks})$ prefers smooth and long tracks, and $E(\text{Tracks}, \text{Data})$ is the difference between track-generated signals and the real dMR image. The energy function is minimized with simulated annealing instead of a deterministic method. Each time, an operation on track segments is randomly proposed: creation, removal, move, join, and so on. The proposed operation is then accepted or rejected stochastically based on the energy change it introduces. Daducci et al. [47] implemented this global tractography method into a software package called Connectome Mapper.

Applications to brain research

Connectivity information captured by dMR imaging has been investigated to understand the brain’s structure and function and its relationship with neurological disorders. A study on five healthy young men conducted by Hagmann et al. [48] revealed

that brain regions are not connected equally, but some play more central roles than others (Figure 15). For each subject, they simulated nearly 3 million fiber tracks with a fixed step size of 1 mm and then derived a connectivity matrix between 998 regions on the cortex surface, each of approximately 1.5 cm². On the basis of graph theory, they found that brain regions within the posterior medial parietal cerebral cortex and several temporal and frontal lobe areas form a highly mutually connected network and constitute hubs linking other regions. The functional activities of these regions are also highly coupled when the brain is at rest.

From childhood to adulthood, the brain experiences profound development to reach its peak of intelligence and mental capacity. Comparing the brains of 439 individuals aged 12–30 years, Dennis et al. [49] found that not all connections are strengthened during the development, but some are “pruned.” They scanned the subjects with high angular resolution diffusion imaging, reconstructed their fiber networks, and analyzed the networks with graph theory and linear regression regarding the subjects’ gender and age. It was found that fiber density relating to the frontal cortex decreases, but that relating to the temporal cortex increases, as shown in Figure 16.

Many mental disorders are related to abnormal functional integration caused by aberrant brain connectivity. Using

diffusion tensor imaging and tractography, Zalesky et al. [50] compared the anatomic connectivity network of 74 schizophrenia patients with 32 controls matched in age and gender. They found statistically significant differences in connectivity involving the medial frontal parieto-occipital lobe and the left temporal lobe between the patients and controls.

Future scope

Advances in dMR imaging have provided a platform for investigating brain connectivity in vivo at unprecedented spatial and angular resolutions. Current acquisition and analyzing techniques have not reached their full potential. Under active development, they will be more efficient, more accurate, and more reliable in the established framework.

As these techniques are becoming more accessible, it is important to explore their clinical applications, such as brain surgical planning and prognosis of recovery from trauma, stroke, and so on. A connectivity network reconstructed from a patient’s dMR images may provide valuable information for surgeons to more precisely locate the spot of intervention or for doctors to better predict the mental impact of trauma or hemorrhage. A single imaging modality such as dMR might be insufficient to explain underlying physiological or pathological changes. To obtain a more comprehensive understanding of

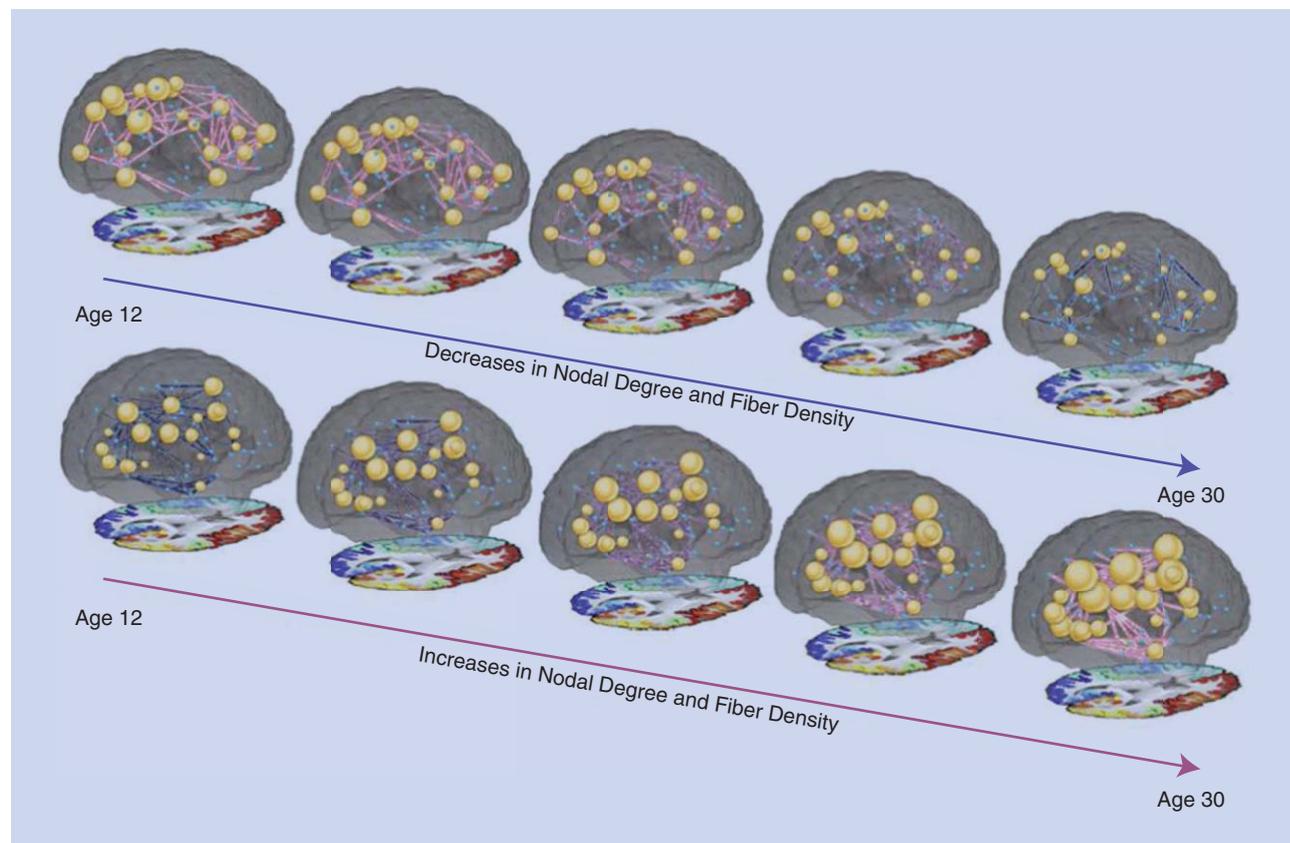


FIGURE 16. Dennis et al. [49] compared the brains of 439 individuals aged 12–30 years by high angular resolution diffusion imaging and found that not all connections are strengthened during development, but some are pruned. Only the connections with significant correlations with age are shown. The node size is proportional to the number of connections, and the thickness of the connection edges is proportional to relative fiber density. (Figure reprinted from [49] with permission.)

these changes, it is important to take advantage of multimodal imaging data such as functional MR imaging and positron emission tomography for biochemical and metabolic information. In return, these clinical applications will inspire more engineering and methodological innovations.

Because dMR imaging provides *in vivo* and indirect views of a brain's structural connectivity at a macroscopic level, it is important to verify and further investigate its relationship with microscopic images. Referring to high-resolution images at the neuron level may help identify signal signatures related to specific tissues. In 2013, the breakthrough CLARITY technology by Chung et al. [51] made it possible to optically create a three-dimensional image of an intact brain structure at the resolution of individual neurons. CLARITY removes light-blocking fatty membranes from the brain and makes it transparent. With high-resolution microcopies, it can image a brain's long-range projections, local circuit wiring, cellular relationships, subcellular structures, protein complexes, nucleic acids, and neurotransmitters. Although dMR imaging cannot achieve such super-resolution, CLARITY does not replace it because CLARITY is a postmortem technique and unavailable for *in vivo* diagnosis. As an *in vivo* and macroscopic technology, dMR imaging can be further investigated with microscopic references to improve its value for clinical diagnosis.

A comprehensive and multiscale description of brain connectivity, function, and development requires integrating information from multiple imaging modalities, functional tasks, and genetic data. Such challenging integration will actively adopt cutting-edge technologies in data science. A connectivity map on the cortical surface forms a four-dimensional space, though sparse. Correlating with genomic data [52] introduces one more dimension with numerous genetic alleles. Overlaying functional tasks brings up even more varieties. Such explosive complexity has pushed researchers to seek solutions beyond classical regression models to recent achievements in machine learning, for instance, deep learning.

The need for standardization will arise as our knowledge about the brain connectome continues to expand. The fusion of complex information will naturally lead to a more detailed and specialized definition of brain regions and fiber bundles, as shown in Zhu et al.'s article "DICCCOL: Dense Individualized and Common Connectivity-Based Cortical Landmarks" [53]. A systematic and precise naming catalog for anatomic structures and functional tasks will be needed for researchers around the world to report and exchange their results in shared databases. Extendable digital storage formats to accommodate complex connectivity information and processing protocols are also important for experiment replication and validation. A standardized neuroinformatics platform will boost collaboration and lead to large-scale investigations as witnessed in the field of genomics.

Acknowledgments

This work is supported by grants P41EB015922, U54EB020406, R01MH094343, and K01EB013633 from the National Institutes of Health.

Authors

Junning Li (Junning.Li@loni.usc.edu) earned his B.S. and M.S. degrees in automation from Tsinghua University, Beijing, in 2002 and 2005, respectively, and his Ph.D. degree in electrical and computer engineering from the University of British Columbia, Vancouver, Canada, in 2009. In 2011, he joined the Laboratory of Neuro Imaging at the University of California, Los Angeles, as a postdoctoral fellow. In 2013, he became an assistant professor in the Mark and Mary Stevens Neuroimaging and Informatics Institute at the University of Southern California. He develops image processing and statistical methods for modeling brain connectivity with functional and diffusion magnetic resonance.

Yonggang Shi (Yonggang.Shi@loni.usc.edu) earned his B.S. and M.S. degrees in electrical engineering from Southeast University, Nanjing, China, in 1996 and 1999, respectively, and his Ph.D. degree in electrical engineering from Boston University in 2005, at which time he joined the Laboratory of Neuro Imaging at the University of California, Los Angeles, as a postdoctoral fellow and became an assistant professor in 2009. He is an assistant professor of neurology and electrical engineering in the Mark and Mary Stevens Neuroimaging and Informatics Institute at the University of Southern California. He is an associate editor of *IEEE Transactions on Image Processing*. He was a winner of the Student Paper Award in the 2005 International Conference on Acoustics, Speech, and Signal Processing student paper competition and also won the Best Paper Award at the 2008 IEEE Workshop on Mathematical Methods in Biomedical Image Analysis.

Arthur W. Toga (toga@loni.usc.edu) earned his B.S. degree in psychology from the University of Massachusetts in 1974 and his M.S. and Ph.D. degrees in neuroscience from St. Louis University in 1976 and 1978, respectively. He is the director of the University of Southern California Mark and Mary Stevens Neuroimaging and Informatics Institute. Previously, he was a distinguished professor of neurology at the University of California, Los Angeles, and the director of the Laboratory of Neuro Imaging. His research is focused on neuroimaging, informatics, and brain atlas. He has written more than 1,000 papers, chapters, or abstracts, including eight books. He has received numerous awards and honors in computer science, graphics, and neuroscience. He is the founding editor-in-chief of *NeuroImage* and has held the chair of numerous committees within the National Institutes of Health and many international task forces.

References

- [1] NIH Blueprint for Neuroscience Research. The human connectome project. [Online]. Available: <http://www.neuroscienceblueprint.nih.gov/connectome/>
- [2] FreeSurfer. (2013). FreeSurfer Software Suite. [Online]. Available: <http://freesurfer.net/>
- [3] P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford, U.K.: Oxford Univ. Press, 1991.
- [4] P. Mansfield and P. K. Grannell, "NMR 'diffraction' in solids?" *J. Phys. C, Solid State Phys.*, vol. 6, no. 22, pp. L422–L426, Nov. 1973.
- [5] R. S. Likes, "Moving gradient zeugmatography," U.S. Patent US4307343A, 1981.
- [6] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient," *J. Chem. Phys.*, vol. 42, no. 1, pp. 288–292, 1965.

- [7] E. L. Hahn, "Spin echoes," *Phys. Rev.*, vol. 80, no. 4, pp. 580–594, Nov. 1950.
- [8] P. Bassler, J. Mattiello, and D. Le Bihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, no. 1, pp. 259–267, Jan. 1994.
- [9] P. Mansfield and A. A. Maudsley, "Medical imaging by NMR," *Br. J. Radiol.*, vol. 50, no. 591, pp. 188–194, Mar. 1977.
- [10] S. N. Sotiropoulos, S. Jbabdi, J. Xu, J. L. Andersson, S. Moeller, E. J. Auerbach, M. F. Glasser, M. Hernandez, et al., "Advances in diffusion MRI acquisition and processing in the human connectome project," *NeuroImage*, vol. 80, pp. 125–143, Oct. 2013.
- [11] P. Mansfield, "Multi-planar image formation using NMR spin echoes," *J. Phys. C, Solid State Phys.*, vol. 10, no. 3, pp. L55–L58, Feb. 1977.
- [12] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, Nov. 1999.
- [13] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1202–1210, June 2002.
- [14] D. J. Larkman, J. V. Hajnal, A. H. Herlihy, G. A. Coutts, I. R. Young, and G. Ehnholm, "Use of multicoil arrays for separation of signal from multiple slices simultaneously excited," *J. Magn. Reson. Imag.*, vol. 13, no. 2, pp. 313–317, 2001.
- [15] S. Moeller, E. Yacoub, C. A. Olman, E. Auerbach, J. Strupp, N. Harel, and K. Ugurbil, "Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI," *Magn. Reson. Med.*, vol. 63, no. 5, pp. 1144–1153, Apr. 2010.
- [16] M. I. Menzel, E. T. Tan, K. Khare, J. I. Sperl, K. F. King, X. Tao, C. J. Hardy, and L. Marinelli, "Accelerated diffusion spectrum imaging in the human brain using compressed sensing," *Magn. Reson. Med.*, vol. 66, no. 5, pp. 1226–1233, Aug. 2011.
- [17] O. Michailovich, Y. Rathi, and S. Dolui, "Spatially regularized compressed sensing for high angular resolution diffusion imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1100–1115, May 2011.
- [18] B. Bilgic, K. Setsompop, J. Cohen-Adad, A. Yendiki, L. L. Wald, and E. Adalsteinsson, "Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries," *Magn. Reson. Med.*, vol. 68, no. 6, pp. 1747–1754, Sept. 2012.
- [19] L. Czervionke, J. Czervionke, D. Daniels, and V. Haughton, "Characteristic features of MR truncation artifacts," *Am. J. Roentgenol.*, vol. 151, no. 6, pp. 1219–1228, Dec. 1988.
- [20] K. T. Block, M. Uecker, and J. Frahm, "Suppression of MRI truncation artifacts using total variation constrained data extrapolation," *Int. J. Biomed. Imag.*, vol. 2008, 184123, 2008.
- [21] P. Jezzard and R. S. Balaban, "Correction for geometric distortion in echo planar images from b0 field variations," *Magn. Reson. Med.*, vol. 34, no. 1, pp. 65–73, July 1995.
- [22] J. L. Andersson, S. Skare, and J. Ashburner, "How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging," *NeuroImage*, vol. 20, no. 2, pp. 870–888, Oct. 2003.
- [23] T. Reese, O. Heid, R. Weisskoff, and V. Wedeen, "Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo," *Magn. Reson. Med.*, vol. 49, no. 1, pp. 177–182, Dec. 2002.
- [24] J. E. Sarlls, C. Pierpaoli, S. L. Talagala, and W.-M. Luh, "Robust fat suppression at 3T in high-resolution diffusion-weighted single-shot echo-planar imaging of human brain," *Magn. Reson. Med.*, vol. 66, no. 6, pp. 1658–1665, May 2011.
- [25] K. Ugurbil, J. Xu, E. J. Auerbach, S. Moeller, A. T. Vu, J. M. Duarte-Carvajalino, C. Lenglet, X. Wu, et al., "Pushing spatial and temporal resolution for functional and diffusion MRI in the human connectome project," *NeuroImage*, vol. 80, pp. 80–104, Oct. 2013.
- [26] F. S. Mikelberg, S. M. Drance, M. Schulzer, H. M. Yidegiligne, and M. M. Weis, "The normal human optic nerve," *Ophthalmology*, vol. 96, no. 9, pp. 1325–1328, Sept. 1989.
- [27] E. Ozarslan and T. H. Mareci, "Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging," *Magn. Reson. Med.*, vol. 50, no. 5, pp. 955–965, Oct. 2003.
- [28] J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly, "Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution," *NeuroImage*, vol. 23, no. 3, pp. 1176–1185, Nov. 2004.
- [29] E. Panagiotaki, T. Schneider, B. Siow, M. G. Hall, M. F. Lythgoe, and D. C. Alexander, "Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison," *NeuroImage*, vol. 59, no. 3, pp. 2241–2254, Feb. 2012.
- [30] B. Dhital, E. Kellner, M. Reisert, and V. G. Kiselev, "Isotropic diffusion weighting provides insight on diffusion compartments in human brain white matter in vivo," presented at the 23rd Annu. Meeting Int. Society for Magnetic Resonance in Medicine, Toronto, May–June 2015.
- [31] T. Behrens, M. Woolrich, M. Jenkinson, H. Johansen-Berg, R. Nunes, S. Clare, P. Matthews, J. Brady, et al., "Characterization and propagation of uncertainty in diffusion-weighted MR imaging," *Magn. Reson. Med.*, vol. 50, no. 5, pp. 1077–1088, Oct. 2003.
- [32] D. C. Alexander, "A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features," *Magn. Reson. Med.*, vol. 60, no. 2, pp. 439–448, Aug. 2008.
- [33] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P. J. Bassler, "Axcaliber: A method for measuring axon diameter distribution from diffusion MRI," *Magn. Reson. Med.*, vol. 59, no. 6, pp. 1347–1354, June 2008.
- [34] S. Y. Huang, A. Nummenmaa, T. Witzel, T. Duval, J. Cohen-Adad, L. L. Wald, and J. A. McNab, "The impact of gradient strength on in vivo diffusion MRI estimates of axon diameter," *NeuroImage*, vol. 106, pp. 464–472, Feb. 2015.
- [35] G. Tran and Y. Shi, "Fiber orientation and compartment parameter estimation from multi-shell diffusion imaging," *IEEE Trans. Med. Imag.*, vol. 34, no. 11, pp. 2320–2332, Nov. 2015.
- [36] F.-C. Yeh, V. J. Wedeen, and W.-Y. I. Tseng, "Estimation of fiber orientation and spin density distribution by diffusion deconvolution," *NeuroImage*, vol. 55, no. 3, pp. 1054–1062, Apr. 2011.
- [37] T. Behrens, H. J. Berg, S. Jbabdi, M. Rushworth, and M. Woolrich, "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?" *NeuroImage*, vol. 34, no. 1, pp. 144–155, Jan. 2007.
- [38] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, pp. 629–639, July 1990.
- [39] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-Euclidean metrics for fast and simple calculus on diffusion tensors," *Magn. Reson. Med.*, vol. 56, no. 2, pp. 411–421, Aug. 2006.
- [40] J. Li, Y. Shi, and A. W. Toga, "Diffusion of fiber orientation distribution functions with a rotation-induced Riemannian metric," *Med. Image Comput. Comput. Assist. Interv.*, vol. 17, pt. 3, pp. 249–256, 2014.
- [41] J. Cheng, A. Ghosh, T. Jiang, and R. Deriche, "A Riemannian framework for orientation distribution function computing," *Med. Image Comput. Comput. Assist. Interv.*, vol. 12, pt. 1, pp. 911–918, 2009.
- [42] M. Reisert and V. G. Kiselev, "Fiber continuity: An anisotropic prior for ODF estimation," *IEEE Trans. Med. Imag.*, vol. 30, no. 6, pp. 1274–1283, June 2011.
- [43] M. Descoteaux, R. Deriche, T. Knosche, and A. Anwander, "Deterministic and probabilistic tractography based on complex fibre orientation distributions," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 269–286, Feb. 2009.
- [44] P. J. Bassler, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "In vivo fiber tractography using DT-MRI data," *Magn. Reson. Med.*, vol. 44, no. 4, pp. 625–632, Oct. 2000.
- [45] J.-F. Mangin, P. Fillard, Y. Cointepas, D. Le Bihan, V. Frouin, and C. Poupon, "Toward global tractography," *NeuroImage*, vol. 80, pp. 290–96, Oct. 2013.
- [46] M. Reisert, I. Mader, C. Anastasopoulos, M. Weigel, S. Schnell, and V. Kiselev, "Global fiber reconstruction becomes practical," *NeuroImage*, vol. 54, no. 2, pp. 955–962, Jan. 2011.
- [47] A. Daducci, S. Gerhard, A. Griffa, A. Lemkaddem, L. Cammoun, X. Gigandet, R. Meuli, P. Hagmann, et al., "The connectome mapper: An open-source processing pipeline to map connectomes with MRI," *PLoS ONE*, vol. 7, no. 12, e48121, Dec. 2012.
- [48] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *PLoS Biol.*, vol. 6, no. 7, e159, July 2008.
- [49] E. L. Dennis, N. Jahanshad, K. L. McMahon, G. I. de Zubicaray, N. G. Martin, I. B. Hickie, A. W. Toga, M. J. Wright, et al., "Development of brain structural connectivity between ages 12 and 30: A 4-tesla diffusion imaging study in 439 adolescents and adults," *NeuroImage*, vol. 64, pp. 671–684, Jan. 2013.
- [50] A. Zalesky, A. Fornito, M. L. Seal, L. Cocchi, C.-F. Westin, E. T. Bullmore, G. F. Egan, and C. Pantelis, "Disrupted axonal fiber connectivity in schizophrenia," *Biol. Psychiatry*, vol. 69, no. 1, pp. 80–89, Jan. 2011.
- [51] K. Chung, J. Wallace, S.-Y. Kim, S. Kalyanasundaram, A. S. Andalman, T. J. Davidson, J. J. Mirzabekov, K. A. Zalocusky, et al., "Structural and molecular interrogation of intact biological systems," *Nature*, vol. 497, no. 7449, pp. 332–337, Apr. 2013.
- [52] N. Jahanshad, P. Rajagopalan, X. Hua, D. P. Hibar, T. M. Nir, A. W. Toga, C. R. Jack, A. J. Saykin, et al., "Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity," *Proc. Nat. Acad. Sci.*, vol. 110, no. 4, pp. 4768–4773, Mar. 2013.
- [53] D. Zhu, K. Li, L. Guo, X. Jiang, T. Zhang, D. Zhang, H. Chen, F. Deng, et al., "DICCOCOL: Dense individualized and common connectivity-based cortical landmarks," *Cereb. Cortex*, vol. 23, no. 4, pp. 786–800, Apr. 2012.

Time-Varying Brain Connectivity in fMRI Data

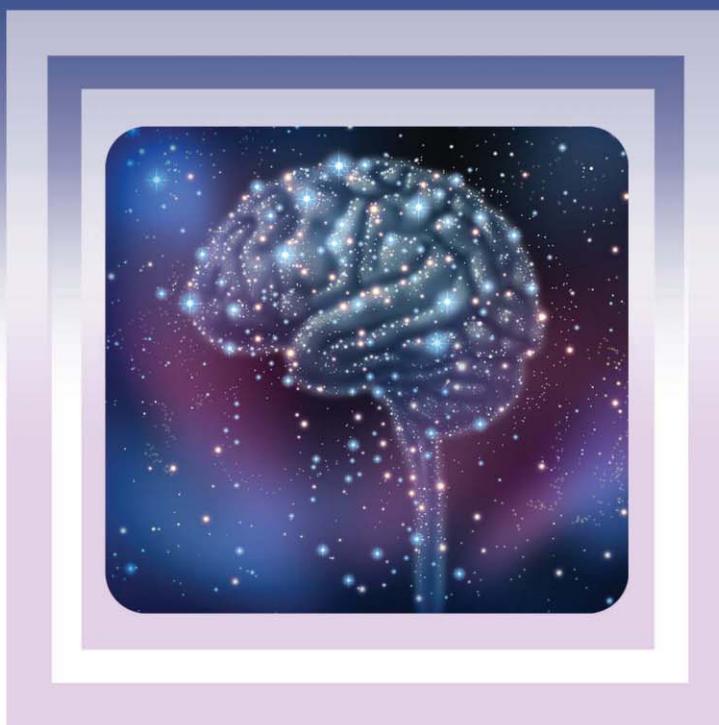


IMAGE LICENSED BY INGRAM PUBLISHING

Whole-brain data-driven approaches for capturing and characterizing dynamic states

Vince D. Calhoun and Tülay Adalı

The study of whole-brain functional brain connectivity with functional magnetic resonance imaging (fMRI) has been based largely on the assumption that a given condition (e.g., at rest or during a task) can be evaluated by averaging over the entire experiment. In actuality, the data are much more dynamic, showing evidence of time-varying connectivity patterns, even within the same experimental condition. In this article, we review a family of blind-source separation (BSS) approaches that have proven useful for studying time-varying patterns of connectivity across the whole

brain. Initial work in this direction focused on time-varying coupling among data-driven nodes, but more recently, time-varying nodes have also been considered. We also discuss extensions of these approaches, including transformations into the time-frequency domain. We provide a rich set of examples of various applications that yielded new information about healthy and diseased brains. Due in large part to developments in the field of signal processing, the fMRI community has seen major growth in the development of approaches that can capture whole-brain systemic connectivity information (connectomics) while also allowing this system to evolve over time as it naturally does (i.e., chronnectomics).

Digital Object Identifier 10.1109/MSP.2015.2478915
Date of publication: 27 April 2016

Introduction

Big data [1], the human connectome [2], the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative [3], and the chronnectome [4] are examples of the major ongoing movement in the United States, European Union, China, and Japan to understand the human brain and brain connectivity [cf., the multiagency Human Connectome Project [2], [3] and BRAIN Initiative, the National Science Foundation's (NSF's) neural and cognitive systems program, the European Union's Human Brain Project, Japan's Brain/MINDS, and others]. Two critical themes across all of these projects are 1) the technology development for studying the human brain and 2) harnessing and adapting ideas resulting from large-scale projects such as the Human Genome Project and the Big Data to Knowledge Initiative. These themes continue to revolutionize our understanding of the complexity of the human brain and are driving the recent focus on scaling science to handle “big data” problems.

The study of changes in brain networks (functional connectomics [5], [6]) over time, termed the *chronnectome*, was recently highlighted as one of the “Best of 2014” by Director Dr. Tom Insel [7] of the National Institute of Mental Health in terms of a concept that brought engineers, physicists, and neurobiologists together to better understand the temporal dynamics of brain-imaging signals. Specifically, Dr. Insel noted the power of the “convergence” or merger of multiple disciplines [7], [8].

The chronnectome is a model of the brain in which nodal activity and connectivity patterns are changing in predictable and meaningful ways through time [4], [9]. Thus, the concept of the chronnectome makes the specific assumption that the dynamics are nonstationary in interesting ways. One can focus on chronnectomic changes at various scales, including milliseconds [as measured by electroencephalogram (EEG) or magnetoencephalogram (MEG)], seconds (as measured by fMRI), and minutes (as measured by changes between experiments using average or static connectivity approaches [10]–[12]) and changes over months and years (at which point, incorporating additional information such as brain structure or epigenetic changes becomes very useful for longitudinal studies) [Figure 1(a)]. Characterization of brain connectivity across the life span is a major priority in both the Human Connectome Project [5] and the BRAIN Initiative [3].

One of the earliest examples of time-varying connectivity is the concept of EEG microstates, or points in time during which there is a common synchrony across multiple brain regions [13]–[17]. More recently, fMRI, which provides a more spatially specific measure of function across the entire brain (at the cost of decreased temporal resolution), has been used to study time-varying connectivity. In this article, we focus primarily on the recent emergence of data-driven approaches that can capture whole-brain patterns of time-varying connectivity within one fMRI experiment (either at rest or during a task). Initial results suggest that such an approach provides more

information than static connectivity approaches, thus motivating methods that acknowledge the dynamically changing brain within a single experiment. Such approaches will likely make the evaluation of connectivity changes over longer timescales even more informative.

There has been great progress in the use of functional connectivity measures to study healthy and diseased brains, and whole-brain measures have proven extremely powerful. The fMRI community has now realized that the assessment of functional connectivity has been limited by an implicit assumption of spatial and temporal stationarities throughout the measurement period [18]. Dynamics are potentially even more prominent in the resting state, during which mental activity is unconstrained [19]. The development or adaptation of approaches to studying time-varying connectivity in the brain has emerged along multiple lines, including the detection of important transition points (e.g., change-point analysis [20]), time-frequency approaches [21], and windowing approaches [22]–[24].

Data-driven approaches, in particular (joint) BSS, have proven useful for taking advantage of the available prior and statistical information to fully characterize both static and dynamic brain connectivity [25], [26]. The term chronnectome describes a focus on identifying time-varying but reoccurring patterns of coupling among brain regions. The chronnectome—in contrast to another interesting concept called the *dynome*, which is focused on time-varying (oscillatory) activity whose basic characteristics (frequency, phase, amplitude, etc.) are generally assumed to be static [27]—makes the specific assumption that the dynamics are nonstationary in interesting ways. In the context of this article, *dynamics* refers to intrinsic nonstationarities rather than to dynamics in its mathematical sense. A number of approaches in this respect are revealing exciting new information about the brain, including information about sleep states [28] and disease [29], and represent a much more natural way to analyze brain-imaging data, especially those that are largely unconstrained, such as resting-state fMRI data.

A high-level summary of the key steps for capturing whole-brain, data-driven, time-varying connectivity is presented in Figure 1(b). Input to the analysis can consist of timecourses from regions or from networks (e.g., component timecourses). Next, timecourse pairs can be analyzed using a fixed-window or adaptive windowing approach [23], [30] or a time-frequency approach [21], [31]. The next step involves estimating the states, which can be done a number of ways, for example, by k -means clustering [30], principal component analysis (PCA) [26], [32], or independent component analysis (ICA) [33], [34]. Finally, the summary measure of the states can be done for each state separately, for example, by dwell time or connectivity within each state matrix [29], [35] or across all states, such as in a metastate approach [33], [34].

In contrast to previous reviews [4], [18], the focus of this article is to review whole-brain, data-driven approaches from a signal processing perspective. In addition, we offer

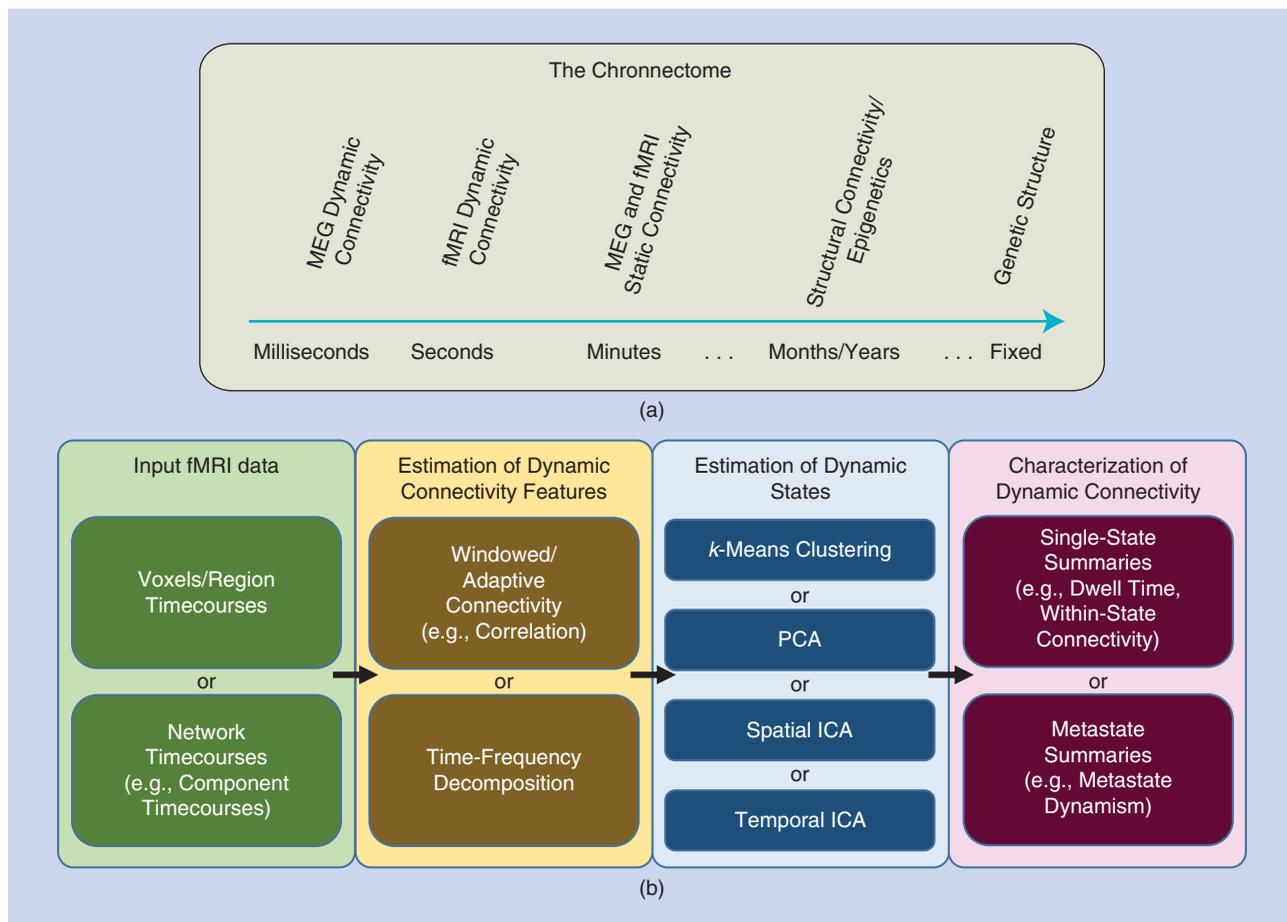


FIGURE 1. (a) The chronnectome concept of studying connectivity at multiple timescales [4]. (b) An overview of some of the key steps and options used in computing time-varying connectivity measures.

new high-level summaries of the various steps in capturing time-varying connectivity [Figure 1(b)], new approaches (e.g., whole-brain time-frequency analyses), new strategies for modeling (e.g., subspace analysis and dynamic model-based connectivity), and new application examples (e.g., results from the EEG/fMRI sleep study and the substance-use study).

Feature generation

One key challenge for studying time-varying connectivity in the brain is generating the features that capture the time-varying dynamics. Approaches include those that make use of a priori information, for example, picking a pair of brain regions (seeds) or using a whole-brain, predefined atlas of regions in fMRI data as well as data-driven approaches. Data-driven approaches include sparsity-based parcellation [36] and latent variables analysis methods, such as PCA, group ICA [37], spatially constrained ICA [38], independent vector analysis (IVA) [25], and tensor decompositions [39]. For example, in [40] and [41], the first event-related potentials in EEG data were detected and then summarized using PCA of time-dependent node correlation matrices. On the other hand, for fMRI data, decompositions that use ICA and IVA can be

adapted to extract dynamic features in multiple ways as demonstrated in, e.g., [30], [34], and [42].

ICA

ICA is based on the assumption that the observations are a linearly mixed set of independent sources/components, an assumption that allows identification of the original sources subject to only scaling and permutation ambiguities and under rather mild conditions for identifiability. If we consider the simple linear mixing model $\mathbf{x}(\nu) = \mathbf{A}\mathbf{s}(\nu)$, $1 \leq \nu \leq V$, $\mathbf{x}(\nu), \mathbf{s}(\nu) \in \mathcal{R}^N$, where ν is the sample index such as voxel, pixel, or time and the mixing matrix \mathbf{A} is full rank, we can obtain the independent component estimates $\mathbf{u}(\nu) = \mathbf{W}\mathbf{x}(\nu)$ by estimating a demixing matrix \mathbf{W} through optimization of an appropriate cost-measuring independence [25], [43].

ICA has proven very useful for fMRI data analysis and can be performed in two different ways [44], [45]: spatial ICA, which extracts independent spatial maps, and temporal ICA, which extracts independent timecourses by considering the transposed version of the data matrix. Spatial ICA is more widely used because the spatial independence assumption is better suited for the systematically nonoverlapping

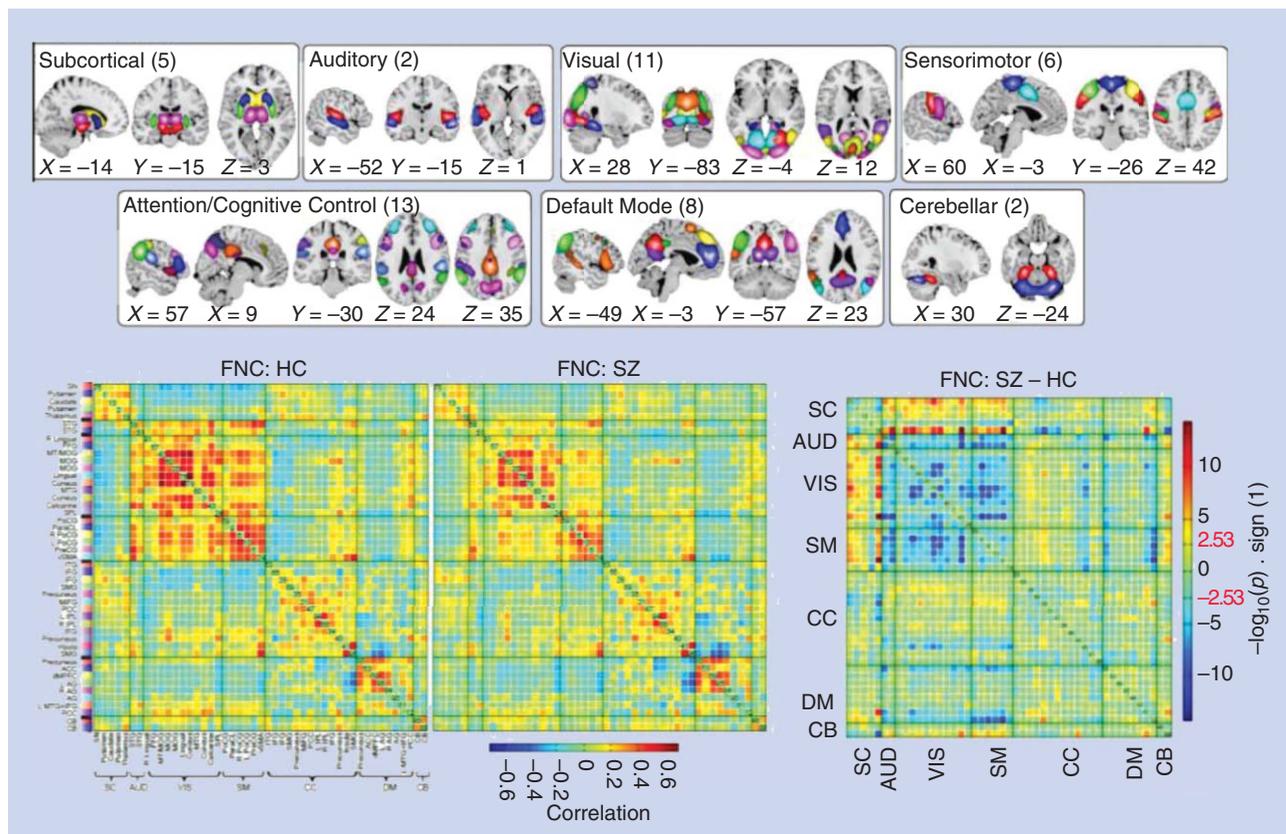


FIGURE 2. Data-driven maps from group ICA provide components that capture information about within-network (within-component) connectivity that is characterized by timecourses that can be used to assess FNC or among-network connectivity, which can be assessed in the simplest manner by computing the cross-correlation among component timecourses. Results are shown for HCs and patients with SZ (reprinted with permission from [29]). The numbers in parentheses indicate the number of components that were included in each group. The X, Y, and Z coordinates of the slice in millimeters in the Montreal Neurological Institute coordinate system are shown.

nature of the spatial patterns [46]. For spatial ICA, the data matrix \mathbf{X} is formed by flattening a given slice at time t as a row, such that \mathbf{X} is time points by voxels ($T \times V$) and dimension T is typically reduced to N using PCA prior to ICA.

In the group ICA model [37], [47], which has been implemented in the Group ICA of fMRI toolbox (GIFT) (<http://mialab.mrn.org/software/gift>), there are double-dimension reduction stages using PCA where the first step is to perform a subject-level PCA, and after vertical concatenation of dimension-reduced subject data, a second-level PCA is applied at the group level to estimate a common group subspace [48]. Individual subject maps are then reconstructed using the group- and subject-level PCA matrices, preserving most of the variability for individual subjects. Other implementations and uses of the group ICA model are also possible and are discussed in [37]. An example of the traditional use of group ICA is shown in Figure 2. The spatial maps are characterized by a single timecourse and provide information about the degree to which each voxel is linearly related to that timecourse. As such, it informs us about within-network connectivity. In the figure, components are divided into anatomical domz indicated within a domain by different colors. Relationships among the timecourses (matrix in Figure 2) capture

the functional network connectivity (FNC) or among-network connectivity. The matrix indicates the degree to which each component is correlated with the other components. Correlations are positive values (red), and anticorrelations are represented as negative values (blue). Results are shown for healthy controls (HCs) and patients with schizophrenia (SZ) [29]. Some approaches have also attempted to combine aspects of both spatial and temporal ICAs [44], [49].

IVA

In IVA, one explicitly assumes a separate source and mixing matrix for each data set and, for K data sets, writes $\mathbf{x}^{[k]}(\nu) = \mathbf{A}^{[k]} \mathbf{s}^{[k]}(\nu)$, $\mathbf{x}^{[k]}(\nu), \mathbf{s}^{[k]}(\nu) \in \mathcal{R}^N$ $k = 1, 2, \dots, K$. The independent decomposition of all K data sets is then achieved jointly by fully taking advantage of the statistical second- and higher-order correlations that exist among the data sets.

The key definition in the formulation of IVA is the source component vector (SCV) that is formed by using the corresponding elements of the source random vectors $\mathbf{s}^{[k]}(\nu)$, such that the n th SCV is given by $[s_n^{[1]} s_n^{[2]}, \dots, s_n^{[K]}]^T$, where the subscript refers to the n th source in each of K data sets. For example, for data from K subjects, this would be the n th spatial map for each subject [25], [50]. The IVA decomposition is achieved by minimizing the

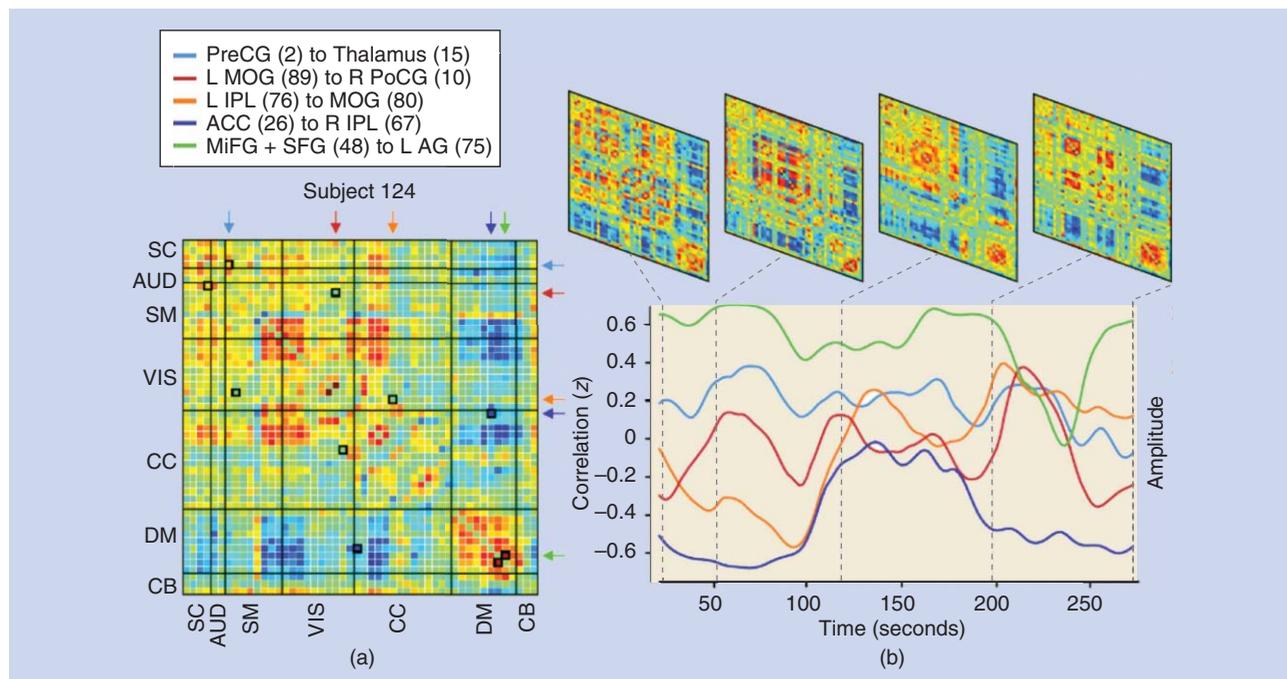


FIGURE 3. FNC dynamics via windowing, single-example subject: (a) average FNC (cross-correlation of ICA timecourses) for a single subject and (b) FNC time series between select components and snapshots of whole-brain FC30. The numbers in parentheses correspond to the component numbers with time courses that were cross-correlated to compute the functional network connectivity matrices.

mutual information among the SCVs (as opposed to sources in ICA), which is equivalent to finding sources that are independent within each data set while maximizing the mutual information within each one of N SCVs [25]. The use of this statistical dependence allows the mitigation of permutation ambiguity for sources (modes) that are dependent across the data sets so that the source estimates across subjects are aligned.

Capturing time-varying connectivity

The capture of time-varying coupling between variables is a topic that has been heavily studied in other fields and in communications for signal processing in particular. However, the specific application to whole-brain functional connectivity is relatively new [22], and its application to brain-imaging data poses particular challenges that are currently being studied. One important challenge is how to best identify relevant features from the high-dimensional brain-imaging data. Both group ICA and IVA can be effectively used to extract features of interest from the fMRI data that in a second step can be used to characterize the dynamic properties. In this section, we provide a brief introduction of the use of both tools in this context.

Time-varying connectivity captured with group ICA

One approach is to use group ICA of multiple subjects and, after selection of components of interest, capture time-varying changes in the coupling (e.g., covariance) among component timecourses using FNC with a tapered window [30]. The FNC information shown in Figure 2 was computed by

assuming that the connectivity was static throughout the experiment. Dynamic approaches capture time-varying connectivity within fMRI data [30], [51] or changes in the spatial maps (spatial FNC).

The simplest approach is to use a windowing method [24], [30], [52]. An example of this is shown in Figure 3, in which group ICA was run on multiple subjects, followed by selection of components of interest and then cross-correlation of the ICA timecourses, called *dynamic FNC (dFNC)*. Figure 3(a) shows a cross-correlation matrix for the entire ICA timecourse for a single subject. A tapered Gaussian window was used to compute time-varying correlation matrices [top of Figure 3(b), with individual correlations shown at the bottom for the black boxes marked in the matrix in Figure 3(a)]. There is considerable variability in the connectivity, which does not appear to be noise due to the modularity of the correlation matrices and the fact that the timecourse tends to be low frequency.

Dynamics of spatial patterns captured with IVA

We can also process overlapping windows jointly using IVA to capture time-varying spatial patterns, as demonstrated in [42]. Because IVA jointly optimizes independence, the use of shorter time windows becomes possible, allowing for sufficient statistical power for the estimation. One example for the use of IVA to capture changes in the spatial coupling (changes in either the within-component maps or the coupling among spatial networks) is in conjunction with a group-level PCA [48]. The data are partitioned into K time windows of equal size T , and then the window from each of the M subjects is

analyzed groupwise, as shown in Figure 4. The dimensionality of each data set X' is reduced from MT to N , resulting in dimension-reduced data sets. This approach enables us to capture changes in the spatial patterns that reflect connectivity over time. In the section “Changes in Time-Varying Spatial Patterns in Patients with SZ,” we show a summary of the results from an application of this approach to evaluate group differences in spatial dynamics.

Characterization of time-varying connectivity

Once the relevant features are extracted from the data, they must be analyzed to evaluate their dynamic properties. Three important approaches to accomplish this are Markov modeling, metastate analysis based on windowing or adaptive approaches (e.g., in which pairwise correlations are computed using small portions of the data), and time-frequency analysis (in which a time-frequency approach is used to transform the data and study patterns of amplitude, phase, and frequency over time [21], [31]).

Markov modeling/state transitions

Markov chain (MC) modeling provides a powerful way to characterize (and distinguish) time-varying connectivity [30], [42]. A data-driven approach can be used to learn both the states and transitions from the data (in both space and time). Figure 5(a) shows the state assignments as a function of time for three representative subjects for the dFNC approach. Transition behavior can be characterized by considering an MC in which the probability to go from the current state to the next state is conditionally independent from all states that occurred (in time) before the current state.

In Figure 5(b), we show the average transition matrix (TM) for our example. Red squares along the diagonal signify a very high probability of staying in the same state. For the off-diagonal elements, hotter colors in column 1 indicate a higher probability of entering state 1 from the other states, and cooler colors in row 3 indicate a lower probability of exiting state 3. Because the MC is irreducible (any state can be reached from any other state in a finite number of steps), its stationary distribution (π) can be obtained as the principal eigenvector of the estimated TM [53].

The vector π , displayed in Figure 5(c), represents the probability distribution over the states of the MC when the chain is in its stationary regime, that is, in the expected behavior of the system in the long run. In our example, the stationary probability for state 3 is far greater than the probabilities for other states, meaning that, in the long run, the system is most likely to be found in state 3. MCs enable us to capture the propagation of probability distribution vectors over the states (i.e., mixed-state vectors) through a network.

Cross-state summary measures (e.g., metastates)

A core challenge for dynamic network connectivity analysis is to summarize the data in ways that simultaneously reduce the data’s dimensionality and expose features that are strongly predictive of important population characteristics. The native dimension of network correlation space can easily exceed 1,000. However, recent approaches have been developed to summarize the dynamic information in a higher-level summary [75]. In this case, the goal is to calculate a tractable characterization of time-varying connectivity

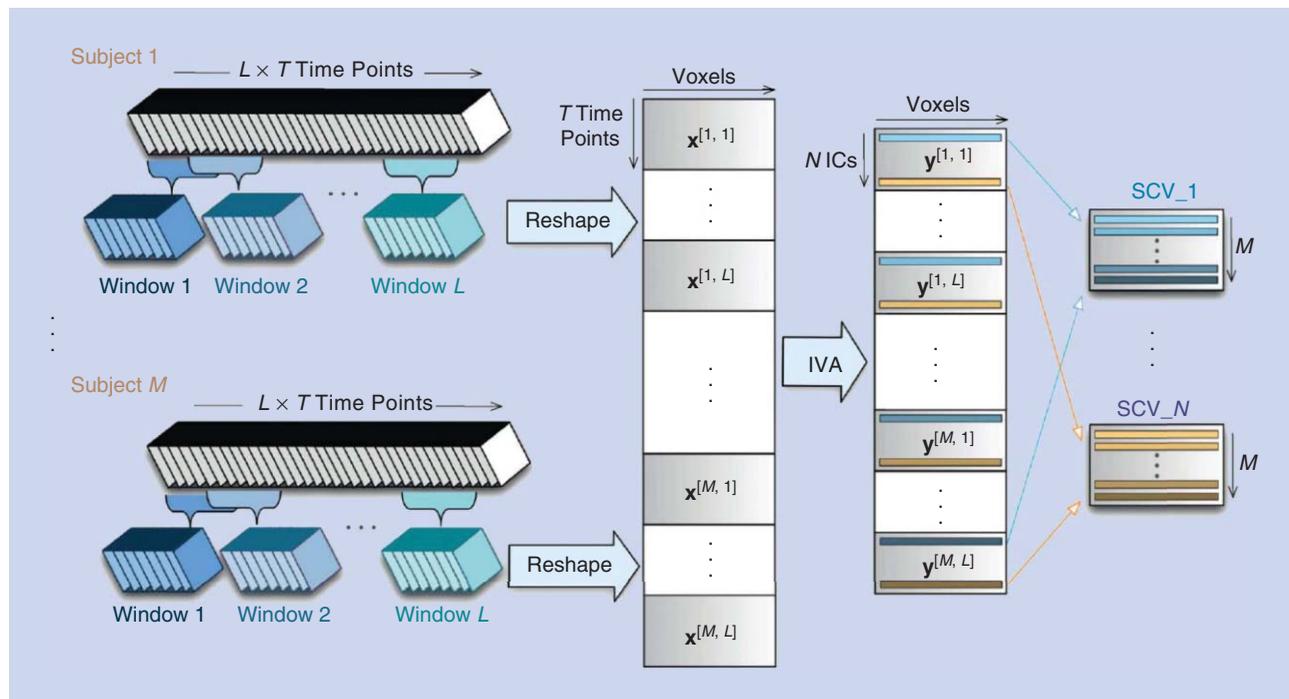


FIGURE 4. The IVA approach to characterize spatially dynamic and static components [4], [18], [42]. Here, spatial maps of a component vector are related over the time windows but should be distinct from the spatial maps of all other components (whether within or outside the current window w_t).

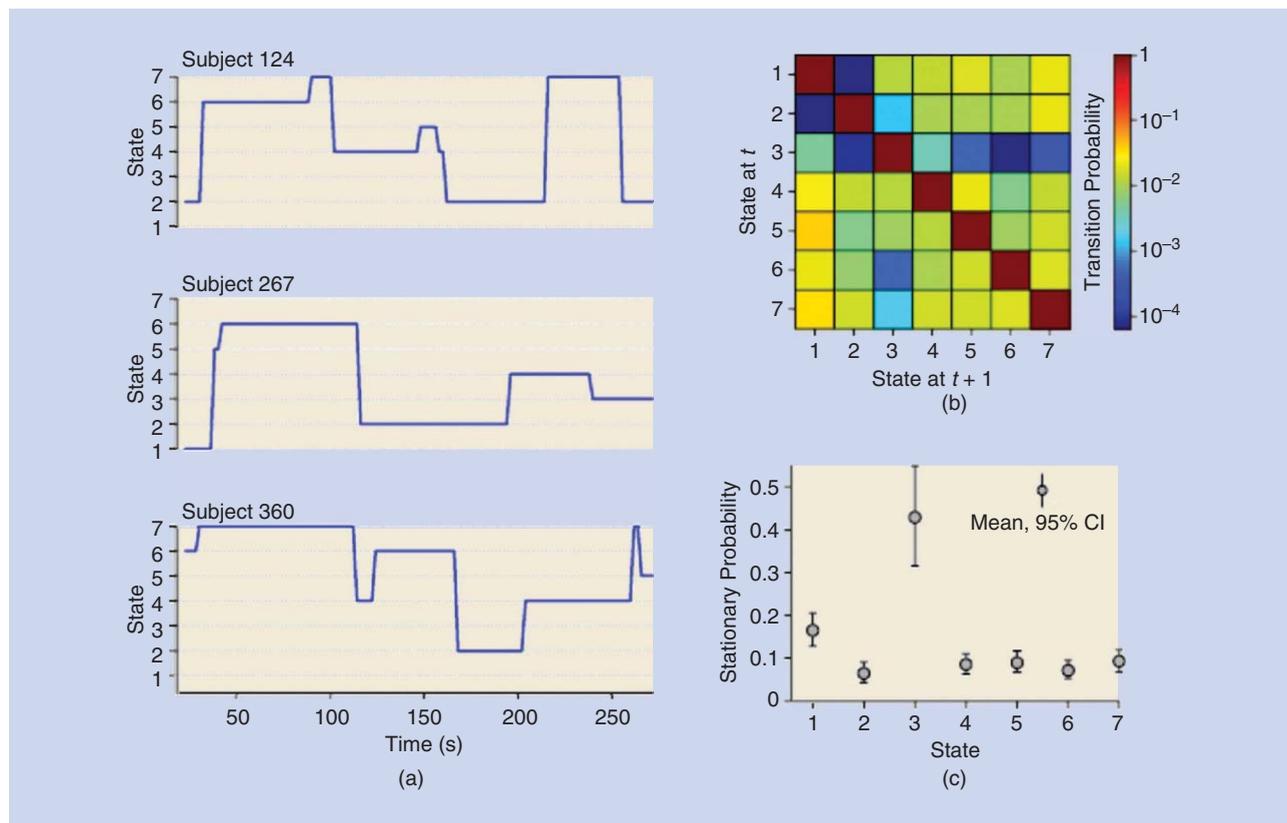


FIGURE 5. (a) State transition vectors for the three example subjects. Assigned states are plotted at the time point corresponding to the center of the sliding window. (b) The state TM averaged over subjects. High values along the diagonal indicate a high probability of staying in a state. Note that transition probability is color mapped on a log-scale. (c) The stationary probability vector (π , principal eigenvector of the TM) shows the steady-state or “long-run” behavior. Error bars indicate the nonparametric 95% confidence intervals (CIs) obtained from 1,000 bootstrap resamples of the average TM (resampling subjects).

in terms of the additive contributions of a set of basis correlation patterns (BCPs) obtained according to some specified optimization criterion (using, for example, temporal ICA, spatial ICA, PCA, or k -means clustering) [34], [54]. A BCP in the context of a PCA-based approach would be called an *eigenconnectivity*.

This is summarized in the “Estimation of Dynamic States” column of Figure 1(b). The time-indexed N -element vectors of BCP weights, discretized according to signed quartile, are the metastates. In a recent work [34] (Figure 6), the results showed a summary of a three-level five-state quantization in 400 healthy subjects that indicates the following.

- 1) Only 22 of these metastates are occupied more than 1% of the time.
- 2) These states include mostly single- or double-state occupancy.
- 3) Females show more single-state occupancy than males, who show more double-state occupancy [34].

Using a large, balanced, multisite data set, we also investigated the effect of SZ diagnosis on four interrelated measures of metastate dynamism, separately evaluated with respect to BCPs obtained from four common algorithms [55]. These analyses yielded consistent and significant evidence for reduced connectivity dynamism in patients with

SZ and provided strong evidence in support of such summary measures.

There are a number of possible ways to compute cross-state summary measures, a topic of ongoing work. One example of such a metric is the concept of a k -level hub (e.g., states that are returned to k or more times). Related concepts include absorbing hubs (subjects stay for extended periods of time) and transient hubs (subjects come in and out multiple times for short periods), both of which appear highly different in SZ [55], [75].

Time-frequency analysis

Chang and Glover [21] first introduced the use of time-frequency methods to study time-varying connectivity (coherence) in a few regions of interest. More recently, a whole-brain time-frequency approach was proposed that enables brain states to be estimated [56]. The proposed approach can be considered an extension and generalization of both the time-domain [30] and coherence [21] approaches (see [56] for more details). Using this approach, we can more fully characterize a state via multiple frequency bands by its connectivity pattern (covariance), frequency contribution, and phase (e.g., anticorrelated pairs would have a 180° phase shift).

Figure 7 shows an example of a state that was defined via k -means clustering after the use of a complex Morlet filter

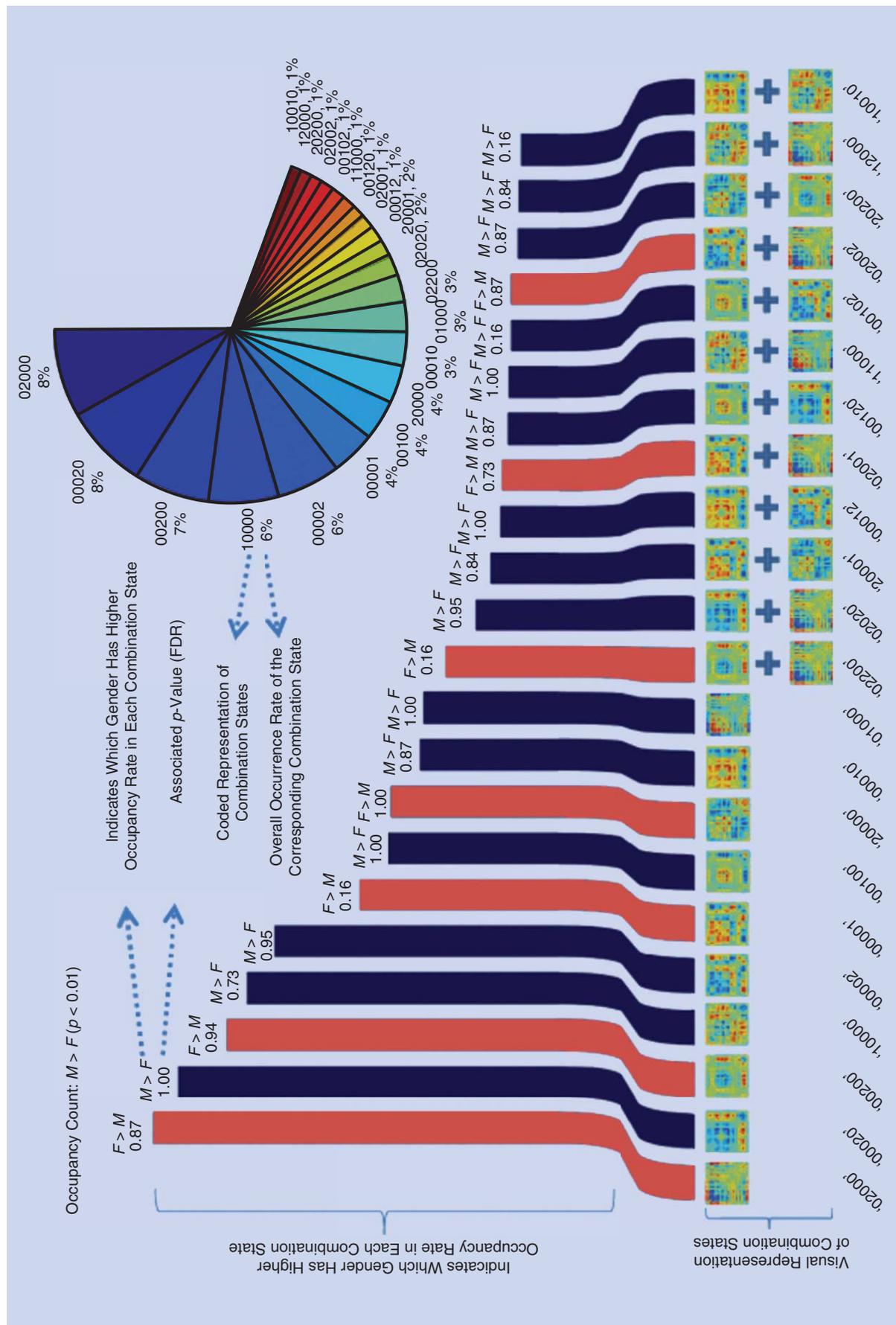


FIGURE 6. The bar chart shows relative occurrence frequency of each combination state. The bottom label of each bar is the coded and visual representation of the associated combination state, and the top is labeled by the gender that, on average, occupies that combination state more, along with the false discovery rate (FDR)-adjusted p -values for that comparison (dark blue: males (M) > females (F); dark red: F > M). Each slice of the pie chart shows the overall occurrence frequency of the associated combination state. (Figure reprinted with permission from [34].)

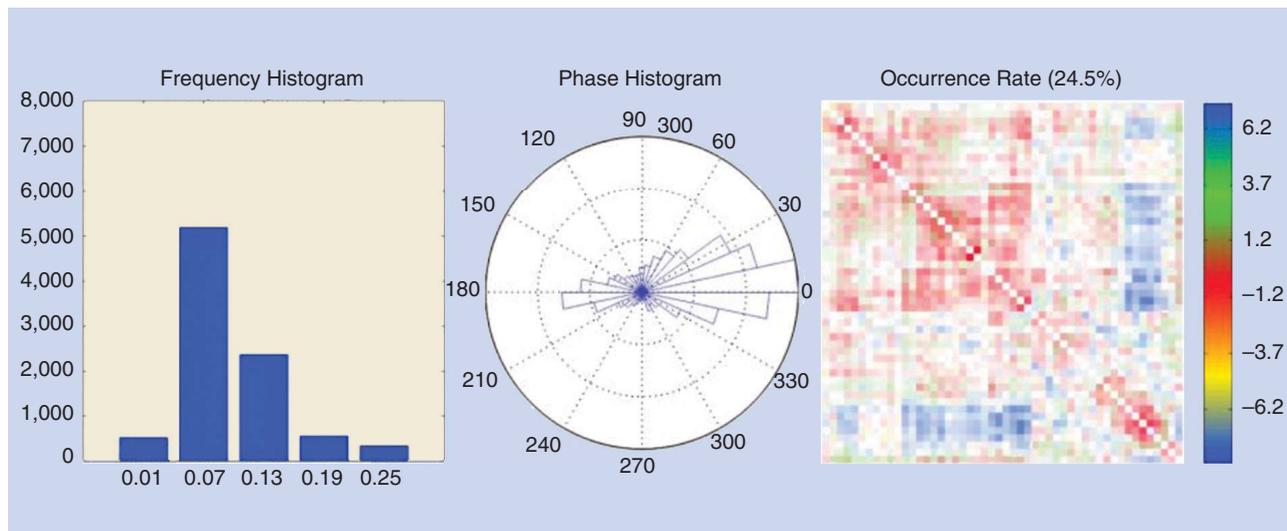


FIGURE 7. The multiband state with 25% occurrence rate showing the most power in the 0.07 and 0.13 frequency bands. Phase histogram and color indicate the phase of the dynamics.

to separate five different frequency bands with magnitude and phase. This particular state has most of its power within 0.07–0.13 Hz, has some strong 0° and 180° phase patterns, and captures some very interesting patterns. The results from a large resting-state fMRI data set ($N = 400$) identified two states, with similar correlation patterns but distinct frequency profiles, one of which was highly predictive of males versus females [31]. This provides additional evidence that ignoring the dynamic information obscures important knowledge.

Validation

Quite a few studies have been published that provide important information validating the presence of chronnectomic information in fMRI data. For example, one study of a large subject ($N = 400$) data set performed a split-half replication and also varied a number of parameters, including the number of estimated states and the window size [30], [57]. Other studies have shown that dynamic connectivity tracks closely with sleep state [28] and psychedelic experience [58], is reflected in both humans and macaques [51], and is associated with daydreaming [59]. Cross-validated classification also appears to be more powerful when applied to dynamic connectivity measures [4], [60]. Comparison of dynamic connectivity measures in the presence of tasks that activate known brain regions also provides powerful evidence to support the presence of connectivity states [18], [61].

Concurrent EEG/fMRI experiments

Concurrent EEG provides a useful way to validate these dynamic changes by providing convergent evidence for them. Although EEG alone cannot provide a ground-truth measure because EEG and fMRI are generated by and sensitive to very different sources, we do expect that fMRI changes in connectivity over time that reflect neuronal changes will also be detectable with EEG.

An illustrative example examining differences in dynamics associated with the eyes-open (EO) versus eyes-closed (EC) state is presented in Figure 8. Concurrent EEG/fMRI data were collected using a Brain Products EEG system (Gilching, Germany) that had been previously used to collect data comparing a variety of frequencies in EEG with fMRI data in the resting state for EO and EC [62], [63]. The results from a preliminary analysis of these data using a group ICA approach to evaluate temporal dynamics are shown in Figure 8 [64]. Figure 8(a)–(f) shows two dynamic states estimated from the fMRI data. Both of these states showed a significant difference between EO versus EC, with state 1 occurring dominantly for EO and state 5 showing significantly more occurrence during the EC stage and demonstrating more EEG alpha power. The anticorrelation with brain regions associated with inner reflection (regions in the widely studied default-mode network [65]) was also stronger in the EO data. As the states were associated with more EEG drowsiness measures, these anticorrelations diminished and then subsided [64].

This is only a relatively simple approach to relating EEG and fMRI data. More advanced methods that take advantage of the joint information during the estimation process would likely be even more fruitful in demonstrating the benefits of dynamic connectivity [63], [66]–[69].

Incorporating dynamics improves contrast to noise

The data shown in Figure 9 were evaluated from a normative resting-state fMRI data set ($N = 200$ HCs) using a simple model that incorporates an explicit static subspace (while modeling the dynamic information in a nuisance subspace). In this case, the model that incorporates the dynamic information [Figure 9(a)] shows a higher contrast-to-noise ratio than when the dynamic information is completely ignored [Figure 9(b)]. This result also provides strong support

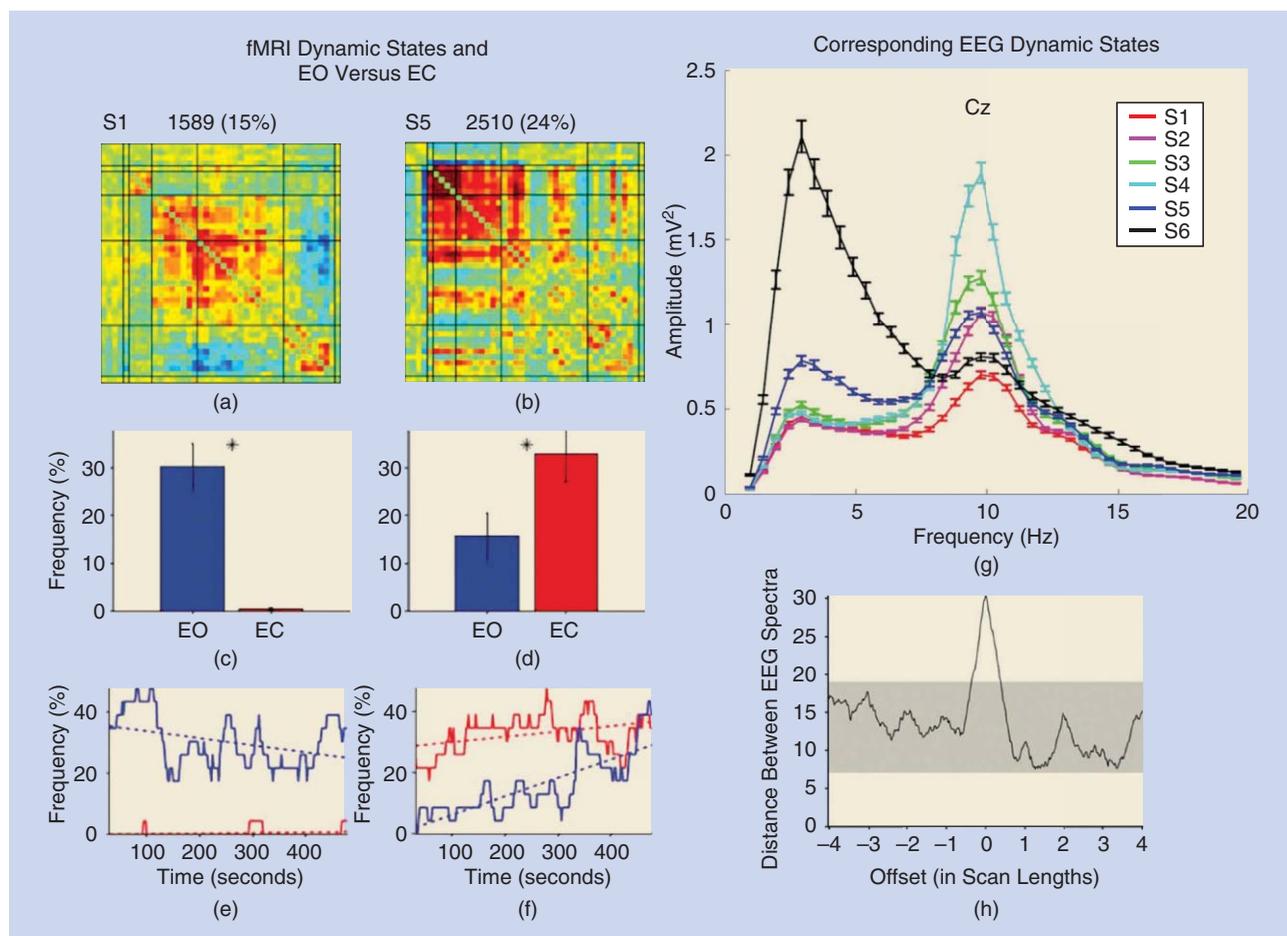


FIGURE 8. Concurrent EEG/fMRI temporal dynamics during EO versus EC [64]. fMRI analysis is shown in (a)–(f), and EEG data analyzed within identified fMRI states for one electrode are shown in (g). EEG data reflected considerably more theta/delta power during the states occurring more in the EC condition. EEG was strongly correlated with the identified fMRI states, as shown in (h) that shows the distance among EEG and fMRI, which reduced as EEG windows were shifted in time away from the fMRI states. The asterisks in (c) and (d) indicate significance at $p < 0.05$ (corrected for multiple comparisons). Cz in (g) indicates the EEG central electrode in the 10–20-coordinate system.

for the use of models that capture both the static and dynamic connectivity information.

Choice of estimation strategy and parameters

One common critique of windowed correlation approaches is that they can introduce spurious correlations [24], [70]. A number of papers have quite carefully evaluated various window parameters and performance in simulations in real data [22], [24], [30]. In particular, spurious changes in connectivity appear if the sliding window length is shorter than the largest period present in the signals [24], suggesting window lengths of at least 30 seconds for fMRI. The combination of multimodal data (e.g., EEG and fMRI) might help mitigate the issue and confirm that the changes are real [71]. Instead of a fixed window, adaptive windowing approaches can also be used [23]. More importantly, fixed-window approaches perform quite similarly in their mean to adaptive windowing approaches [23]. In addition, the combination of multivariate approaches with windowing appears to be more robust in preventing spurious correlations than are univariate approaches [4]. Another choice involved is the number of states. This

has not yet been evaluated comprehensively, although in multiple papers, an evaluation of results with various numbers of states has been presented to ensure that results are not heavily dependent on the final choice (see, for example, [30]).

Applications

There have already been numerous uses of time-varying connectivity in fMRI data. In this section, we review three interesting applications:

- 1) a study of changes in spatial connectivity patterns in SZ
- 2) an evaluation of the relationship between the sleep stage and connectivity
- 3) an evaluation of the differences in connectivity states in individuals who are either heavy smokers or heavy drinkers.

Changes in time-varying spatial patterns in patients with SZ

It is challenging to consider changes over time in both spatial and temporal aspects of connectivity, but spatial patterns are also an important aspect of the dynamic information. In an analysis of patients with SZ versus HCs, we used the

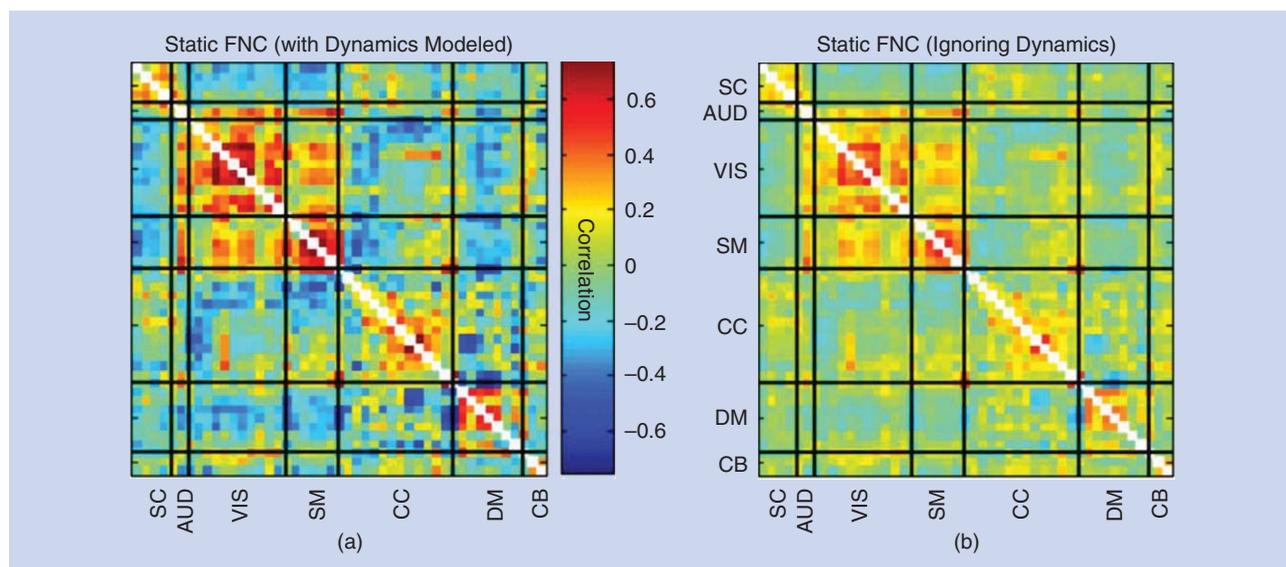


FIGURE 9. The data show that the static FNC pattern estimated by (a) a model incorporating dynamics has higher contrast to noise than (b) one ignoring dynamics.

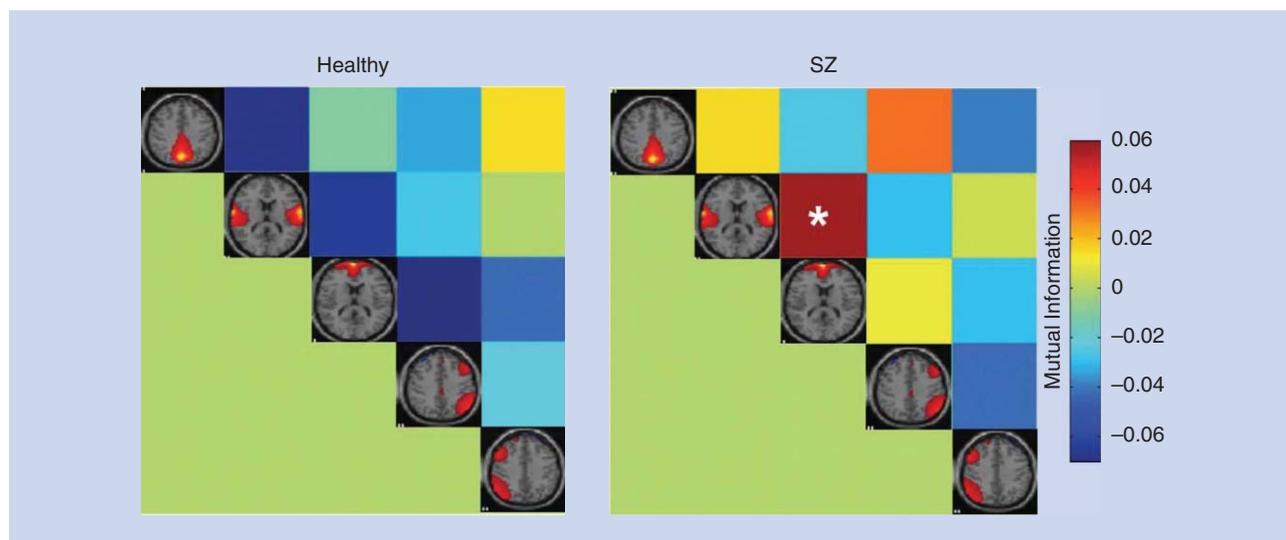


FIGURE 10. SZ patients exhibit significant changes in the spatial dependency between default-mode and temporal-lobe networks. (Figure reprinted from [76] with permission.)

windowed IVA approach shown in Figure 4 with seven windows, each of which overlapped by 50% to cover a 200-time point resting-state fMRI data set. Thirty components were estimated, and 12 of them were determined to be related to brain function and not artifact. Computation of MC transition probabilities between multiple states demonstrated that the controls showed significantly less probability to transition between states. This provides a way to summarize changes in the spatial patterns over time. We can also evaluate changes in the dependencies between pairs of spatial networks over time. To estimate spatial dependencies,

we computed a mutual information matrix for each subject and each window. The spatiotemporal dependency dynamics are very interesting. Some showed significant differences between patients with SZ and HCs. For example, patients with SZ showed more coupling between the network dynamics of the medial prefrontal cortex (brain regions thought to mediate cognition) and the temporal lobe (regions that process sound and language and are known to be disrupted in SZ) than did controls (Figure 10). This is a simple summary measure of only seven windows, but it indicates that spatial dynamics are a sensitive measure of disease state.

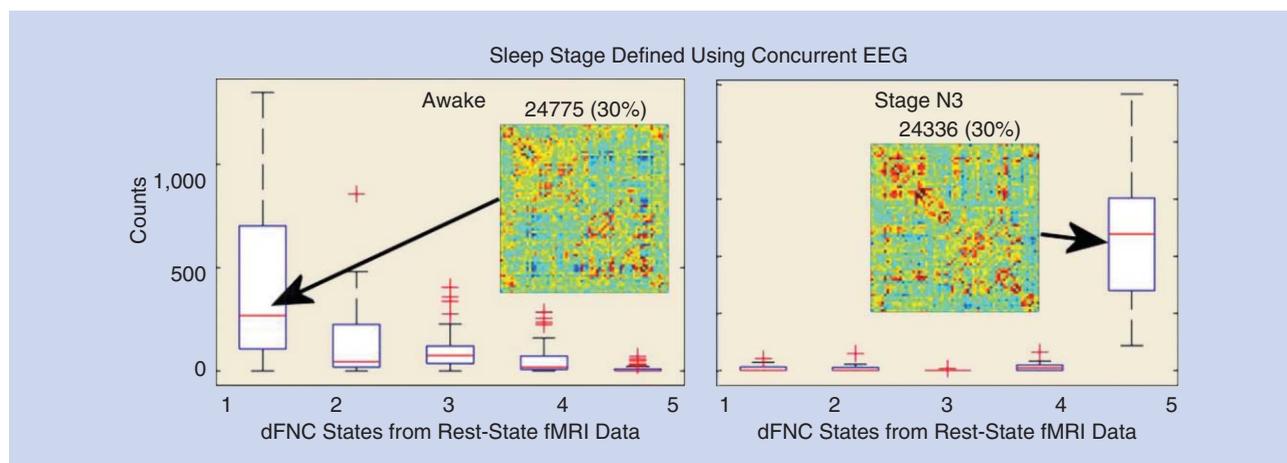


FIGURE 11. Dynamic states 1 and 5 map to awake and deeper-sleep stages. (Figure reprinted from [73] with permission.)

Time-varying connectivity in fMRI maps to EEG-defined sleep stages

For evidence of the utility of the dynamic patterns, we evaluated resting-state fMRI data collected from 55 subjects for 50 minutes each (1,500 volumes, TR = 2.08 seconds) with a Siemens 3T Trio scanner (Malvern, Pennsylvania) while the subjects transitioned from wakefulness to, at most, sleep stage N3 (for more details, see [72]). Simultaneous EEG was acquired, facilitating sleep staging according to the American Academy of Sleep Medicine (AASM) criteria, resulting in one hypnogram per subject (a vector assignment of consecutive 30-second EEG epochs to one of awake, N1, N2, and N3 sleep). Following our recent work [30], we estimated dFNC between components following a group ICA. We then computed the counts of these dFNC windows for each hypnogram state. The results show that states 1 and 5 mapped strongly to the awake and deeper-sleep stages, respectively (Figure 11) [73]. More work is needed, but the results strongly support the utility of capturing dynamic connectivity.

Time-varying connectivity is significantly changed in substance users

A greater understanding of individual differences in the neurobiology of substance use is integral to developing more effective interventions. A large body of evidence shows aberrant brain structure and function in substance users. Whereas some specific regions are implicated in cravings and loss of controls (e.g., mesocorticolimbic regions), for the most part, these studies are heterogeneous and do not provide the ability to discriminate between substance users and controls at the level of the individual. In part, we believe this is because the connectivity methods focused on static measures and did not fully capture the variability of the patterns within the patient groups. In particular, it is clear that 1) large heterogeneity in the substance-use disorder brain function makes analysis challenging; and 2) although certain brain pathways have been hypothesized as most affected, all of these disorders encompass multiple

interacting brain regions. The ability to evaluate the dependencies between multiple functional brain networks is critical to understanding the disorders.

The dFNC results were computed for smokers and drinkers ($N = 50$) and identified significant changes in correlation among multiple brain networks. Figure 12 shows dFNC matrices for two dynamic states demonstrating differences among smokers and drinkers. State 1 lacks most of the anticorrelation between default-mode and other networks (pink boxes), as well as the connectivity within sensorimotor regions. Some interesting differences are also apparent when evaluating the dwell time each group spent in the dynamic states. For example, smokers and drinkers both spent significantly more of their time within state 2. The percentage of time each group spent in these two states is illustrated in Figure 13 and was significantly different among controls, smokers, and drinkers. Neither of these interesting results is observable from the static results. The importance of such a result is that the ability of methods that focus on dynamics to separate out information about the neurobiology of substance use may teach us more about how the brain is different in nicotine or alcohol use. In addition, this information may provide a more accurate biomarker that can be used to predict, for example, treatment outcomes.

Conclusions

In summary, time-varying connectivity is a powerful tool for improving our understanding of the brain. There are still plenty of avenues of ongoing investigation that require creative thinking and the development of advanced signal processing methods to improve the estimation performance and the extraction and characterization of meaningful information. For example, some specific directions of interest include the development of approaches that can capture both static and dynamic connectivity patterns. Moreover, approaches that can capture spatiotemporal patterns of connectivity would be very desirable because it is clear that both static and dynamic connectivity patterns are changing in systematic and interesting ways. In

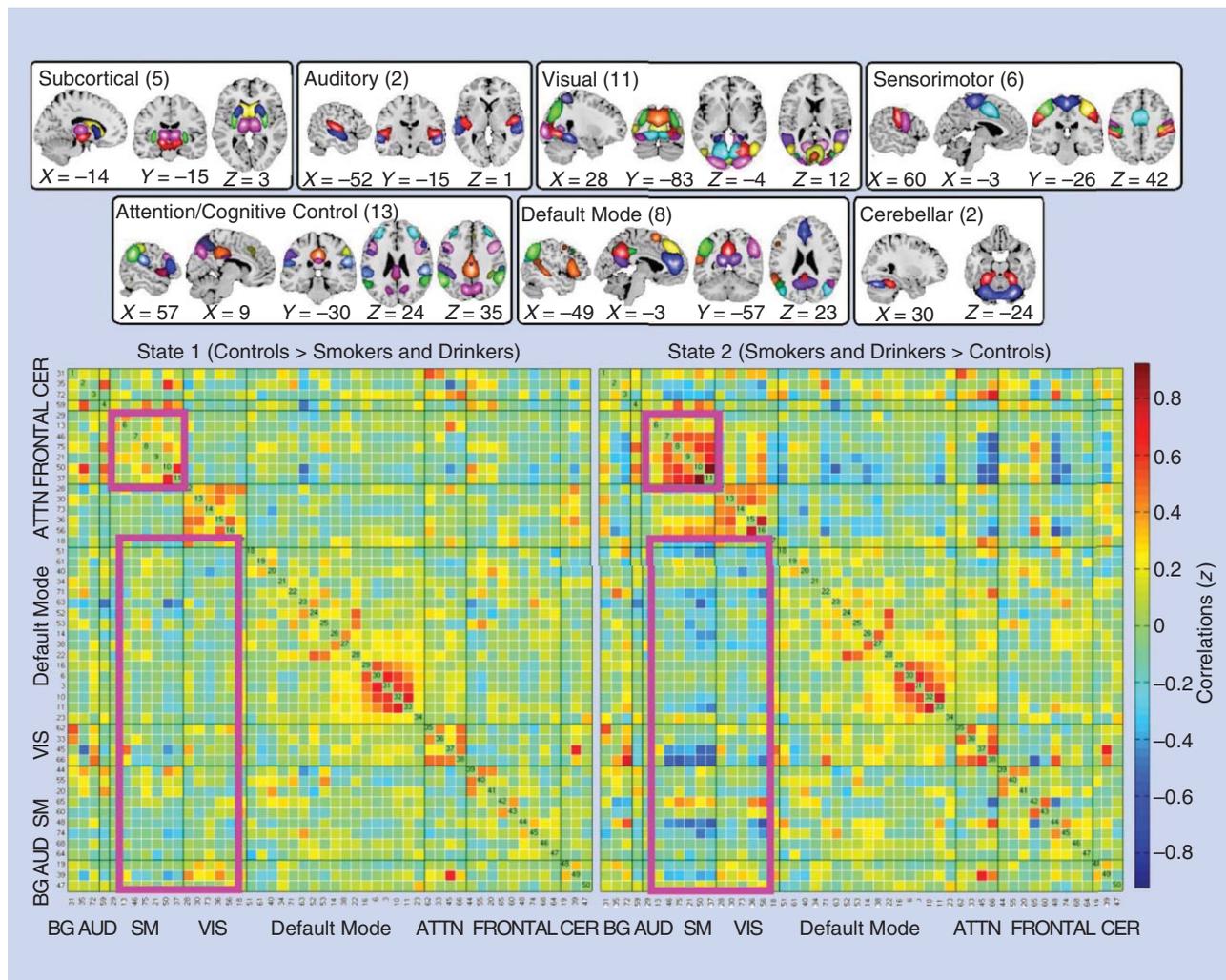


FIGURE 12. The dFNC states that show significant group differences. In particular, smokers and drinkers spent more time in state 1 versus state 2 (of five estimated states). Notably, state 1 lacks the predominant anticorrelation between default-mode regions that is visible in state 2, suggesting that their lack may serve either as a protective factor or as a marker of substance use. The numbers in parentheses indicate the number of components that were included in each group. The X, Y, and Z coordinates of the slice in millimeters in the Montreal Neurological Institute coordinate system are shown.

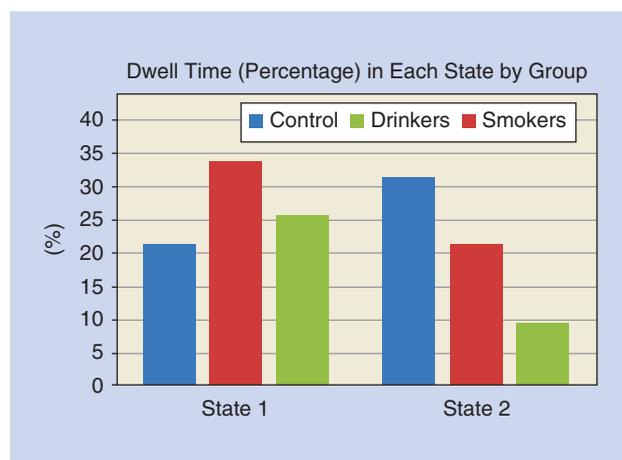


FIGURE 13. The dwell time (percentage) for states 1 and 2. Smokers and drinkers spent significantly more time ($p < 0.0001$) in state 1, whereas controls spent more time in state 2.

addition, more studies that map task information to the states will improve our understanding of the function of these connectivity states [74].

Finally, there is an important need for continued work in characterizing single and multiple states or other summary measures that provide intuitive ways of conveying brain connectivity in a way that respects the dynamic nature of the brain. Such approaches should inform us about the healthy brain and direct us to important aspects of disease, especially for complex mental illnesses, such as SZ and autism spectrum disorder.

Acknowledgments

We thank Victor Vergara, Robyn Miller, Maziar Yaesoubi, and Eswar Damaraju for their input and help with data processing. The work was supported in part by the National Institutes of Health under the Centers of Biomedical Research Excellence

(COBRE) grant P20GM103472 and grants R01EB005846, IR01EB006841, and R01EB020407 and National Science Foundation EPSCoR grant 1539067.

Authors

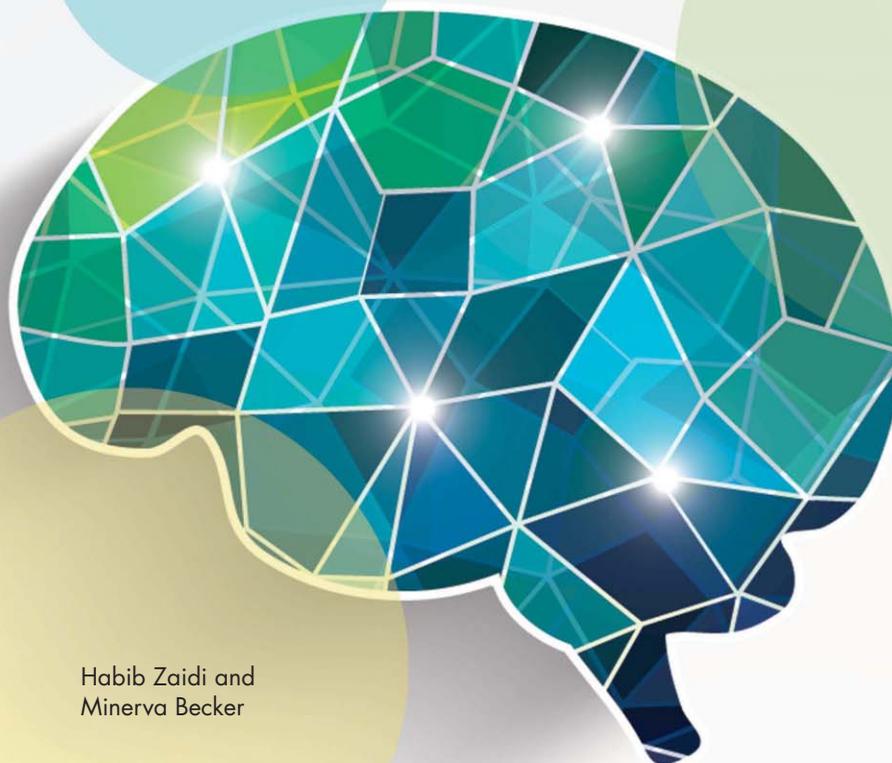
Vince D. Calhoun (vcalhoun@unm.edu) received his Ph.D. degree in electrical engineering from the University of Maryland, Baltimore County, in 2002. He is currently the executive science officer at the Mind Research Network and a distinguished professor in the Department of Electrical and Computer Engineering at the University of New Mexico. He is the author of more than 400 full journal articles and 500 technical reports, abstracts, and conference proceedings. His work includes independent component analysis for functional magnetic resonance imaging, data fusion of multimodal imaging and genetics data, and the identification of biomarkers for disease. He leads a P20 COBRE center grant on multimodal imaging of schizophrenia, bipolar disorder, and major depression and a National Science Foundation Experimental Program to Stimulate Competitive Research grant focused on brain imaging and the epigenetics of adolescent development. He is a Fellow of the IEEE, the Association for the Advancement of Science, the American Institute for Medical and Biological Engineering, and the International Society for Magnetic Resonance in Medicine. He is currently the chair of the IEEE Machine Learning for Signal Processing Technical Committee.

Tülay Adalı (adali@umbc.edu) received her Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, in 1992 and joined the faculty at the University of Maryland, Baltimore County (UMBC), the same year. She is currently a distinguished university professor in the Department of Computer Science and Electrical Engineering at UMBC. She was the general or technical cochair of the IEEE Workshops on Machine Learning for Signal Processing and Neural Networks for Signal Processing 2001–2008 and helped organize a number of conferences, including the IEEE International Conference on Acoustics, Speech, and Signal Processing. She is a Fellow of the IEEE and American Institute for Medical and Biological Engineering; a Fulbright Scholar; and recipient of a 2010 IEEE Signal Processing Magazine Best Paper Award, a 2013 University System of Maryland Regents' Award for Research, and a National Science Foundation CAREER Award. She was an IEEE Signal Processing Society Distinguished Lecturer for 2012 and 2013. Her research interests include statistical signal processing, machine learning for signal processing, and biomedical data analysis.

References

- [1] N. B. Turk-Browne, "Functional interactions as big data in the human brain," *Science*, vol. 342, no. 6158, pp. 580–584, Nov. 2013.
- [2] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. Behrens, R. Bucholz, A. Chang, L. Chen et al., "The Human Connectome Project: A data acquisition perspective," *NeuroImage*, vol. 62, no. 4, pp. 2222–2231, Oct. 2012.
- [3] P. S. Weiss, "President Obama announces the BRAIN Initiative," *ACS Nano*, vol. 7, pp. 2873–2874, Apr. 2013.
- [4] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı, "The chronnectome: Time-varying connectivity networks as the next frontier in fMRI data discovery," *Neuron*, vol. 84, no. 2, pp. 262–274, Oct. 2014.
- [5] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, and K. Ugurbil, "The WU-Minn Human Connectome Project: An overview," *NeuroImage*, vol. 80, pp. 62–79, Oct. 2013.
- [6] R. C. Craddock, S. Jbabdi, C. G. Yan, J. T. Vogelstein, F. X. Castellanos, A. Di Martino, C. Kelly, K. Heberlein et al., "Imaging human connectomes at the macroscale," *Nat. Methods*, vol. 10, no. 6, pp. 524–539, June 2013.
- [7] T. Insel. (2014, Dec. 16). *Director's Blog: Best of 2014*. [Online]. Available: <http://www.nimh.nih.gov/about/director/2014/best-of-2014.shtml>
- [8] P. A. Sharp and R. Langer, "Research agenda. Promoting convergence in biomedical science," *Science*, vol. 333, no. 6042, p. 527, July 2011.
- [9] R. F. Betzel, L. Byrge, Y. He, J. Goni, X. N. Zuo, and O. Sporns, "Changes in structural and functional connectivity among resting-state networks across the human lifespan," *NeuroImage*, vol. 102, pt. 2, pp. 345–357, Nov. 2014.
- [10] M. J. Brookes, J. R. Hale, J. M. Zumer, C. M. Stevenson, S. T. Francis, G. R. Barnes, J. P. Owen, P. G. Morris et al., "Measuring functional connectivity using MEG: Methodology and comparison with fcMRI," *NeuroImage*, vol. 56, no. 3, pp. 1082–1104, June 2011.
- [11] M. J. Brookes, M. Woolrich, H. Luckhoo, D. Price, J. R. Hale, M. C. Stephenson, G. R. Barnes, S. M. Smith, et al., "Investigating the electrophysiological basis of resting state networks using magnetoencephalography," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 108, no. 40, pp. 16783–16788, Oct. 2011.
- [12] J. M. Houck, M. Cetin, A. Mayer, J. Bustillo, M. Brookes, and V. D. Calhoun, "Comparison of resting intrinsic connectivity networks in schizophrenia and healthy controls computed via group spatial ICA of MEG and fMRI data," in *Proc. Organ. Human Brain Mapping*, Hamburg, Germany, June 8–12, 2014.
- [13] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Segmentation of brain electrical activity into microstates: Model estimation and validation," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 7, pp. 658–665, July 1995.
- [14] T. Koehnig, L. Prichep, D. Lehmann, P. V. Sosa, E. Braeker, H. Kleinlogel, R. Isenhardt, and E. R. John, "Millisecond by millisecond, year by year: Normative EEG microstates and developmental stages," *NeuroImage*, vol. 16, no. 1, pp. 41–48, May 2002.
- [15] J. Kindler, D. Hubl, W. K. Strik, T. Dierks, and T. Koehnig, "Resting-state EEG in schizophrenia: Auditory verbal hallucinations are related to shortening of specific microstates," *Clin. Neurophysiol.*, vol. 122, no. 6, pp. 1179–1182, June 2011.
- [16] D. Lehmann, "Topography of spontaneous alpha EEG fields in humans," *Electroencephalogr. Clin. Neurophysiol.*, vol. 30, no. 2, pp. 161–162, Feb. 1971.
- [17] D. Lehmann, H. Ozaki, and I. Pal, "EEG alpha map series: Brain micro-states by space-oriented adaptive segmentation," *Electroencephalogr. Clin. Neurophysiol.*, vol. 67, no. 3, pp. 271–288, Sept. 1987.
- [18] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. Bandettini, V. D. Calhoun, M. Corbetta, S. D. Penna, J. H. Duyn et al., "Dynamic functional connectivity: Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360–378, Oct. 2013.
- [19] Z. Shehzad, A. M. Kelly, P. T. Reiss, D. G. Gee, K. Gotimer, L. Q. Uddin, S. H. Lee, D. S. Margulies, et al., "The resting brain: Unconstrained yet reliable," *Cereb. Cortex*, vol. 19, no. 10, 2209–2229, Feb. 2009.
- [20] I. Cribben, R. Haraldsdottir, L. Y. Atlas, T. D. Wager, and M. A. Lindquist, "Dynamic connectivity regression: Determining state-related changes in brain connectivity," *NeuroImage*, vol. 61, no. 4, pp. 907–920, July 2012.
- [21] C. Chang and G. H. Glover, "Time–frequency dynamics of resting-state brain connectivity measured with fMRI," *NeuroImage*, vol. 50, no. 1, pp. 81–98, Mar. 2010.
- [22] U. Sakoglu, G. D. Pearlson, K. A. Kiehl, Y. M. Wang, A. M. Michael, and V. D. Calhoun, "A method for evaluating dynamic functional network connectivity and task-modulation: Application to schizophrenia," *Magn. Reson. Mater. Phys.*, vol. 23, pp. 351–366, Dec. 2010.
- [23] M. A. Lindquist, Y. Xu, M. B. Nebel, and B. S. Caffo, "Evaluating dynamic bivariate correlations in resting-state fMRI: A comparison study and a new approach," *NeuroImage*, vol. 101, pp. 531–546, Nov. 2014.
- [24] N. Leonardi and D. Van De Ville, "On spurious and real fluctuations of dynamic functional connectivity during rest," *NeuroImage*, vol. 104, pp. 430–436, Jan. 2015.
- [25] T. Adalı, M. Anderson, and G. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Processing Mag.*, vol. 31, no. 3, pp. 18–33, May 2014.
- [26] N. Leonardi, W. R. Shirer, M. D. Greicius, and D. Van De Ville, "Disentangling dynamic networks: Separated and joint expressions of functional connectivity patterns in time," *Hum. Brain Mapp.*, vol. 35, no. 12, pp. 5984–5995, Dec. 2014.
- [27] N. J. Kopell, H. J. Gritton, M. A. Whittington, and M. A. Kramer, "Beyond the connectome: The dynamo," *Neuron*, vol. 83, no. 6, pp. 1319–1328, Sept. 2014.
- [28] E. Tagliazucchi and H. Laufs, "Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep," *Neuron*, vol. 82, no. 3, pp. 695–708, May 2014.
- [29] E. Damaraju, E. A. Allen, A. Belger, J. M. Ford, S. McEwen, D. H. Mathalon, B. A. Mueller, G. D. Pearlson et al., "Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia," *NeuroImage Clin.*, vol. 5, pp. 298–308, 2014.

- [30] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking whole-brain connectivity dynamics in the resting state," *Cereb. Cortex*, vol. 24, no. 3, pp. 663–676, Nov. 2014.
- [31] M. Yaesoubi, R. L. Miller, and V. D. Calhoun, "Dynamic coherence analysis of rest fMRI with state specific phase, frequency and temporal profile," presented at the *OHBM 2015 Annu. Meeting*, Honolulu, HI, June 14–18, 2015.
- [32] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J. M. Annoni, M. Schluep, P. Vuilleumier, and D. Van De Ville, "Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, pp. 937–950, Dec. 2013.
- [33] R. L. Miller, M. Yaesoubi, and V. D. Calhoun, "Meta-state analysis reveals reduced resting fMRI connectivity dynamism in schizophrenia across multiple multivariate analytic techniques," presented at the *Biennial Conf. on Resting State/Brain Connectivity*, Boston, Sept. 11–13, 2014.
- [34] M. Yaesoubi, R. L. Miller, and V. D. Calhoun, "Mutually temporally independent connectivity patterns: A new framework to study resting state brain dynamics with application to explain group difference based on gender," *NeuroImage*, vol. 107, pp. 85–94, Feb. 2015.
- [35] B. Rashid, E. Damaraju, G. D. Pearson, and V. D. Calhoun, "Dynamic connectivity states estimated from resting fMRI identify differences among schizophrenia, bipolar disorder, and healthy control subjects," *Front. Hum. Neurosci.*, vol. 8, p. 897, Nov. 2014.
- [36] R. C. Craddock, G. A. James, P. E. Holtzheimer, III, X. P. Hu, and H. S. Mayberg, "A whole brain fMRI atlas generated via spatially constrained spectral clustering," *Hum. Brain Mapp.*, vol. 33, no. 8, pp. 1914–1928, Aug. 2012.
- [37] V. D. Calhoun and T. Adali, "Multi-subject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery," *IEEE Rev. Biomed. Eng.*, vol. 5, pp. 60–73, Aug. 2012.
- [38] Q.-H. Lin, J. Liu, Y.-R. Zheng, H. Liang, and V. D. Calhoun, "Semi-blind spatial ICA of fMRI using spatial constraints," *Hum. Brain Mapp.*, vol. 31, no. 7, pp. 1076–1088, Jul. 2010.
- [39] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Trans. Knowledge Data Eng.*, vol. 21, no. 1, pp. 6–20, Jan. 2009.
- [40] A. Y. Mutlu, E. Bernat, and S. Aviyente, "A signal-processing-based approach to time-varying graph analysis for dynamic brain network identification," *Comput. Math. Methods Med.*, vol. 2012, 451516, 2012.
- [41] D. M. Zoltowski, E. M. Bernat, and S. Aviyente, "A graph theoretic approach to dynamic functional connectivity tracking and network state identification," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2014, pp. 6004–6007, 2014.
- [42] S. Ma, V. D. Calhoun, R. Phlypo, and T. Adali, "Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis," *NeuroImage*, vol. 90, pp. 196–206, Apr. 2014.
- [43] P. Comon and C. Jutten, *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. New York: Academic Press Inc., 2010.
- [44] J. V. Stone, J. Porrill, C. Buchel, and K. Friston, "Spatial, temporal, and spatio-temporal independent component analysis of fMRI data," in *Proc. Leeds Statistical Research Workshop*, Leeds, U.K., July 7–9, 1999, pp. 1–4.
- [45] V. D. Calhoun, T. Adali, G. D. Pearson, and J. J. Pekar, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," *Hum. Brain Mapp.*, vol. 13, no. 1, pp. 43–53, May 2001.
- [46] M. J. McKeown and T. J. Sejnowski, "Independent component analysis of fMRI data: Examining the assumptions," *Hum. Brain Mapp.*, vol. 6, no. 5–6, pp. 368–372, 1998.
- [47] V. D. Calhoun, T. Adali, G. D. Pearson, and J. J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis," *Hum. Brain Mapp.*, vol. 14, no. 3, pp. 140–151, Nov. 2001.
- [48] E. B. Erhardt, S. Rachakonda, E. J. Bedrick, E. A. Allen, T. Adali, and V. D. Calhoun, "Comparison of multi-subject ICA methods for analysis of fMRI data," *Hum. Brain Mapp.*, vol. 32, no. 12, pp. 2075–2095, Dec. 2011.
- [49] E. Seifritz, F. Esposito, F. Hennel, H. Mustovic, J. G. Neuhoff, D. Bilecen, G. Tedeschi, K. Scheffler et al., "Spatiotemporal pattern of neural processing in the human auditory cortex," *Science*, vol. 297, pp. 1706–1708, Sept. 2002.
- [50] J. H. Lee, T. W. Lee, F. A. Jolesz, and S. S. Yoo, "Independent vector analysis (IVA): Multivariate approach for fMRI group study," *NeuroImage*, vol. 40, no. 1, pp. 86–109, Mar. 2008.
- [51] R. M. Hutchison, T. Womelsdorf, J. S. Gati, S. Everling, and R. S. Menon, "Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques," *Hum. Brain Mapp.*, vol. 34, no. 9, pp. 2154–2177, Sept. 2013.
- [52] A. Zalesky and M. Breakpear, "Towards a statistical test for functional connectivity dynamics," *NeuroImage*, vol. 114, pp. 466–470, July 2015.
- [53] R. Meyer and N. Christensen, "Bayesian reconstruction of chaotic dynamical systems," *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*, vol. 62, no. 3, pt. A, pp. 3535–3542, Sept. 2000.
- [54] R. L. Miller and V. D. Calhoun, "Data fusion for multimodal multi-feature brain data using Markov-style dynamics in a feature meta-space," presented at the 2015 *Int. Workshop on Pattern Recognition in NeuroImaging (PRNI)*, Palo Alto, CA, June 10–12, 2015.
- [55] R. Miller and V. D. Calhoun, "Meta-state analysis reveals reduced resting fMRI connectivity dynamism in schizophrenia, with dynamic fluidity and range further suppressed by many individual symptoms," presented at the *Society for Neuroscience Annu. Meeting*, Washington, DC, Nov. 15–19, 2014.
- [56] M. Yaesoubi, E. A. Allen, R. Miller, and V. D. Calhoun, "Dynamic coherence analysis of resting fMRI data to jointly capture state-based phase, frequency, and time-domain information," *NeuroImage*, vol. 120, pp. 133–142, Oct. 2015.
- [57] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking whole-brain connectivity dynamics in the resting-state," presented at the *18th Annu. Meeting Organization for Human Brain Mapping*, June 10–14, 2012, Beijing.
- [58] E. Tagliazucchi, R. Carhart-Harris, R. Leech, D. Nutt, and D. R. Chialvo, "Enhanced repertoire of brain dynamical states during the psychedelic experience," *Hum. Brain Mapp.*, vol. 35, no. 11, pp. 5442–5456, Nov. 2014.
- [59] A. Kucyi and K. D. Davis, "Dynamic functional connectivity of the default mode network tracks daydreaming," *NeuroImage*, vol. 100, pp. 471–480, Oct. 2014.
- [60] U. Sakoglu, A. M. Michael, and V. D. Calhoun, "Classification of schizophrenia patients vs healthy controls based on dynamic functional network connectivity," presented at the *15th Annu. Meeting Organization for Human Brain Mapping*, San Francisco, June 18–23, 2009.
- [61] J. Gonzalez-Castillo, Z. S. Saad, D. A. Handwerker, S. J. Inati, N. Brenowitz, and P. A. Bandettini, "Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 109, no. 14, pp. 5487–5492, Apr. 2012.
- [62] L. Wu, T. Eichele, and V. D. Calhoun, "Reactivity of hemodynamic responses and functional connectivity to different states of alpha synchrony: A concurrent EEG-fMRI study," *NeuroImage*, vol. 52, no. 4, pp. 1252–1260, Oct. 2010.
- [63] D. A. Bridwell, L. Wu, T. Eichele, and V. D. Calhoun, "The spatio-spectral characterization of brain networks: Fusing concurrent EEG spectra and fMRI maps," *NeuroImage*, vol. 69, pp. 101–111, Apr. 2013.
- [64] E. A. Allen, T. Eichele, L. Wu, and V. D. Calhoun, "EEG signature of functional connectivity states," presented at the *19th Annu. Meeting Organization for Human Brain Mapping*, Seattle, WA, June 16–20, 2013.
- [65] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 98, no. 2, pp. 676–682, Jan. 2001.
- [66] L. Wu and V. D. Calhoun, "An approach for fusion between EEG and fMRI data," presented at the *16th ISMRM Scientific Meeting and Exhibition*, Toronto, May 3–9, 2008.
- [67] T. Eichele, V. D. Calhoun, and S. Debener, "Mining EEG-fMRI using independent component analysis," *Int. J. Psychophysiol.*, vol. 73, no. 1, pp. 53–61, Jul. 2009.
- [68] T. Deneux and O. Faugeras, "EEG-fMRI fusion of paradigm-free activity using Kalman filtering," *Neural Comput.*, vol. 22, no. 4, pp. 906–948, Apr. 2010.
- [69] M. Luessi, S. D. Babacan, R. Molina, J. R. Booth, and A. K. Katsaggelos, "Bayesian symmetrical EEG/fMRI fusion with spatially adaptive priors," *NeuroImage*, vol. 55, no. 1, pp. 113–132, Mar. 2011.
- [70] D. A. Handwerker, V. Roopchansingh, J. Gonzalez-Castillo, and P. A. Bandettini, "Periodic changes in fMRI connectivity," *NeuroImage*, vol. 63, no. 3, pp. 1712–1719, Nov. 2012.
- [71] E. Tagliazucchi and H. Laufs, "Multimodal imaging of dynamic functional connectivity," *Front. Neurol.*, vol. 6, p. 10, Feb. 2015.
- [72] E. Tagliazucchi, F. von Wegner, A. Morzelewski, V. Brodbeck, K. Jahnke, and H. Laufs, "Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 110, no. 38, pp. 15419–15424, Sept. 2013.
- [73] E. Damaraju, E. Tagliazucchi, H. Laufs, and V. D. Calhoun, "Dynamic functional network connectivity from rest to sleep," presented at the *OHBM 2015 Annu. Meeting*, Honolulu, HI, June 14–18, 2015.
- [74] J. Gonzalez-Castillo, C. W. Hoy, D. A. Handwerker, M. E. Robinson, L. C. Buchanan, Z. S. Saad, and P. A. Bandettini, "Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 112, no. 28, pp. 8762–8767, July 2015.
- [75] R. L. Miller, M. Yaesoubi, J. Turner, D. Mathalon, A. Preda, G. Pearson, T. Adali, and V. D. Calhoun, "Higher dimensional meta-state analysis reveals reduced resting fMRI connectivity dynamism in schizophrenia patients," *PLoS ONE*, to be published.
- [76] S. Ma, T. Eichele, V. D. Calhoun, W. Du, and T. Adali, "Modulations of functional connectivity in healthy and schizophrenia groups during task and rest," *NeuroImage*, vol. 62, no. 3, pp. 1694–1704, Sept. 2012.



©GRAPHIC STOCK

Habib Zaidi and
Minerva Becker

The Promise of Hybrid PET/MRI

Technical advances and clinical applications

During the last few decades, positron emission tomography (PET)-based molecular imaging has advanced elegantly and steadily gained importance in the clinical and research arenas. However, the lack of structural information provided by this imaging modality motivated its correlation with structural imaging techniques such as X-ray computed tomography (CT) or magnetic resonance imaging (MRI), which are well established in the clinical setting. The additional capability of simultaneous acquisition of PET and MRI data bridges the gap between molecular and morphologic diagnoses. Since diagnostic imaging methods evolve from the anatomical to the molecular level, the mission of multimodal and multiparametric imaging increasingly becomes more essential. Since 2010, whole-body hybrid PET/MRI has been investigated in the clinical setting for clinical diagnosis and staging, treatment response monitoring, and radiation therapy treatment planning of a wide range of malignancies. However, quantitative PET/MRI is still challenged by the lack of accurate and robust attenuation and motion compensation strategies to enable the production of artifact-free and quantitative PET images. This article briefly summarizes the historical development of PET/MRI and gives

Digital Object Identifier 10.1109/MSP.2015.2482225
Date of publication: 27 April 2016

an overview of the state of the art and recent advances in the design and construction of clinical systems. Progress in quantitative imaging, including MRI-guided image reconstruction and correction, and potential clinical applications of this novel technology are also discussed.

Introduction

PET is considered to be one of the key molecular imaging modalities enabling noninvasive characterization and quantitative evaluation of a multiplicity of molecular and physiologic biomarkers in vivo at the cellular level in healthy and disease states, including neurology, psychiatry, cardiology, and oncology. However, PET produces blurred and noisy images that inherently lack the anatomical information required for localization of metabolic abnormalities. This limitation has motivated the combination of PET with structural imaging modalities, such as X-ray CT and MRI. Currently, PET is capitalizing and complementing other anatomical imaging modalities, such as CT and MRI, to address basic research and clinical questions. However, multimodality imaging requires robust registration of images generated by various modalities. Initially, multimodality imaging was accomplished through the use of software-based image registration (rigid body or deformable) and fusion to correlate anatomical and molecular information [1]. However, the challenges and inherent limitations of software-based image registration approaches motivated the emergence of hardware-based approaches for multimodality imaging. The advent of combined PET/CT and PET/MRI systems, their commercial introduction, and the fast and wide acceptance of the former in the clinic have had a significant impact on patient management and clinical research. However, the latter is still an “embryonic” technology, having the potential to become a powerful tool and likely to play a pivotal role in clinical diagnosis and research [2], [3].

This article reviews the state-of-the-art developments and the latest advances in hybrid PET/MRI instrumentation along with quantitative procedures developed to address the challenges of this modality. An outlook outlining potential promising developments and current and future clinical applications of this technology is also discussed.

History of hybrid PET/MRI

The history of PET/MRI can be traced back to 1986, when the first attempts to perform PET imaging within strong static magnetic fields were initiated, motivated by the need to reduce positron range prior to annihilation through magnetic confinement of emitted positrons [4], [5]. Indeed, the static magnetic field of the MRI subsystem influences the trajectory of positrons, causing them to spiral between successive interactions with matter, thus reducing the in-plane spatial resolution of the PET subsystem. Monte Carlo simulation studies demonstrated that the use of a magnetic field collinear with a PET scanner's axis improves the transaxial spatial resolution without impinging on the axial spatial resolution [5]. For instance, Wirtzwar et al. [6] reported foreseen improvements in spatial resolution for high-energy positron-emitting tracers ranging between

18.5% (2.73 mm instead of 3.35 mm) for ^{68}Ga and 26.8% (2.68 mm instead of 3.66 mm) for ^{82}Rb at a field strength of 7 T. Another effect, which has been characterized only very recently, is the degradation of the axial spatial resolution owing to the elongation of the positron range distribution along the magnetic field B_0 or the so-called shine-through artifact [7]. It was reported that this effect might cause severe artifacts in PET images for malignant lesions located close to air cavities, particularly when using high-energy positron-emitting radionuclides (see the section “Pitfalls and Artifacts”).

Contrary to the history of PET/CT, which began with the design of hybrid systems suitable for clinical use, PET/MRI began with systems dedicated to preclinical imaging. Surprisingly, the history of hybrid PET/MRI instrumentation per se can be traced back to 1995, prior to the introduction of PET/CT [8]. Early designs of MR-compatible PET detector modules focused on modifying detector blocks of an existing small-animal PET scanner to avoid mutual interference by placing photomultiplier tubes (PMTs) at a realistic distance from the strong magnetic field of a clinical MRI unit [9]. For the sake of avoiding or reducing mutual interference between imaging modalities, the coupling of detector blocks, position-sensitive PMTs, and readout electronics located outside of the magnetic field was achieved through long (4–5 m) optical fibers. The main disadvantage was, however, the non-negligible loss of scintillation light through the long fibers, resulting in a weak signal, which negatively impacts energy and timing resolution, impairs deteriorating crystal identification, and decreases PET signal performance, reducing overall PET performance [10].

Although this design concept bears inherent limitations, analogous approaches were adopted in academic settings [11]. Other associated approaches based on conventional PMT-based PET detectors included split-magnet [12] and field-cycled [13] MRI, which rely on more complex magnet designs. In the split-magnet design, an 8-cm gap in the axial direction of a 1-T magnet enables accommodation of the microPET Focus 120 small-animal PET scanner (Siemens Healthcare, Erlangen, Germany) and 1.2-m-long optical fiber bundles [12], making it possible to place the PMTs at very low field strength (~30 mT). The main advantage of this design is the need for only minor modifications of conventional PET detectors and associated readout technologies, although the magnet and gradient coil design is more complex and costly compared with technologies used on current-generation MRI systems. In the field-cycled design, PMTs are assembled into the magnet, although PET data acquisition is barely permitted within short time intervals (~2.5 seconds) when MRI polarizing and readout fields are switched off [13]. The challenges associated with this design still need to be addressed before a viable hardware realization can be achieved. Moreover, the need for electromagnets instead of conventional superconducting magnets requires noteworthy compromises.

The introduction of MR-compatible readout technologies, such as avalanche photodiodes (APDs) and silicon photomultipliers (SiPMs), was essential to achieve this goal.

Avalanche photodiode-based readout technology was successfully employed on a commercial preclinical scanner [14] and various prototypes for small-animal [15] and breast [16] PET/MRI. Small pixelated APDs or SiPMs operated in “Geiger mode” and more recent readout technologies, such as analog [17] and digital [18] SiPMs, have been investigated as possible candidates for PET/MRI, and their current performance is sufficient for the design of combined PET/MRI systems [19], given that the bulk of MRI electronics could be significantly reduced [20]. Convincing experimental results and *in vivo* mouse images obtained on APD-based PET/MRI design demonstrate the capability for simultaneous PET/MRI [15]. More importantly, experimental measurements confirmed that each subsystem performs equally well when the other is on or off, reinforcing that each modality is barely visible to the other. These technological advances motivated additional exploration of the clinical potential of PET/MRI [21].

Design considerations of hybrid PET/MRI systems

Contrary to sequential PET/CT, where the design concept is straightforward and consists of putting together two separate modalities, the design of fully integrated PET/MRI systems is less obvious. Indeed, such development requires not only modifications of the PET subsystem to deal with MR computability but also significant redesign of the MRI subsystem [3], [10], [22]. Basically, two major design concepts for PET/MRI have emerged: sequential and concurrent [23] (Figure 1). In the former design concept, a serial arrangement of two separate scanners enables sequential data acquisition of both modalities using a single patient’s bed to transfer the patient from one modality to the other. Conversely, the latter consists of either an MR-compatible PET insert that can be placed with the MRI gantry or a compact integrated system enabling truly simultaneous data acquisition.

The sequential design is the more straightforward and by far the more economical concept, requiring only minor modifications of both subsystems (e.g., shielding the PET detectors) and arranging for a common patient bed. Sequential PET/

MRI systems were designed in anticipation of the availability of mature and economically viable simultaneous whole-body PET/MRI systems, which appeared later and became commercially available. Two design concepts have materialized depending on the configuration adopted for patient bed shuttling from one modality to the other. Systems belonging to the first category include the Ingenuity TF PET/MRI system (Philips Healthcare, Best, The Netherlands), in which a common sliding/rotating bed transfers the patient from MRI to PET and vice versa [24]. The PET/CT/MR trimodality imaging system (GE Healthcare, Little Chalfont, United Kingdom) consists of commercial PET/CT and MRI systems placed in separate but nearby rooms, and a specially designed patient transfer tabletop, docked on both imaging systems, is used to shuttle the patient from the PET/CT to MRI examination rooms [25]. A similar design concept dedicated to brain imaging was pursued by Cho et al. [26] by docking a high-resolution research tomograph and 7-T MRI.

The concurrent design of hybrid PET/MRI is possibly more attractive but is technically more challenging because it involves addressing many difficulties to deal with space restrictions and to avoid interference between the two modalities. To this end, MR-compatible photodetector technologies that are insensitive to magnetic fields and readout electronics producing the least amount of heat radiation have to be used [10]. In addition, the PET detector modules should not affect the operation of the MRI subsystem through electronic interference with the radio frequency (RF) and gradient coils. In essence, the operation of both modalities should not be affected by their integration, and both subsystems should retain their full performance, similar to what can be achieved with two separate PET and MRI scanners.

As mentioned in the previous section, recent developments in solid-state detectors have led to the replacement of conventional PMTs by MR-compatible position-sensitive APDs and SiPMs for the practical implementation of fully simultaneous PET/MRI systems. The PET insert concept, consisting of placing the detector ring inside an MRI scanner,

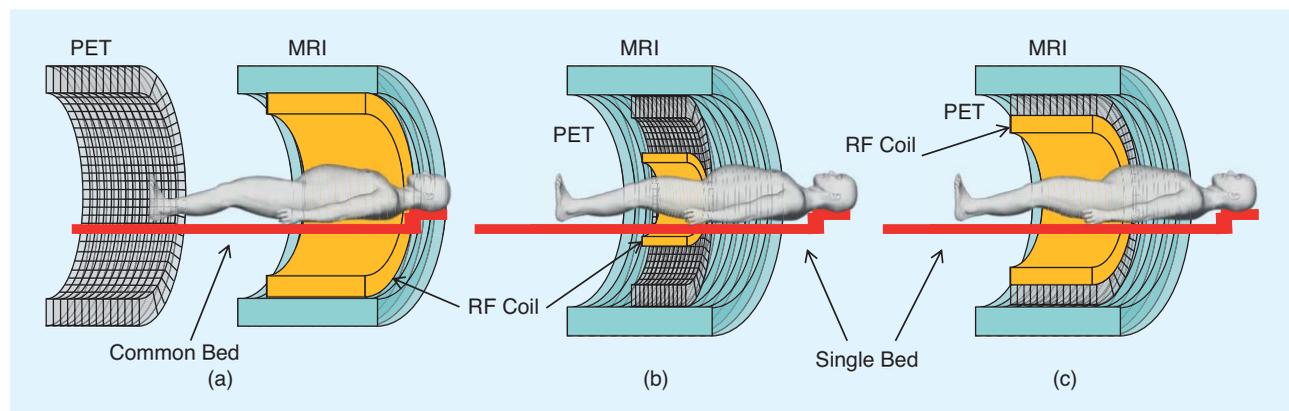


FIGURE 1. Schematic cross-sectional views of potential designs for combined PET/MRI systems: (a) a tandem design with two imagers mounted back to back (similar to that in PET/CT instrumentation) to allow sequential rather than simultaneous acquisition, (b) an insert design with the PET imager inserted between the RF coil and gradient set of the MR imager, and (c) a fully integrated design with two imagers in the same gantry. The RF coil, gradient set, PET imager, and patient bed are shown for all configurations. (Figure adapted with permission from [23].)

was the first landmark, and a number of studies have described different design trends, focusing particularly on the integration of small-bore, small-animal PET scanners inside existing clinical MRI scanners. The small diameters of these devices allows them to fit into the MRI system without crowding the MRI gradients.

As mentioned previously, fully integrated compact systems combining PET and MRI components in a single apparatus, such as Siemens Healthcare's Biograph mMR and GE Healthcare's SIGNA, constitute the most promising design concept for PET/MRI. The exploitation of the most advanced technologies available for both systems is advised to achieve the best performance. For instance, using a PET scanner equipped with time-of-flight (TOF) capability is certainly a bonus, as discussed in the following section. In this regard, SiPMs have many advantages compared with other solid-state photodetectors, such as APDs, because they have better performance characteristics, including high gain, signal-to-noise ratio (SNR), and timing resolution, enabling the implementation of TOF PET on potential PET/MRI systems.

Instrumentation for clinical PET/MRI

The successful design of small-animal PET/MRI systems spurred the development of clinical systems, with the first

prototype (called BrainPET) for brain imaging manufactured by Siemens Healthcare in collaboration with the University of Tübingen in Germany [27]. The system performance was characterized and its suitability for various clinical applications assessed at a number of academic institutions. Special attention was paid to the possibilities offered by high-resolution structural MRI, including high soft-tissue contrast sensitivity and advanced functional MRI techniques [28]. A sequential PET/MRI system was also developed to meet the needs of molecular and genetic brain imaging by docking separate PET and 7-T MRI scanners together with a shared common bed for interscanner patient translation [26].

Subsequent to early groundbreaking developments, different design concepts of PET/MRI systems have materialized during the last decade in both academic and corporate settings. Figure 2 shows photographs of current commercial clinical whole-body PET/MRI systems with potential design concepts. Table 1 summarizes the main characteristics of clinical PET/MRI systems developed so far.

The Ingenuity TF PET/MRI system, with TOF Gemini TF PET and Achieva 3T X-series MRI systems, is one such example, allowing for sequential acquisition of aligned PET and MR images. A number of such systems were deployed worldwide, and the PET subsystem was fully characterized using

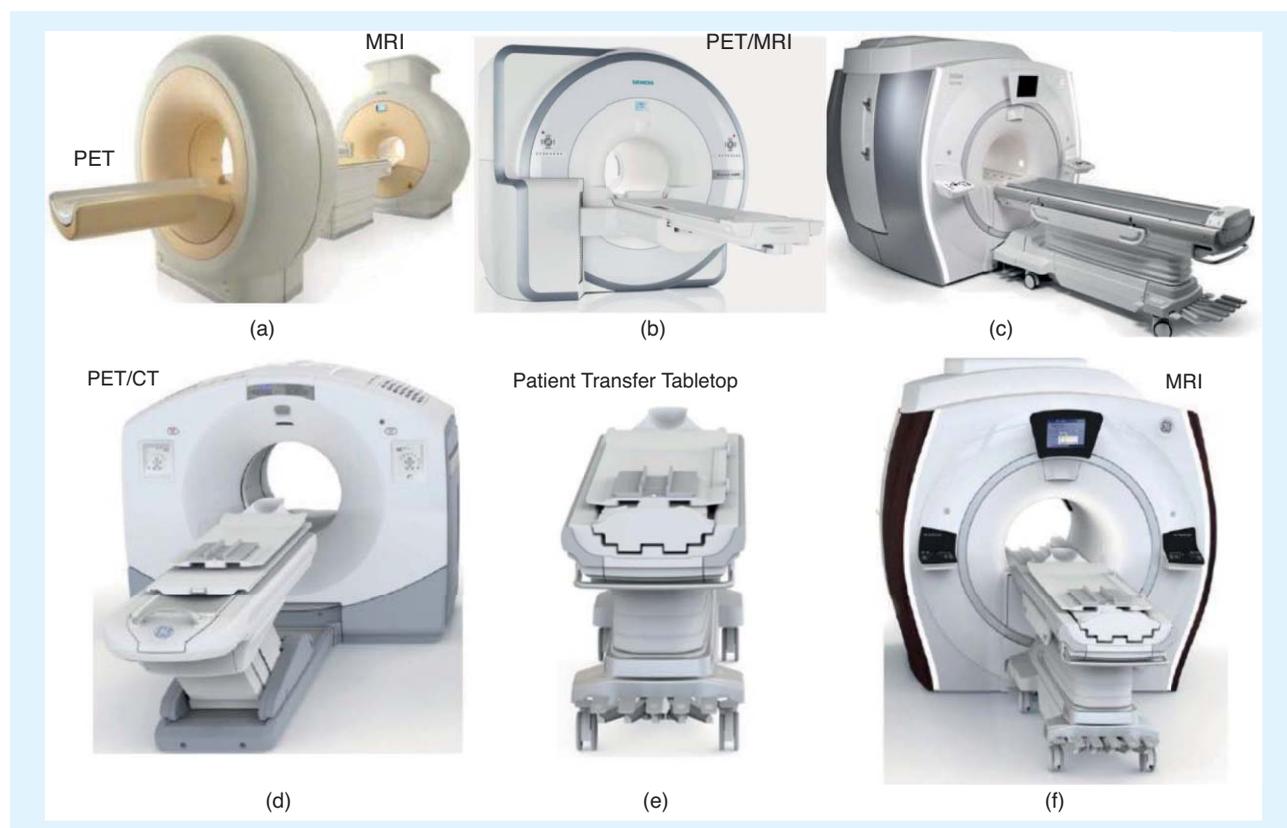


FIGURE 2. (a)–(c) The Philips Healthcare whole-body Ingenuity TF PET/MRI system [in which a turntable patient-handling system facilitates patient motion between the PET subsystem shown in (a) and the Achieva 3T X-series MRI system shown in (c) for sequential acquisition], the Siemens Healthcare Biograph mMR system, and the GE Healthcare SIGNA PET/MRI system, enabling simultaneous acquisition of PET and MRI data. (d)–(f) The GE Healthcare trimodality (PET/CT and MRI) setup using a dedicated patient transporter tabletop. [(a) and (c) used courtesy of Philips Healthcare, (b) courtesy of Siemens Healthcare, and (d)–(f) courtesy of GE Healthcare.]

Table 1. The main features of currently available clinical PET/MRI systems.

| System | Manufacturer | Operation | PET detector/readout | Axial FOV (cm) | TOF | MRI | Reference |
|---------------|--------------------|--------------|----------------------|----------------|-----|------------------------|-----------|
| Biograph mMR | Siemens Healthcare | Simultaneous | LSO/APDs | 25.8 | No | Verio 3T (modified) | [29] |
| Ingenuity TF | Philips Healthcare | Sequential | LYSO/PMTs | 18 | Yes | Achieva 3T | [24] |
| Signa PET/MRI | GE Healthcare | Simultaneous | LYSO/SiPMs | 25 | Yes | MR750w 3.0T (modified) | [17] |
| Trimodality | GE Healthcare | Sequential | LYSO/PMTs | 15.7 | Yes | MR750w 3.0T | [25] |
| BrainPET | Siemens Healthcare | Simultaneous | LSO/APDs | 19.2 | No | Trio 3T (modified) | [30] |
| Brain MGI | Academia | Sequential | LSO-LYSO/PMTs | 25.2 | No | Magnetom 7T | [26] |

Adapted with permission from [23].

the National Electrical Manufacturers Association (NEMA) NU 2-2007 standard, demonstrating that its performance was not compromised by the presence of the strong MR magnet [24]. Most performance parameters were comparable to those reported for the commercial Gemini TF PET/CT system.

The design concept of the concurrent BrainPET was further exploited to build the Siemens Healthcare Biograph mMR whole-body PET/MRI system, which was also installed in a relatively large number of institutions [29]. More recently, a simultaneous PET/MRI system (SIGNA) based on MR-compatible SiPMs and enabling the implementation of TOF capability was introduced in the market by GE Healthcare [17].

Most current PET/MRI systems have been tested within a high field and proved to produce PET and MR images that appear to be free of distortion, confirming the premise that the interference between the two systems is almost negligible and that each modality is practically invisible to the other [17], [24], [27], [29], [30]. Switching clinical workflows to PET/MRI introduces a number of image registration challenges that were not of major concern with traditional PET/CT scanners. These relate to the additional artifacts within MRI, such as bias fields, the range and number of MRI sequences, and the range of fields of view (FOVs) and orientations of the acquired images [31].

During the last decade, hardware and software advances have enabled improved localization of the position of annihilation along the line of response. The precise measurement of the difference between the arrival times of the two annihilation photons, referred to as TOF, enables more accurate localization of the annihilation point. However, the annihilation point could be located only with limited precision owing to inherent uncertainty in the detector modules and readout electronics, causing some ambiguity in the photon arrival times. As such, the incorporation of TOF information in the image reconstruction process enables improved SNR and tumor detectability in addition to reduction of patient scanning time and/or injected dose, all depending upon patient size and coincidence time resolution (CTR). The SNR improves as the CTR decreases, and this improvement becomes more significant for overweight patients. In a clinical setting, this results in a more homogeneous image quality across different (and increasing) patient sizes and overall

yields a much-improved image quality in shorter acquisition times, thus providing the possibility to investigate novel acquisition protocols, such as whole-body dynamic imaging. The SNR gain when using TOF is equivalent to a non-TOF image reconstructed using higher statistics; in this way, adding TOF information to PET increases the sensitivity of the scanner. In addition, TOF PET scanners are less sensitive to inaccuracies in normalization and data correction procedures, including attenuation compensation [32]. The first commercial TOF PET/MRI scanners using lutetium oxyorthosilicate (LSO)/lutetium-yttrium oxyorthosilicate (LYSO) crystals and PMT/SiPM photodetectors have a time resolution of 400–600 picoseconds [17], [24]. APD-based hybrid PET/MRI systems, including the BrainPET and Biograph mMR scanners, are not equipped with TOF capability owing to the poor timing resolution of APDs. A CTR of fewer than 100 picoseconds has been obtained with short crystals of 3–5 mm [33], [34]. The interaction length of 511-keV photons in LSO is 12 mm. As such, achieving a sensible detection efficiency requires 15–20-mm-long crystals. However, the CTR degrades with increasing length owing to the reduction in the speed of light in the high refractive index of the scintillator because the position along the length of the crystal where the interaction of the 511-keV photon occurred is unknown. With advances in detector technology and fast electronics, a TOF PET/MRI scanner with sub-100-picosecond CTR will likely be possible in the near future (Figure 3). The target in the long term is to attain the physical limit of spatial resolution for clinical scanners (~2 mm), and by definition, a target CTR of 20 picoseconds would be required to obviate the need of image reconstruction.

Hybrid small-animal PET/MRI is also flourishing in both academic and corporate settings, with several prototypes based on different design concepts and a number of companies already offering commercial solutions [15], [22]. The potential benefits of compact and integrated systems were already recognized, and it is expected that this technology will find a niche in preclinical research, which is well under way [35].

Quantitative PET/MRI

PET/MRI shows promise for radiotracer uptake quantification via image fusion of molecular and structural data to assist in anatomical localization of functional abnormalities

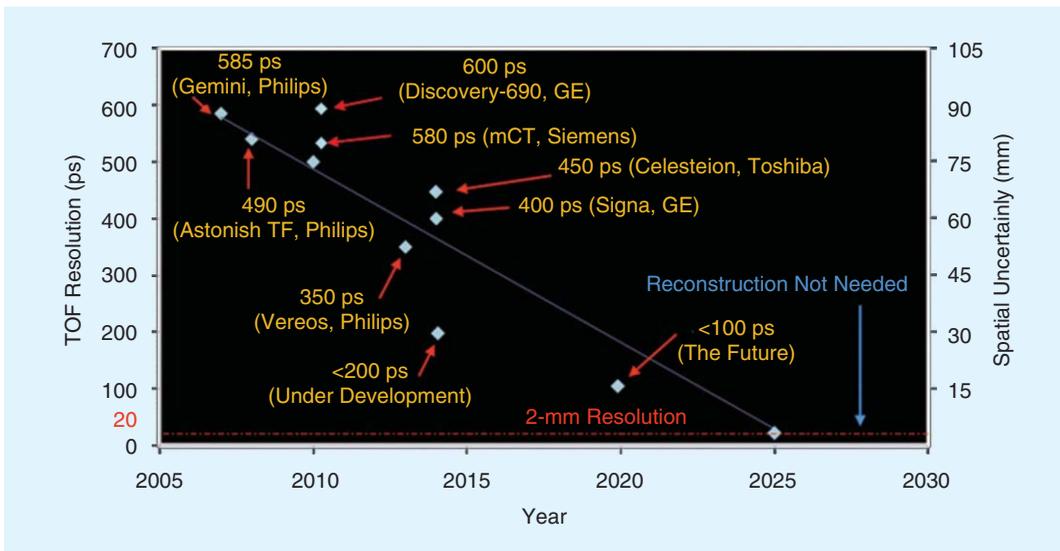


FIGURE 3. The evolution of TOF resolution performance characteristics of current-generation and future-generation TOF PET scanners. PS: picoseconds.

and delineation of regions of interest (ROIs) for quantitative analysis. However, there are several challenges undermining the widespread adoption of this technology, which may, in fact, represent inherent limitations. Similar to CT in PET/CT, MRI provides the structural information suitable for implementation of attenuation compensation techniques and introduction of a priori anatomical information into image reconstruction, partial-volume correction, and motion correction schemes. However, contrary to PET/CT, in which CT-based attenuation correction is straightforward, MRI-guided attenuation correction is challenging and still requires further development [36]. Owing to its clinical relevance and the challenges faced, the latter issue is addressed in more detail in this article.

MRI-guided attenuation correction in PET/MRI

The development of MRI-guided attenuation correction algorithms has received considerable attention during the last

decade. This was motivated by the lack of space in PET/MRI systems, precluding placement of external radionuclide sources within the gantry. MRI-guided attenuation correction is, however, still in its infancy and remains extremely challenging for whole-body imaging. The impact of this limitation on clinical interpretation of findings and patient outcome is not yet clear.

MRI-guided attenuation correction is complex because MRI signal intensity is not correlated with electron density, thus making conversion of signal intensity to attenuation coefficients complicated (Figure 4). MRI-guided attenuation map derivation consists of locating and mapping various biological tissues with different attenuation properties in the body. This can be achieved by one of the three main categories of techniques: 1) MRI segmentation-based techniques, in which the body is segmented into regions corresponding to tissues/organs with different attenuation properties, followed by assignment of corresponding linear attenuation coefficients at 511 keV to the segmented tissues/organs; 2) atlas-based and machine-learning

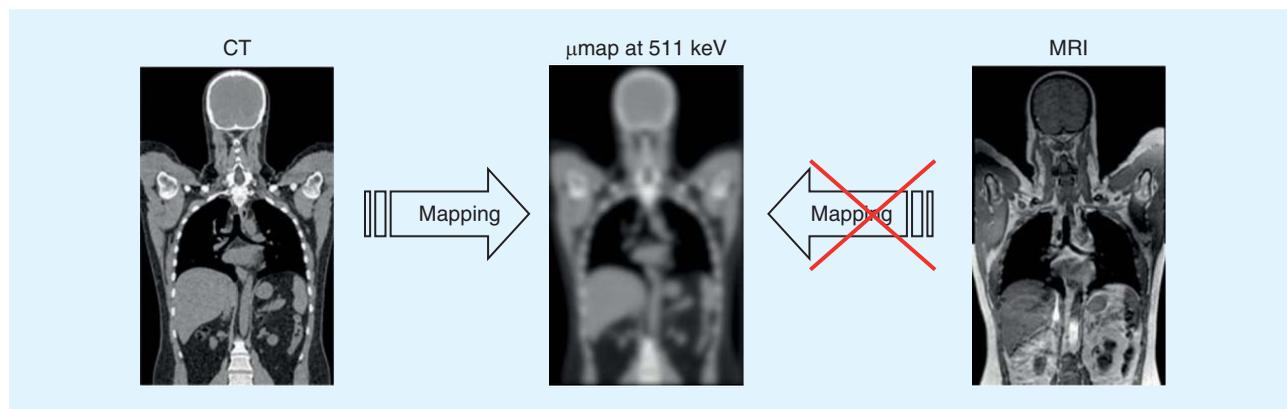


FIGURE 4. The conversion of CT images (Hounsfield units) to an attenuation map at 511 keV is evident (in the absence of sources of error), whereas the MR intensity level is not directly related to electronic density, which renders the conversion of MR images to attenuation maps less evident compared with CT.

techniques, in which an aligned MR/CT atlas combined with a learning strategy enables the prediction of the pseudo-CT from an actual patient's MR image; and 3) emission-based and transmission-based algorithms, in which the TOF emission or transmission data are exploited to derive the attenuation map (Figure 5) [37].

Segmentation-based methods are simple to implement and usually require a single and fast MRI sequence. However, they suffer from limited accuracy in the determination of attenuation coefficients owing to the limited number of segmented clusters (usually three to five, including air, lungs, fat, soft tissue, and fat/nonfat mixture) and the assignment of theoretical rather than actual patient-specific attenuation coefficients. In these techniques, bones and air pockets are replaced by soft tissue, and the variability of attenuation coefficients is ignored, especially in the lungs. Tissues such as bone and lung and various pathological abnormalities with varying attenuations are among the most challenging in whole-body imaging. With the exception of the use of ultrashort echo time [38] and zero echo time [39] pulse sequences, cortical bone has very low signal intensity on conventional MRI sequences and is difficult to distinguish from air cavities and gas in the body. These sequences were designed to portray tissues with low proton density and short T2 relaxation time (e.g., cortical bone and lungs) and, as such, to separate the bone signal from soft tissue. The main drawback of these techniques is the long acquisition

time and susceptibility to artifacts when using a large FOV, which limits their application to only brain imaging [37]. A number of studies have shown that ignoring bone might not be adequate for quantification of osseous lesions with bias in estimation in tracer uptake [standardized uptake value (SUV)], varying between 5 and 15% in most cases but going up to 30% in some cases [40]–[45].

The second category of approaches consists of using representative anatomical atlas registration, in which an MRI template is registered to a patient's MRI, and prior knowledge of the atlas attenuation properties, obtained by registration to a corresponding CT template combined with a learning algorithm based on the use of support vector machines, is applied to derive a patient-specific attenuation map [46]. Reliable deformable registration algorithms play a pivotal role in this approach, and failure of the registration process in the case of large deformations will produce incorrect results [47]. The critical issue is the extent to which the global anatomy depicted by an atlas will predict individual and patient-specific attenuation maps. For this and a few other reasons, most techniques proposed so far that belong to this category were developed specifically for brain imaging [48], [49]. Adaptation of these techniques for whole-body imaging applications required few modifications to be made, consisting mainly of generating a four-class segmentation of the MR images to improve the registration process and optimal selection of

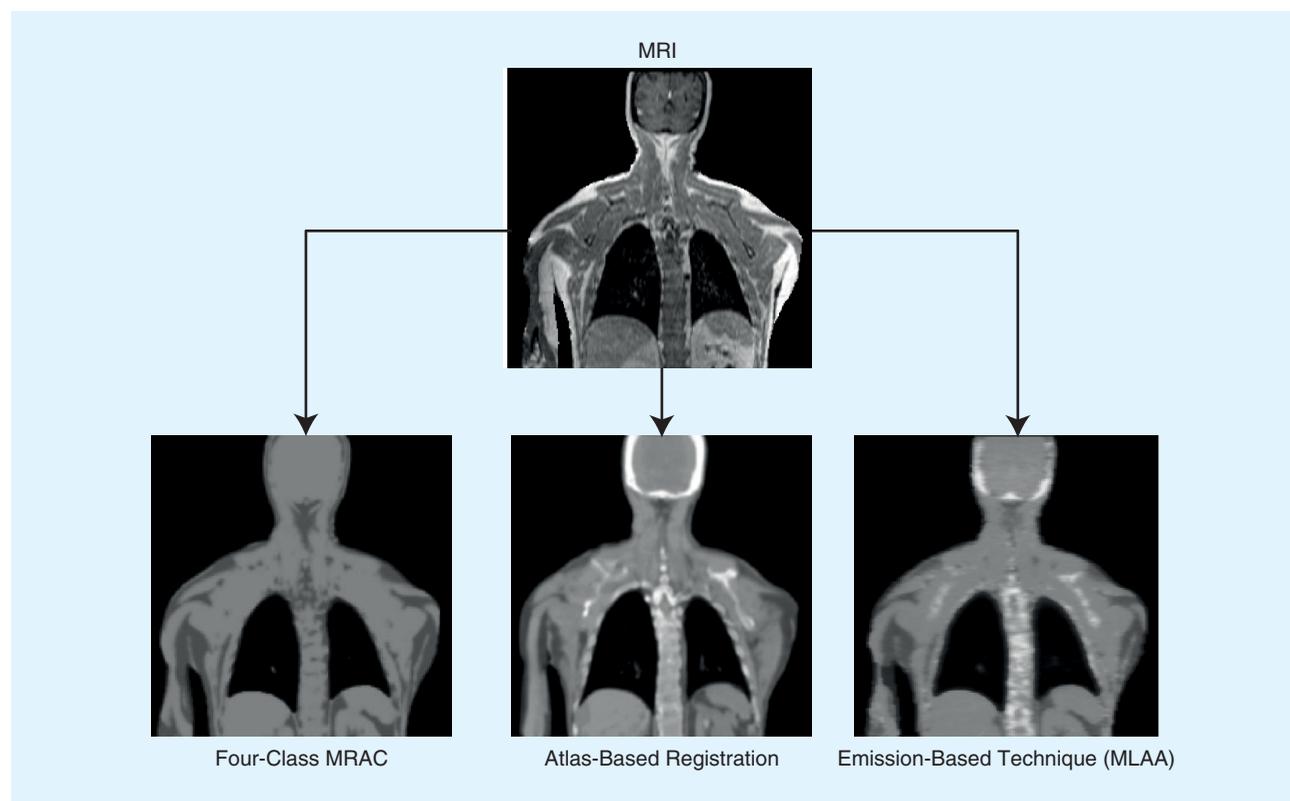


FIGURE 5. Strategies for MRI-guided attenuation map generation, including the four-class segmentation-based method, atlas-based registration and machine learning, and MRI-guided emission-based technique (MLAA). MRAC: MRI-based attenuation correction. (Figure adapted with permission from [37].)

regions for the learning process and applying postprocessing techniques to determine the tissue class for which sufficient information is available from the MRI. In a more recent contribution, Arabi and Zaidi [50] improved the robustness of the aforementioned technique [46] to nonsystematic registration bias and anatomical abnormalities by discarding locally gross misalignment errors from the training and pseudo-CT generation process through local sorting of the atlas images using the local normalized cross-correlation criterion as a metric to assess the similarity to the target image prior to providing it to the training step. Despite promising preliminary results reported in a number of studies using more advanced approaches [48], [50], more research is still required to make the procedure completely automated and suitable for clinical usage in whole-body PET/MRI.

Emission-based techniques form the last category of algorithms and have gained substantial momentum during the last decade. They are now recognized as valuable approaches for estimation of the attenuation map in PET/MRI through the simultaneous estimation of activity and attenuation within a maximum-likelihood (MLAA) framework [51]. However, these techniques suffer from cross-talk, depend on tracer distribution, and are susceptible to counting statistics. The use of TOF information proved to partially mitigate the cross-talk issue and stabilize the joint estimation problem [52]. It is worth emphasizing that TOF PET is less sensitive to attenuation artifacts than conventional non-TOF PET. Recent advances in emission-based techniques demonstrated the promise of an MRI-guided MLAA algorithm for attenuation correction in whole-body PET/MRI [53]. In this work, the estimation of attenuation maps takes advantage of a constrained Gaussian mixture model and Markov random field smoothness prior imposed by MRI spatial and CT statistical constraints. These techniques proved to outperform previous approaches reported in the literature [54]. Overall, each category of techniques has its own pros and cons, and it is expected that hybrid techniques combining at least two (and ideally the three categories of attenuation correction methods) will result in more accurate and robust techniques.

The many other challenging issues that still have to be addressed in this regard, including attenuation of MRI hardware (tables, rigid and non-rigid RF coils, pillows, headphones, medical probes, and other objects that are MRI-invisible but contribute to photon attenuation), patient positioning aids in the FOV, and conductive MR-compatible or nonconductive but MRI-invisible implants, should also be taken into account. Another challenging issue is transaxial plane truncation owing to the limited MRI FOV, which results in incomplete attenuation maps, producing artifacts on corresponding attenuation-corrected PET images.

MRI-guided image reconstruction in PET/MRI

One of the important limitations of statistical iterative reconstruction techniques, such as the maximum-likelihood-expectation-maximization (ML-EM) algorithm, is that a large number of iterations deteriorate image quality and

amplify noise in PET images [55]. An elegant way to control the noise characteristics consists of incorporating a prior distribution to depict the statistical properties of the image to be determined and thus generate a posteriori probability distributions from the image conditioned upon the data [10]. The well-established Bayesian reconstruction framework forms a prevailing expansion of the popular ML-EM algorithm. The maximum a posteriori (MAP) estimate is derived from maximization of the a posteriori probability over the set of probable images [56]. There are many benefits associated with this approach in the sense that the diverse mechanisms of the prior, including the pseudo-Poisson nature of statistics, nonnegativity of the solution, local voxel correlations, or identified presence of anatomical boundaries (from correlated structural imaging), may be incorporated into the estimation process, evaluated independently, and employed during the algorithm's implementation [10]. Prior anatomical information obtained from correlated anatomical imaging can also be included in PET reconstruction within a Bayesian framework to avoid resolution loss resulting from regularization, albeit to recover resolution by taking advantage of the better resolution of anatomical images [57]. This has been achieved with various degrees of success over the years using MRI [58].

A coupling term is usually incorporated in this category of reconstruction techniques, which favors the preservation of edges in PET images related to the location of relevant anatomical boundaries extracted from corresponding anatomical images. A Gibbs prior distribution is typically used to encourage the piecewise smoothness of PET images, which can be included in the Bayesian model. Promising preliminary results were reported by various investigators using segmentation-free anatomical priors based on similarity measures analogous to mutual information, but further research and development efforts are still required. Therefore, the advent of simultaneous hybrid PET/MRI systems creating perfectly aligned molecular and anatomical images is stimulating the further development and assessment in the clinical setting of Bayesian MAP reconstruction algorithms.

As an example, a MAP algorithm for PET image reconstruction incorporating MRI information with joint entropy between PET and MRI features serving as the regularization constraint was proposed [59]. A nonparametric method was then used to estimate the joint probability density of PET and MR images. It was demonstrated that incorporation of the anatomical information using this approach, following parameter optimization, produces significant improvement in the noise versus bias tradeoff in ROI-based quantitative analysis compared with conventional MAP reconstruction.

MRI-guided partial-volume correction in PET/MRI

The accuracy of PET for measuring regional radiotracer concentrations is limited by the finite spatial resolution capability of current-generation clinical PET scanners and the resulting partial-volume effect (PVE). Accurate PET quantification requires that the source size be greater than two to three times the scanner's spatial resolution, expressed in terms of

full-width at half-maximum. Any object with smaller dimensions only partly occupies this characteristic volume, such that acquired counts are spread over a larger volume owing to the limited spatial resolution of the PET scanner. Under these circumstances, corresponding PET images still reproduce the total amount of radiotracer uptake within an object but do not represent the regional activity distribution within this volume. A number of strategies have been proposed to correct for PVE [60]. The most straightforward approach uses recovery coefficients, which can be determined through experimental studies involving the use of spheres of different sizes. This simple approach produces acceptable results for objects with similar shape as the calibration phantom used for derivation of recovery coefficients (e.g., tumors of spherical shape). More refined approaches rely on anatomically guided postreconstruction techniques, in which the size and shape of corresponding objects assessed by structural imaging (MRI or CT) [61] are used instead to correct for this effect.

The PVE is among the major concerns in brain PET imaging in connection with quantification of cerebral metabolism in the atrophied brain, such as with Alzheimer's disease. Various voxel-based MRI-guided PVE correction methods have been proposed. The most popular technique consists of segmenting MR images into white and gray matter after PET/MRI registration. This is followed by convolving the segmented white and gray matter images by a Gaussian point spread function representing the PET scanner's spatial resolution. The PVE-corrected gray matter PET image is achieved by subtracting the convolved PET white matter image from the original PET image, followed by division by the convolved gray matter MR image. The final step involves the application of a binary mask to the gray matter region [62].

The overall accuracy achieved by MRI-guided PVE correction in PET depends upon the accuracy achieved by each procedural step, including image registration and MRI segmentation. This has been investigated in detail for the voxel-based approach [61]. The high soft-tissue contrast provided by MRI provides reasonable accuracy in terms of differentiation between gray and white matter. Nevertheless, errors in segmentation of brain tissue components have been found to be of greater significance [63]. For instance, a 25% error in total volume produces a 5% decrease in the caudate nucleus apparent recovery coefficient [64]. It is interesting to note that the effect of segmentation error is limited to the missegmented region. Inaccuracies from segmentation can be regarded in the framework of a more broad question of tissue heterogeneity. In fact, the main limiting feature of these algorithms is the assumption regarding the homogeneity of radiotracer distribution in each region or tissue component. Overall, it appears that the success of MRI segmentation has a higher impact on the accuracy of the corrected estimates [63] compared with the influence of image registration, although some studies seem to suggest that registration errors have the greatest impact on data accuracy and precision [61].

More refined strategies using multiresolution synergetic approaches merging anatomical and functional information

seem to have the potential to overcome the limitations of classical techniques. However, their feasibility in a clinical setting still needs to be demonstrated [65]. PVE correction can also be included directly into statistical reconstruction algorithms through the use of an appropriate mathematical formulation of PVE in the forward model along with other physical degrading factors governing the physics of PET [58].

MRI-guided motion compensation in PET/MRI

Recent advances in PET instrumentation have made it possible to achieve high spatial resolution, which motivates further development and clinical implementation of sophisticated motion correction strategies. The various sources of motion, including unwanted patient motion, cardiac motion, and respiratory motion, and correction strategies specifically developed to reduce or eliminate them have been reviewed recently [66]. Overall, three broad approaches were reported in the literature: 1) nonrigid registration of independently reconstructed images; 2) initial estimation of motion information from gated PET or MR/CT images, subsequently used in a new reconstruction applied to all gated frames; and 3) simultaneous estimation of motion parameters and images.

Motion between or during anatomical/molecular data acquisition remains an important challenge for PET/MRI protocols. The characteristic misalignment between PET and CT images at the level of the diaphragm in PET/CT systems resulting from breathing pattern differences is expected to be partly addressed by PET/MRI owing to the longer acquisition time of MRI sequences used for attenuation correction, which results in temporal averaging that would improve PET and MRI registration in some situations. In addition, the use of a specific respiratory protocol in PET/MRI can improve the spatial correspondence between PET and MRI. Owing to the typical duration of PET data acquisition (2–3 minutes/bed position), a PET image corresponds to an average of several respiratory cycles and is susceptible to motion-related distortion. Similarly, typical low-resolution MR images suitable for attenuation correction involve averages over multiple respiratory cycles, although the averaging process in MRI is different from that in PET. More importantly, severe motion artifacts may appear when there is marked organ motion with increased noise and smaller-appearing organ size on the MRI attenuation map, with subsequent bias in the attenuation correction procedure. Ideally, PET and MR images should correspond to the same phase of the respiratory cycle and be matched to achieve accurate attenuation correction and improved spatial resolution. To achieve good matching between PET and MR images at a specific respiratory phase, the patient's breathing during scanning should be synchronized to reduce distortional effects of respiratory motion. Provision of breathing instructions to patients prior to scanning may also be useful.

An assortment of MRI motion-tracking methods predominantly for rigid-body motion have been employed in the clinical setting, including, but not limited to, embedded cloverleaf navigators [67]. One such technique uses motion estimates derived from high temporal resolution MRI during simultaneous

acquisition of structural or functional MRI data for motion correction of corresponding brain PET data, demonstrating that MRI-derived motion can be used to improve PET image quality, thus increasing confidence in interpretation, reproducibility, and quantitative accuracy [68].

Current trends focus on four-dimensional MR-derived motion correction strategies to reduce artifacts observed in PET/CT by developing MRI-guided motion-compensated PET attenuation correction schemes, and research efforts should focus on designing suitable protocols to minimize MRI artifacts while also reducing mismatch between MR and PET images. Motion-free PET images are obtained by correcting motion-related blurring through MRI-derived motion estimates, allowing for improved image quality and accurate quantification of PET data compared with correction using PET-only motion information. Concurrent PET/MRI can also enable potential nonrigid motion compensation in whole-body PET imaging without increasing acquisition time [69]. Furthermore, three-dimensional cine MRI sequences for tracking the position and deformation of organs can be used to derive deformation fields for incorporation into statistical PET image reconstruction algorithms, although this approach can be complex [70], [71].

The use of tagged MRI for motion tracking in the phase domain to derive motion estimates in deformable registration during concurrent PET/MRI data acquisition was recently reported [72]. The conventional harmonic phase tracking technique is regularized, and the derived motion fields are incorporated in the system matrix of a statistical PET reconstruction algorithm. Preliminary results using computer simulations and a deformable phantom appear promising. Further investigation reported in more recent studies demonstrated the full potential of MRI-guided motion correction and its feasibility in clinical and research settings [73]–[75].

Clinical applications of PET/MRI

Hybrid whole-body human PET/MRI systems have been available since 2010; however, despite the initial excitement, the implementation of these systems in a clinical environment is still in an early phase. Among all potential clinical PET/MRI applications, recent advances in adult and pediatric oncology emerge as the most promising application fields. So far, only very few studies based on relatively small sample sizes have addressed the clinical workflow, feasibility, and optimized PET/MRI protocols in oncology [76]–[78]. Although some authors report no added diagnostic benefit in comparison with PET/CT or MRI [76], [79], [80], [81], others describe added value in selected cases [82]–[84]. During the past two years, there has been only a slow increase in the rate of new PET/MRI system installations, mainly owing to lack of clearly defined applications, and currently, there is little evidence to validate key applications based on the experience of multiple centers. This lack of significant progress can probably be explained by the fact that both MRI and PET/CT are powerful methods that are already well implemented in everyday clinical oncology.

Clinical workflow and protocols

A particular problem hampering the clinical implementation of PET/MRI in oncology is the lack of standardized imaging protocols and workflows because large variations in MRI protocols, sequences, and image requirements exist. Currently, most PET/MRI examinations in oncology are obtained using multistep protocols, which are very similar in design to PET/CT protocols. PET/CT protocols include a total-body low-dose CT, followed by a PET acquisition. In some institutions, contrast-enhanced CT is additionally obtained to avoid false-negative or false-positive PET readings, thereby improving the overall diagnostic yield [85]. In analogy to PET/CT, PET/MRI examinations can be performed by obtaining a rapid total-body MRI sequence for attenuation correction and localization of focal uptake, which is then followed by a whole-body PET. Although this approach is time-effective regardless of scanner type (simultaneous or sequential), the approach is not optimal for pretherapeutic tumor staging because it provides neither detailed anatomical nor functional MRI information.

The second PET/MRI approach consists of a rapid total-body PET/MRI acquisition and an additional high-resolution MRI examination of anatomical ROIs depending on the clinical situation. In simultaneous systems, this full diagnostic MRI (anatomical, diffusion-weighted, and perfusion sequences) can be performed during the PET acquisition, whereas in sequential systems, it is usually performed during the 60 minutes necessary for tracer uptake and prior to PET acquisition. However, not all MRI sequences can be acquired during the PET acquisition even in simultaneous scanners, and the total PET/MRI in-room time may be quite long (60–90 minutes), with most of the time being allotted to the MRI acquisition [77], [86], [87]. Therefore, in practical terms, the length of the MRI acquisition is a major limiting factor in the clinical implementation of PET/MRI in everyday routine provided that full use of multiparametric MRI capabilities is sought.

To compete with PET/CT acquisitions, which usually take around 30–40 minutes, some authors [31], [93] have suggested limiting the number of MRI sequences to the absolute minimum necessary for the oncologic diagnosis. However, there is no consensus today on the essential sequences necessary for tumor imaging, and different investigators have proposed different protocols. These are based not only on institutional preferences and technical parameters specific to different vendors but also on ongoing research protocols, time-effectiveness issues, cost, and—last but not least—the type of tumor to be imaged [76], [77], [88]. For example, there is increasing evidence supporting the utility of routinely obtaining diffusion-weighted imaging (DWI) in head and neck cancer and lymphoma. Whole-body MRI with DWI has a high sensitivity (96–97%) in the detection of lymphoma [89]. Gu et al. [89] evaluated whole-body MRI without and with DWI in the detection of known ^{18}F -Fluorodeoxyglucose (^{18}F -FDG)-avid lymphomas in 17 adult patients. By adding DWI to anatomical MRI, sensitivity was increased from 89% to 97% ($p = 0.002$); in particular, the accuracy for detecting nodal and extranodal disease in the abdomen and pelvis was improved, but without

affecting overall staging. Lin et al. [90] suggested that whole-body DWI may help to identify additional lymphoma lesions relative to the lesions already identified by ^{18}F -FDG PET/CT, whereas Punwani et al. [91] reported the complementary value of DWI to ^{18}F -FDG PET for prediction of site-specific interim response to chemotherapy, thereby supporting the need to incorporate DWI sequences into integrated PET/MRI protocols for lymphoma. In a similar fashion, Varoquaux et al. [92] showed that, although DWI and FDG PET may both reflect increased cell proliferation in head and neck squamous cell carcinoma, the two modalities refer to different biological phenomena, and their respective metrics, apparent diffusion coefficient (ADC) and SUV, are independent biomarkers, thereby having the potential to provide complementary information (Figures 6 and 7). The same authors showed that measurements of ADC and SUV values are reproducible with almost perfect interobserver and intraobserver agreements for both methods, and they observed a trend toward higher SUV and lower ADC values in poorly differentiated head and neck cancers compared to their well-differentiated or moderately differentiated counterparts [92].

In an attempt to reduce the total number of MRI sequences, some investigators have questioned the use of contrast-

enhanced MRI sequences [93]. However, in head and neck squamous cell carcinoma, contrast-enhanced MRI sequences are superior to CT and PET/CT for a variety of findings that are essential for the therapeutic choice, such as invasion of the skull base, perineural spread, and retropharyngeal lymph nodes or detection of extranodal spread in metastatic lymph nodes [86]. In addition, the combination of contrast-enhanced T1-weighted and T2-weighted sequences allows more precise differentiation between tumor and peritumoral inflammation, and it appears that the differentiation between these two conditions on the basis of MRI signal intensity characteristics can have direct implications on patient outcome after radiation therapy [86]. Morphologic MRI also appears to provide a higher accuracy than FDG PET/CT in detecting residual and/or recurrent nasopharyngeal carcinoma, and the combination of PET/CT and MRI seems to be superior to either modality alone for the detection and precise locoregional evaluation of recurrent disease [94]. In other tumor types, such as in breast cancer, the use of contrast-enhanced MRI sequences is essential. As shown by Taneja et al. [81], the morphologic MRI appearance of a breast lesion (size, shape, and pattern of enhancement) and its time–signal intensity curve after intravenous contrast material (progressive, plateau, or washout) yield

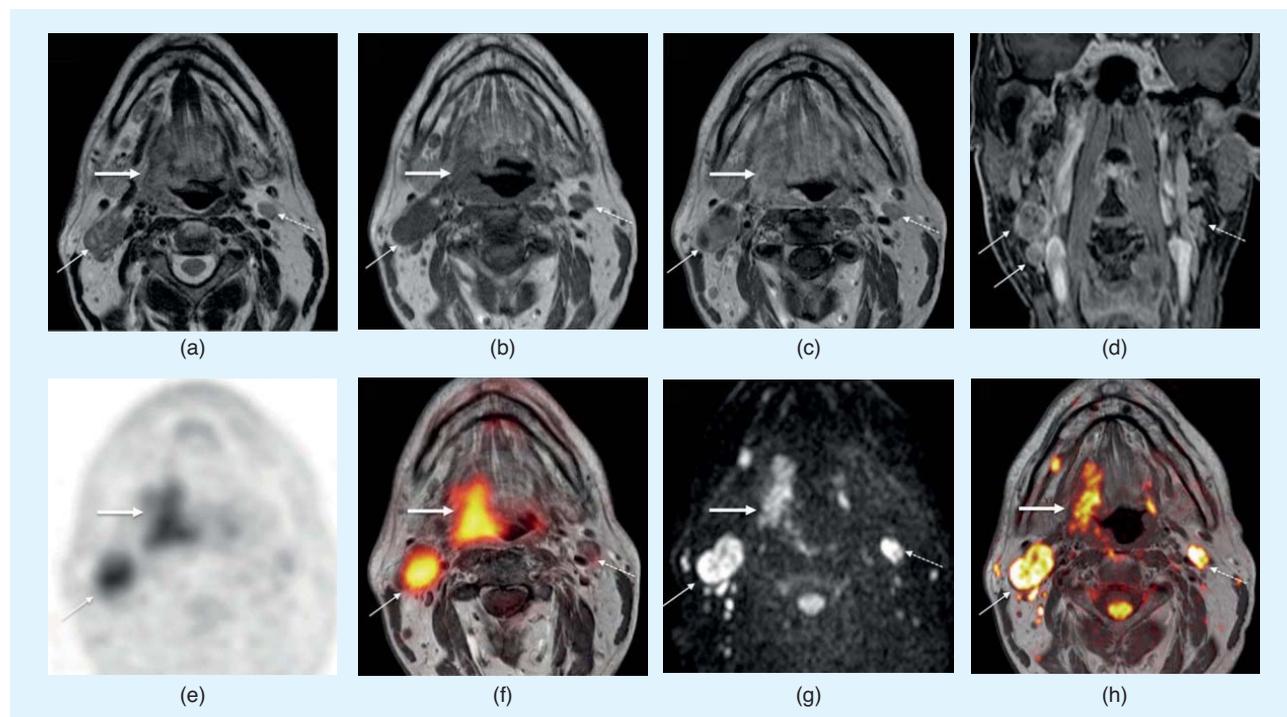


FIGURE 6. Multiparametric PET/MRI in head and neck squamous cell carcinoma. Complementary values of DWI, PET, and contrast-enhanced images: (a) Axial T2, (b) T1, (c) contrast-enhanced T1, (d) coronal contrast-enhanced fat-saturated T1, (e) axial PET, (f) fused PET with contrast-enhanced T1, (g) axial b1000 image from DWI, and (h) fused b1000 with contrast-enhanced T1. Right tonsillar cancer (thick arrows) invading the base of the tongue and posterior oropharyngeal wall and two right level 2 lymph node metastases are shown [thin arrows in (a)–(d)]. (c) Contrast-enhanced T1 reveals nodal necrosis, and (d) the corresponding fat-saturated T1 with gadolinium shows extranodal spread particularly well [spiculated margins in (d), thin arrows] not revealed by (a) and (b). (e) and (f) show high FDG uptake in the tumor ($\text{SUV}_{\text{max}} = 9.2$) and, on the right, level 2 lymph nodes ($\text{SUV}_{\text{max}} = 7.5$). (g) and (h) show restricted diffusivity in the tumor (mean $\text{ADC} = 1.0 \times 10^{-3} \text{ mm}^2/\text{second}$) and in the ipsilateral lymph nodes (mean $\text{ADC} = 1.16 \times 10^{-3} \text{ mm}^2/\text{second}$). These findings were confirmed histologically. A contralateral 8-mm large level 2 lymph node (dashed arrows) with restricted diffusivity ($\text{ADC} = 1.08 \times 10^{-3} \text{ mm}^2/\text{second}$) is seen in (g) and (h). Based on DWI, this left level 2 lymph node is considered suspicious for metastasis, although the FDG uptake and the morphologic aspect instead suggest a benign lymph node. Histology revealed moderately differentiated squamous cell carcinoma of the tonsil with bilateral lymph node metastases.

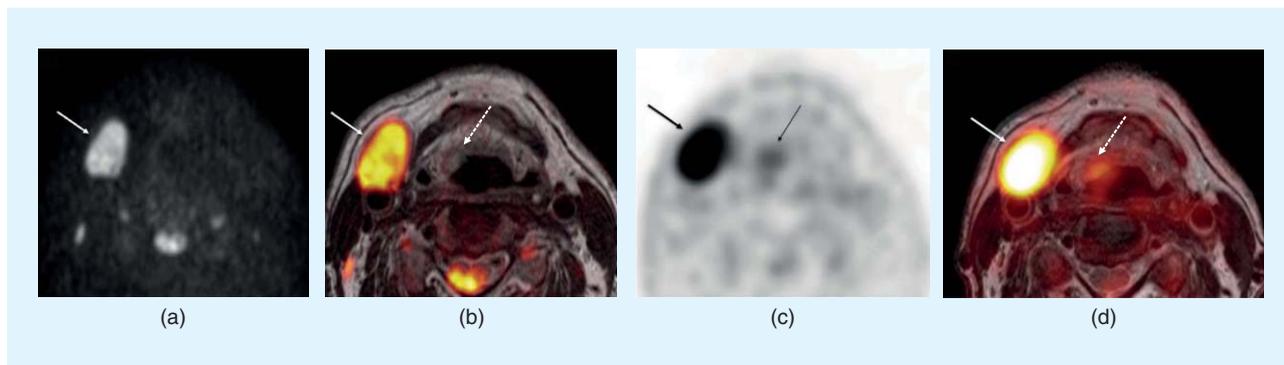


FIGURE 7. The detection of an unknown primary tumor with PET/MRI in a 60-year-old patient with nodal metastasis from squamous cell carcinoma. Endoscopy performed prior to PET/MR did not reveal a primary tumor. (a) b1000 image from DWI and (b) fused b1000 with T2-weighted image show a large level 2 lymph node metastasis (thick arrows) but no clearly identifiable primary tumor with restricted diffusivity. (c) A corresponding PET image and (d) fused PET with T2-weighted image confirm the metastatic node (thick arrows) and also reveal a small suspicious area located in the right pre-epiglottic space (thin dashed arrows). A repeat deep biopsy showed poorly differentiated squamous cell carcinoma located beneath an intact mucosa of the right epiglottis. Retrospectively, a small nodule can be seen on the T2-weighted image (dashed arrow).

a higher sensitivity for the detection of breast cancers than ^{18}F -FDG PET. Therefore, PET/MR in breast cancer staging cannot be performed without contrast material in particular because breast lesions and metastatic axillary nodes may be ^{18}F -FDG-negative. Punwani et al. [95] have shown the indispensable role of dynamic contrast-enhanced MRI in Hodgkin lymphoma staging to detect splenic involvement. There are many other examples in which contrast-enhanced sequences are indispensable in the oncologic context, a detailed description being beyond the scope of this article. It therefore appears that contrast-enhanced sequences cannot be excluded from the MRI protocol in most oncologic situations without compromising MRI performance. Nevertheless, standardized PET/MRI protocols and harmonized data acquisition across multiple institutions are desirable. As suggested at the Third International Symposium on PET/MRI [87], an alternative to institutional series would be to create a PET/MRI registry for pooling data from multiple centers. Such a registry would facilitate evaluation of clinical data in terms of diagnostic performance and would equally expedite the evaluation of the impact of PET/MRI on patient management [87].

Finally, the third PET/MRI approach is to perform a total-body, full diagnostic, high-resolution MRI in addition to the total-body PET acquisition. Currently, this option cannot be implemented in clinical settings owing to the unacceptably long in-room time.

Feasibility studies and PET image quality in PET/MRI versus PET/CT

Several studies have shown that PET/MRI is feasible with both simultaneous and sequential systems in patients with a variety of tumors, including lung cancer, breast cancer, brain tumors, head and neck cancers, and pediatric tumors [76]–[78], [80], [88], [96]. A few authors [76], [78], [79], [93], [96] have compared PET/MRI results with PET/CT results, with all patients undergoing a single dose injection of ^{18}F -FDG. Although some authors [78], [79] first performed PET/CT and immediately

thereafter PET/MRI, others chose to perform PET/MRI first and then PET/CT. Experienced observers, who were blinded to clinical data, evaluated the PET/CT and PET/MRI data sets. Despite variable protocols, in all studies, PET/MR image quality, fusion quality, lesion conspicuity, and anatomical lesion localization were good to excellent, and no statistically significant difference was found between the rating scores for image quality, fusion quality, lesion conspicuity, and anatomical localization, as well as with respect to the number of detected focal uptake lesions in PET/MRI and PET/CT, respectively [76], [78], [79], [96].

In terms of quantification, all authors reported a high correlation for SUV values measured in PET/MRI and PET/CT for organs and for malignant and benign focal uptake [76], [78], [79]. However, several investigators have reported that SUVs for focal uptake and normal organs may be underestimated by 11–20% in PET/MRI compared with PET/CT; this observed underestimation results in a limited concordance of SUV measurements between the two modalities [76], [78], [79]. In summary, although quantification issues are not yet completely solved, as discussed in the section on quantitative PET/MRI, several studies have shown that for the detection and localization of FDG-avid lesions and for differentiating between benign and malignant lesions in pediatric and adult patients, PET/MRI results are comparable with PET/CT results. In other words, from the clinical point of view, lesion detection and characterization with PET/MRI do not appear to be significantly affected by limitations in quantitative accuracy. In a clinical setting, the significantly lower radiation exposure when using PET/MRI compared with PET/CT constitutes an important benefit, especially for serial studies and in the pediatric population.

Pitfalls and artifacts

Several artifacts may hamper the interpretation of PET/MRI examinations. They are caused mainly by PET/MRI hardware, MRI and PET physics, physiologic phenomena, the presence

of medical devices, and MRI contrast agents [97]. Artifacts related to the technical aspects and conceptual design of PET/MRI systems can be caused by the presence of RF coils in the FOV and may result in additional attenuation and scatter with more complex patterns. Because the position of RF coils is not visualized on MR images, integrated systems use fixed coils at known positions and do not commonly correct for surface coils (which have a negligible effect), whereas sequential systems use “coil identification scans” to account for the attenuation of RF coils.

Truncation artifacts can be observed in both PET/CT and PET/MRI and typically occur in large patients scanned with arms down. In PET/MRI, they are caused by the fact that the transaxial FOV of the MRI acquisition (~45 cm) is smaller than the FOV of the PET acquisition (~70 cm). Inhomogeneity of the static magnetic field (B_0) and gradient field nonlinearity at the FOV periphery are additional factors predisposing to truncation artifacts. Because parts of the body are outside the FOV of the MRI scan, the resulting attenuation map is incomplete, thereby leading to visible artifacts on the corresponding PET images and underestimation of SUV values. Truncation artifacts can be corrected by obtaining a compensated attenuation map from nonattenuation-corrected emission PET data, which is then fused with the truncated map [98], or through more advanced approaches, such as the MLAA algorithm described in the previous section.

Fold-over artifacts occur along the phase-encoding direction if the chosen FOV of the MRI acquisition is smaller than the part of the body that needs to be imaged. Fold-over artifacts may lead to incorrect PET quantification but can be easily corrected by changing the direction of phase encoding. However, this approach results in a prolonged MRI acquisition time.

Pulsation artifacts occur along the MRI phase-encoding direction and are often seen in the chest, head and neck area, or upper abdomen. They are the consequence of vascular and cardiac pulsation or turbulent flow and may lead to erroneous quantification of tracer uptake mainly in lymph nodes located along vascular structures. Possible solutions to correct pulsation artifacts include changing the phase-encoding direction or applying flow compensation techniques.

Local destructive interferences, eddy currents and standing-wave effects, which are more common at 3 T than at 1.5 T, can lead to signal loss in the area of interest, thereby impairing PET/MRI interpretation. Eddy currents and standing-wave artifacts may occur in the upper abdomen of overweight patients or in patients with peritoneal fluid. Potential solutions include the use of parallel RF coil technology and placing cushions with an ionic solution on the abdomen [99].

A recently described effect observed in simultaneous PET/MRI systems is the shine-through artifact [7]. As reported by Kolb et al. [7], the static magnetic field (B_0) of the MRI scanner affects the trajectory of positrons by reducing the positron range in the plane perpendicular to B_0 (axial plane) and by elongating the positron range along the direction of B_0 (cranio-caudal direction). In PET/MRI, the shine-through artifact

can be seen with low-energy PET radionuclides only if the area of high tracer uptake is in the immediate vicinity of an air cavity, a situation typically observed in the head and neck region. Depending on the orientation of the larynx and trachea in the magnetic field, the magnitude of the artifact may vary considerably. It results in an elongated shape of tracer uptake on coronal and sagittal images and in an apparently increased tracer concentration diametrically opposed to the location of the actual lesion. The shine-through artifact can potentially lead to overestimation of tumor involvement in PET/MRI [7]. Adequate compensation techniques for this artifact are not yet available.

Bulk and respiratory motion-induced mismatch between MRI and PET data acquisition in sequential scanning can result in misregistration. Misregistration can hinder precise tumor localization in small-sized lesions and, in severe cases, overall image interpretation. To avoid this pitfall, anatomical MRI sequences need to be carefully analyzed, and interpretation of PET findings should always take morphology into account. Motion artifacts have been reported mainly in head and neck cancer patients and pediatric oncology patients. Motion artifacts are caused by patient stress from long scanning times, dyspnea, or pain. Careful patient instruction, breaks between sequences, or pain medication, whenever necessary, can significantly reduce the number of poor-quality images caused by motion. Diaphragmatic excursion during respiration may affect interpretation of the basal lung, liver, pancreas, and spleen. PET images in both PET/CT and PET/MR tend to be blurred, the effective resolution above and below the diaphragm is diminished, and tracer uptake may be underestimated. To avoid misalignment, several authors recommended acquiring the MRI attenuation correction sequence and the anatomical reference sequences during shallow free-breathing or in end-expiratory breath hold [97], [100]. Current developments in PET/MRI utilizing motion-sensitive MRI pulse sequences, such as velocity-encoded phase-contrast MRI and tagged MRI, have the potential to outperform PET/CT, for which similar correction strategies do not exist [99], [101].

Although these methods improve PET image quality, they require a modification of the PET/MRI protocol. Recently, Manber et al. [102] showed that a respiratory signal could be extracted from raw PET list-mode data, thus substantially improving clinical PET image quality only by adding an additional 1-minute dynamic MR scan.

A particular problem encountered in pelvic PET/MRI is the misalignment of the bladder owing to the continuous physiologic secretion of urine and resulting bladder volume change during the procedure. This problem is encountered mainly when patients are imaged with sequential PET/MRI systems. Several practical solutions have been proposed, including obtaining an additional fast MRI sequence for image fusion with PET just before starting the PET acquisition, restricting water intake 4 hours before the exam, and asking the patient to void just before beginning the image acquisition [77].

Susceptibility artifacts around ferromagnetic objects, typically dental restorations or osteosynthesis material, result in

distortion of the surrounding anatomy and in bright and dark areas on all sequences, but most frequently on gradient echo sequences, DWI sequences, and fat-saturated sequences. Geometric distortion is a well-recognized problem inherent to DWI sequences; it is caused by B_0 susceptibility differences. Geometric distortion of DWI images is a common cause of misregistration of anatomical and b1000 images. In addition to the effects on image quality, susceptibility artifacts caused by ferromagnetic objects can have an impact on the attenuation coefficient on the MRI-based or CT-based attenuation map. In consequence, artifacts caused by dental implants, hip prostheses, sternal wires, metallic port catheter systems, or other metallic implants can affect SUVs measured in PET/MRI and PET/CT. In PET/MRI, metallic implants result in an apparently decreased focal uptake. A priori knowledge of these pitfalls avoids erroneous image interpretation. Nevertheless, artifacts caused by metallic implants are, in general, larger and more disturbing on CT and PET/CT than on MRI and PET/MRI (Figure 8). Strategies to reduce metal artifacts in PET/MRI are quite challenging, and up until now, this area is still a work in progress. New sequences, such as slice encoding for metal artifact correction with view angle tilting, have been developed for standalone MRI systems; however, they have not yet been tested in hybrid PET/MRI systems [103]. A different approach for metal artifact correction in PET/MRI has been proposed by Ladefoged et al. [104], who developed an automatic algorithm for correction of dental artifacts in PET/MRI by first using a template of artifact regions and then representing the artifactual regions with a combination of active shape models and k -nearest neighbors.

Last but not least, MRI contrast agents, such as iron oxide nanoparticles, which are used mainly to detect focal hepatocellular lesions, may hamper the interpretation of PET/MRI studies. The main effect of iron oxide nanoparticles is on T_2^* relaxation; they cause signal loss in T_2^* -weighted and T_2 -weighted images owing to the susceptibility effects of the iron oxide core. In the liver, the particles accumulate in the Kupffer cells of the normal reticuloendothelial system

while sparing lesions lacking Kupffer cells, such as metastases. Because this effect may last for several weeks, it may affect MRI attenuation maps and should be considered when interpreting PET/MRI examinations, although some studies reported negligible quantification bias because MRI contrast agents have almost the same linear attenuation coefficient as water [105].

Clinical data on diagnostic accuracy of PET/MRI in oncology

Very little data are currently available regarding the diagnostic performance of hybrid PET/MRI systems in oncology. This fact is due to the difficulty in obtaining a rigorous standard of reference based on histology and/or long-term follow-up. In most published studies, PET/CT is used as the standard of reference, whereas one could argue that false-positive and false-negative readings may equally occur with this modality. From a clinical perspective, some incremental progress has been reported over the past two years in breast cancer [81], pediatric oncology [96], prostate cancer [82], head and neck chondrosarcoma [83], and neck irradiation [106], whereas no added value compared with other imaging modalities (PET/CT or MRI) could be demonstrated for lung cancer [107], detection of lung nodules [108], and nodal staging in head and neck squamous cell carcinoma [80].

Taneja et al. [81] assessed the utility of whole-body ^{18}F -FDG PET/MRI in the initial staging of breast cancer in 36 patients with histologically confirmed invasive ductal carcinoma. Primary lesions, lymph nodes, and distant metastases were evaluated with PET, MRI, and PET/MRI for lesion count and diagnostic confidence (DC). The study yielded the highest DC score of 5 with PET/MRI compared with PET (median DC score = 4) and MRI (median DC score = 4) alone. MRI detected 47 satellite lesions, of which only 23 (49%) were FDG-avid with multifocality and multicentricity in 21 (58%) patients. The study equally showed sensitivities of 60 and 93% for PET and MRI, respectively, in the detection of axillary lymph nodes, with a specificity of 91% for both. Combined PET/MRI increased the DC for nodal involvement. Ninety-one metastatic lesions were detected in PET ($\text{DC} \geq 4$) and 105 in MRI ($\text{DC} \geq 4$), with the difference being statistically significant ($p = 0.001$). The authors concluded that PET/MRI is useful as an initial staging modality in breast cancer patients because the DC is higher with PET/MRI compared with PET or MRI alone; however, no statistical comparison was performed to evaluate the added value of PET/MRI compared with MRI alone [81]. In particular, on the basis of the published figures, the authors reported a similar sensitivity and staging accuracy with MRI and PET/MRI [81].

Regarding pediatric oncology, several studies have shown that PET/MRI is technically feasible in children as young as six years old without general anesthesia, as well as in adolescents, showing adequate quantitative accuracy with SUVs compared with those obtained in PET/CT [96]. Schäfer et al. [96] demonstrated that PET/MRI achieved equivalent lesion detection rates compared with PET/CT, with the former offering

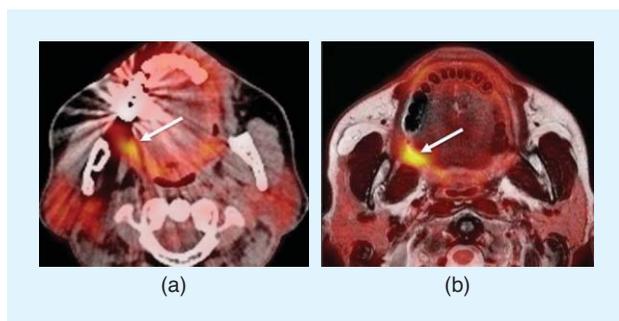


FIGURE 8. Dental artifacts affect PET/MRI less than PET/CT. (a) A PET/CT image shows major streak artifacts from dental implants hampering image interpretation. Note the poorly delineated area of higher FDG uptake on the right (arrow). (b) A corresponding hybrid PET/MRI (PET fused with axial T_2 -weighted image) image obtained from the same patient clearly shows an FDG-avid tumor in the right anterior tonsillar pillar (arrow) extending into the retromolar trigone. The biopsy revealed squamous cell carcinoma.

markedly reduced radiation exposure. Currently, reported effective radiation doses for PET/CT are in the range of 25 mSv, whereas for PET/MRI, they are in the range of 7 mSv [87]. Child-specific patient preparation procedures are, however, mandatory to obtain good-quality PET/MRI examinations. Among whole-body PET/MRI applications identified in pediatric oncology, staging of Hodgkin lymphoma with DWI and dynamic contrast-enhanced MRI sequences has been identified as promising because PET/MRI can enhance the accuracy of lesion detection compared with PET/CT [91]. Data on other pediatric tumors, such as sarcomas and neuroblastomas, are not yet available.

As opposed to the high signal intensity of the peripheral zone, prostate cancers typically display a low signal on T2-weighted images. However, tumors arising from the transitional zone are often difficult to detect owing to the heterogeneous T2 appearance of the transitional zone, with this mixed signal being often caused by benign hyperplasia. Therefore, it has been suggested that a multiparametric imaging approach including T2-weighted sequences, DWI, dynamic contrast-enhanced imaging, and ^{18}F -choline PET may improve pretherapeutic diagnostic accuracy. In a recent publication based on a series of 24 patients with total prostatectomy, de Perrot et al. [82] demonstrated that ^{18}F -choline PET/MRI had an improved diagnostic accuracy in the peripheral zone compared with multiparametric MRI but had no added value in the transition zone owing to adenomatous hyperplasia. ADC and SUV_{max} were not correlated biomarkers, suggesting that they may provide complementary information in the workup of these tumors [82].

Chondrosarcoma of the larynx is a rare, low-grade malignancy; in a minority of cases, a dedifferentiated component can occur within a chondrosarcoma. Histologically, a well-differentiated cartilaginous component is juxtaposed to the dedifferentiated component, with an abrupt transition between the two tissue types [83]. Purohit et al. [83] reported that the diagnosis of dedifferentiation can be suggested in PET/MRI owing to the morphologic and metabolic findings because the well-differentiated component has a low signal on T1-weighted images, slight peripheral enhancement, a high signal on T2-weighted images, high ADC values, and low SUVs, whereas the dedifferentiated component has a low signal on T1-weighted images, major inhomogeneous enhancement, a moderately high signal on T2-weighted images, low ADC values, and high SUVs. The authors therefore concluded that PET/MRI can provide additional functional information to supplement the morphologic mapping and histopathology of these tumors [83].

^{18}F -FDG PET/CT is widely accepted as the evaluation method of choice for staging nonsmall cell lung cancer (NSCLC). Heusch et al. [107] compared a dedicated pulmonary ^{18}F -FDG PET/MRI protocol with PET/CT for primary and locoregional lymph node staging in NSCLC patients using histopathology as the standard of reference. The results from PET/MRI and PET/CT agreed on T stages in all 16 patients (100%). There was no statistically significant difference between PET/CT and PET/

MRI regarding detection of lymph node metastases ($p = 0.48$) and SUV measurements, and tumor size measurements derived from PET/CT and PET/MRI showed a high correlation. The authors concluded that compared with ^{18}F -FDG PET/CT, PET/MRI with a dedicated pulmonary MRI protocol does not provide advantages in thoracic staging in NSCLC patients [107].

Most investigators currently consider multidetector CT (MDCT) the imaging modality of choice for the detection of pulmonary nodules. Although MDCT has high sensitivity in the detection of pulmonary nodules, its capability to differentiate between benign and malignant nodules in patients with primary malignancy is limited [108]. Discrimination between malignant and benign nodules is, however, facilitated by the use of ^{18}F -FDG PET/CT. One of the major potential disadvantages of PET/MRI over PET/CT is the lower sensitivity of MRI compared with CT in the detection of small pulmonary nodules. Chandarana et al. [108] compared the performance of PET, MRI, and combined PET/MRI in the detection of lung nodules in oncologic patients with clinically indicated PET/CT. PET/CT was considered the standard of reference. The combination of PET and MRI acquired using a hybrid PET/MRI system with a radial T1-weighted gradient echo sequence had a higher sensitivity for lung nodules compared with PET or MRI alone. The sensitivities of PET/MRI were 70, 96, and 23% for all nodules together, FDG-avid nodules, and non-FDG-avid nodules, respectively. When nodule size was analyzed, PET/MRI had a sensitivity of 89% for the detection of nodules with a diameter of at least 5 mm and a sensitivity of 38% for the detection of lesions smaller than 5 mm. The authors therefore concluded that PET/MRI has a high sensitivity for FDG-avid lung nodules and for nodules 5 mm or larger in diameter but a lower sensitivity for small non-FDG-avid nodules [108]. Results from our institution [86] confirm these findings and show that although the conspicuity of lung lesions may be less clear in PET/MRI compared with PET/CT, FDG-avid lung nodules and nodules larger than 5 mm are equally well detected with both modalities (Figure 9).

In a prospective study including 38 patients with squamous cell carcinoma of the head and neck, Platzeck et al. [80] evaluated the performance of PET, MRI, and PET/MRI in the detection of lymph node metastases. Results were compared on the basis of receiver operating characteristic analysis, whereas histology served as the standard of reference. Metastatic nodes were present in 42% of the 38 patients and in 10% of the 391 dissected lymph node levels. There were no significant differences among PET/MRI, MRI, and PET ($p > 0.05$) regarding accuracy for cervical metastatic disease. The authors therefore concluded that in head and neck cancer, FDG PET/MRI does not significantly improve accuracy for cervical lymph node metastases compared with MRI or PET [80]. Nevertheless, it is important to mention that MRI did not include DWI acquisitions in this study.

The variable appearance of recurrent tumors after radiation therapy of the head and neck and treatment-induced expected tissue alterations and complications often render MRI or

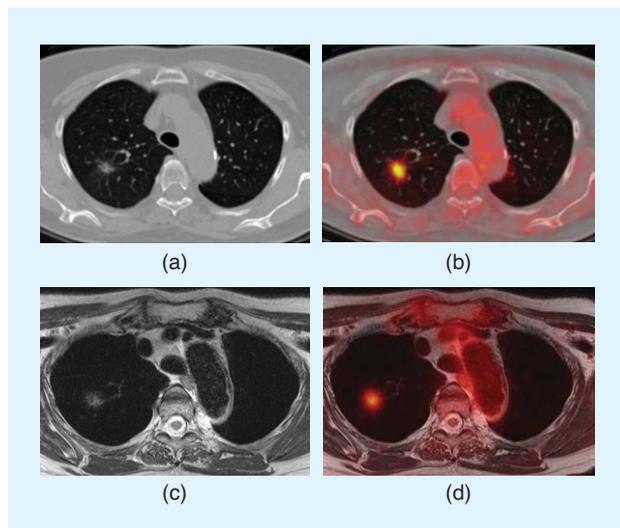


FIGURE 9. The detection of larger FDG-avid pulmonary nodules with PET/MRI and PET/CT. This patient was a follow-up case of a salivary gland adenocarcinoma. (a) and (b) Axial PET/CT and (c) and (d) corresponding PET/MR images obtained within 1 hour and after a single-dose injection of FDG show a spiculated $10 \times 9 \times 9$ -mm large lung nodule with a central cavitory area in the upper right lobe and with focal FDG uptake (SUV_{max} PET/CT = 3.9, SUV_{max} PET/MR = 4.8), suggesting either metastatic disease or a synchronous second primary tumor. A biopsy revealed a second primary pulmonary squamous cell carcinoma.

PET/CT image interpretation very demanding. As reported by Varoquaux et al. [106], PET/MRI with DWI holds promise to facilitate differentiation between tumor recurrence and radiation-induced changes and complications. In particular, multiparametric imaging, including anatomical, contrast-enhanced, DWI, and PET information, can be very beneficial in solving diagnostic dilemmas. In their recent article on DWI/MRI and PET/MRI of the irradiated neck, the authors provided a comprehensive approach to the understanding of key features of radiation-induced edema, fibrosis and scar tissue, soft-tissue necrosis, osteochondronecrosis, brain necrosis, and thyroid disorders by discussing the complementary role of DWI/MRI and PET/MRI in these entities and in the detection of recurrent disease. The authors concluded that multiparametric PET/MRI leads to a major reduction of interpretation pitfalls, thereby increasing the DC in interpreting examinations of the irradiated neck [106].

Diagnostic challenges related to multiparametric imaging

Interpreting hybrid PET/MRI studies with multiparametric data sets can be quite challenging in clinical routine owing to the huge amount of complex information and the difficulty taking all measured parameters into consideration. Current research protocols based on multiparametric data acquisitions already use classification algorithms in the analysis of PET/MRI data, including algorithms based on Gaussian distribution models or support vector machine analysis. Within the same tumor, these classification algorithms can separate tissue regions on the basis of their different PET, ADC, or perfusion maps, therefore enabling more accurate differentiation

between areas of increased proliferation, apoptosis, fibrosis, and viable cells. Differentiating between these entities may have a tremendous impact on future radiotherapy regimens in view of tailored treatment options.

A practical problem that has not yet been solved is how to deal with discrepant multiparametric data from PET/MRI during everyday image interpretation. Should one rely on the morphologic and perfusion information or on the PET or DWI information? How should one weight the value of each parameter to increase the diagnostic yield and avoid unnecessary biopsy? Certainly, the clinical experience of the radiologists and nuclear medicine physicians interpreting the data plays a major role, and the interdisciplinary collaboration with oncologists, surgeons, pathologists, and radiation oncologists for a meaningful integration of all imaging and biological patient data is crucial. Although recent publications on the clinical value of multiparametric PET/MRI show promising results [106], future studies based on larger patient cohorts are required.

Summary and future directions

The bulk of PET/MRI instrumentation research to date has focused on building MR-compatible PET detectors and read-out technologies, reducing the interferences between the two imaging modalities, addressing the challenges of quantitative PET/MRI in general and MRI-guided PET attenuation correction in whole-body imaging in particular, devising tools for advanced multiparametric imaging, and finding a primary clinical role (killer application) for PET/MRI [2], [21]. In this regard, much worthwhile research in instrumentation and quantitative PET/MRI is well under way, and the technical and methodological challenges in this area are likely to be resolved in the near future. While in the clinic, radiologists and nuclear medicine physicians are in search of a primary clinical use of PET/MRI that differentiates it from PET/CT; in doing so, they are making use of tools designed for PET/CT and assessing PET/MRI in the same way as PET/CT.

In summary, the expectations for PET/MRI are high owing to the potential to obtain morphologic, functional and metabolic, qualitative, and quantitative information in the same examination. The overall consensus among active research groups is that multiparametric PET/MRI may add diagnostic certainty in difficult oncologic situations and may also help to tailor treatment plans. Recent research has shown that compared with PET/CT, PET/MRI can demonstrate equivalent lesion detection rates while offering markedly reduced radiation exposure. Although some authors [76], [79], [80], [81] have reported no added diagnostic benefit compared with PET/CT or MRI, others have described added value in selected cases. However, PET/MRI is currently still a long way from providing multiparametric information within an acceptable time window. In addition, there is no consensus regarding which parameters and how many of them are needed to influence relevant clinical endpoints, and decisional algorithms based on multiparametric data still need to be developed. Currently, the interpretation of multiparametric PET/MRI requires a team effort of imaging experts

with different backgrounds. Although this methodological approach is part of ongoing research protocols, it may be more difficult to implement in clinical everyday routine. Last but not least, cost and reimbursement issues are still a matter of debate in many countries.

To aid the implementation of PET/MRI in a clinical environment, future studies will need to address several questions: What is the advantage of PET/MRI compared with separate PET/CT and MRI, in particular, as radiation doses with recent CT scanners continue to decrease substantially? What is the value of each modality (PET, MRI, and PET/MRI) in staging and restaging, personalized treatment decisions, and treatment outcome? What are the unique key applications to PET/MRI? What is the value of multiparametric quantitative analysis tools? Only larger studies based on a solid standard of reference, such as histology and/or long-term follow-up, will be able to answer these questions.

Acknowledgments

This work was supported by the Swiss National Science Foundation under grants SNSF 31003A_149957 and SNSF 320030_135728/1.

Authors

Habib Zaidi (habib.zaidi@hcuge.ch) is the chief physicist and head of the PET Instrumentation and Neuroimaging Laboratory at Geneva University Hospital, faculty member at the medical school of Geneva University, professor of medical physics at the University Medical Center of Groningen, and professor of molecular imaging at the University of Southern Denmark. He was guest editor for nine special issues of peer-reviewed journals and serves on the editorial board of leading journals in medical physics and molecular imaging. He is a Senior Member of the IEEE and liaison representative of the International Organization for Medical Physics to the World Health Organization. He received the 2003 Young Investigator Medical Imaging Science Award, 2004 Mark Tetalman Memorial Award, 2007 Young Scientist Prize in Biological Physics, 2010 Kuwait Prize of Applied Sciences, 2013 John S. Laughlin Young Scientist Award, 2013 Vikram Sarabhai Orion Award, and 2015 Godfrey Hounsfield Award.

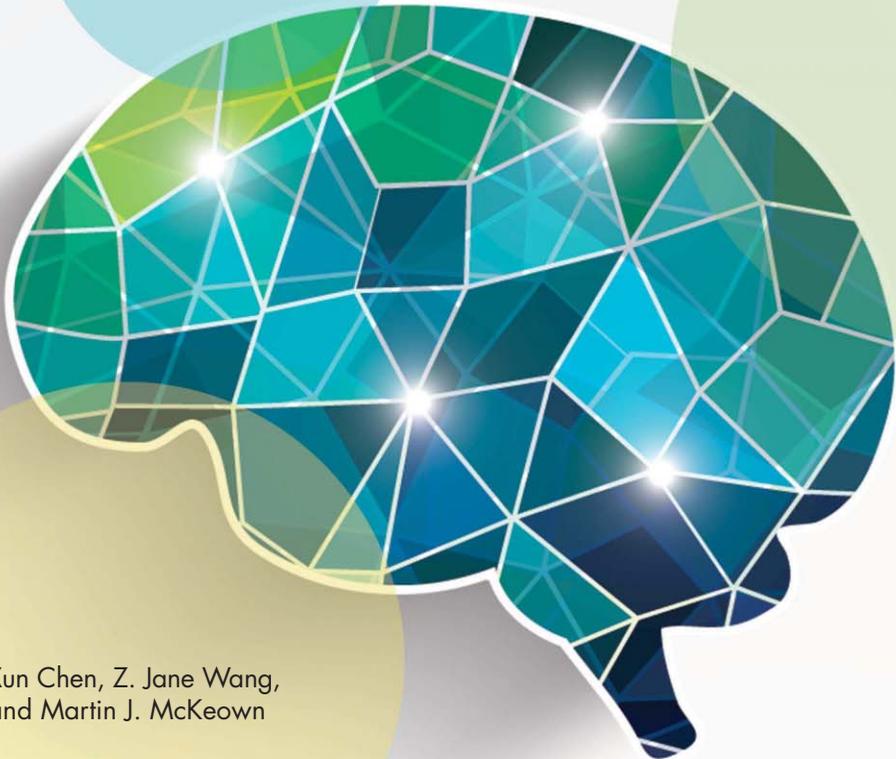
Minerva Becker (minerva.becker@hcuge.ch) studied medicine at the University of Berne in Switzerland and specialized in radiology at the University Hospital in Berne. She is the past president of the European Society of Head and Neck Radiology (ESHNR), author of more than 120 scientific publications and review articles, and the recipient of more than 25 prestigious international and national scientific awards. She is a coeditor of *Valvassori's Imaging of the Head and Neck* and a reviewer for 12 scientific journals. She is currently an associate professor of radiology and the head of the Imaging Unit of Head and Neck and Maxillofacial Radiology at Geneva University Hospital. As chair of the Education Committee of the ESHNR, she is in charge of the European training curriculum in head and neck radiology and European board of head and neck radiology diploma.

References

- [1] P. J. Slomka and R. P. Baum, "Multimodality image registration with software: State-of-the-art," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 36, pp. 44–55, Mar. 2009.
- [2] D. A. Torigian, H. Zaidi, T. C. Kwee, B. Saboury, J. K. Udupa, Z.-H. Cho, and A. Alavi, "PET/MR imaging: Technical aspects and potential clinical applications," *Radiology*, vol. 267, no. 1, pp. 26–44, Apr. 2013.
- [3] H. F. Wehrl, A. W. Sauter, M. R. Divine, and B. J. Pichler, "Combined PET/MR: A technology becomes mature," *J. Nucl. Med.*, vol. 56, no. 2, pp. 165–168, Feb. 2015.
- [4] H. Iida, I. Kanno, S. Miura, M. Murakami, K. Takahashi, and K. Uemura, "A simulation study of a method to reduce positron annihilation spread distribution using a strong magnetic field in positron emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 33, no. 1, pp. 597–600, Feb. 1986.
- [5] R. R. Raylman, B. E. Hammer, and N. L. Christensen, "Combined MRI-PET scanner: A Monte-Carlo evaluation of the improvements in PET resolution due to the effects of a static homogeneous magnetic field," *IEEE Trans. Nucl. Sci.*, vol. 43, no. 4, pp. 2406–2412, Aug. 1996.
- [6] A. Wirrwar, H. Vosberg, H. Herzog, H. Halling, S. Weber, and H.-W. Muller-Gartner, "4.5 Tesla magnetic field reduces range of high-energy positrons-potential implications for positron emission tomography," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 2, pp. 184–189, Apr. 1997.
- [7] A. Kolb, A. W. Sauter, L. A. Eriksson, A. Vandenbrouke, C. C. Liu, C. S. Levin, B. J. Pichler, M. Rafecas, "Shine-through in PET/MRI: Effects of the magnetic field on positron range and subsequent image artifacts," *J. Nucl. Med.*, vol. 56, no. 6, pp. 951–954, June 2015.
- [8] N. L. Christensen, B. E. Hammer, B. G. Heil, and K. Fetterly, "Positron emission tomography within a magnetic field using photomultiplier tubes and light-guides," *Phys. Med. Biol.*, vol. 40, no. 4, pp. 691–697, Apr. 1995.
- [9] Y. Shao, S. R. Cherry, K. Farahani, and K. Meadors, "Simultaneous PET and MR imaging," *Phys. Med. Biol.*, vol. 42, no. 10, pp. 1965–1970, Oct. 1997.
- [10] H. Zaidi and A. Del Guerra, "An outlook on future design of hybrid PET/MRI systems," *Med. Phys.*, vol. 38, no. 10, pp. 5667–5689, Oct. 2011.
- [11] S. Yamamoto, M. Imaizumi, Y. Kanai, M. Tatsumi, M. Aoki, E. Sugiyama, M. Kawakami, E. Shimosegawa, et al., "Design and performance from an integrated PET/MRI system for small animals," *Ann. Nucl. Med.*, vol. 24, no. 2, pp. 89–98, Feb. 2010.
- [12] R. C. Hawkes, T. D. Fryer, S. Siegel, R. E. Ansorge, and T. A. Carpenter, "Preliminary evaluation of a combined microPET-MR system," *Technol. Cancer Res. Treat.*, vol. 9, no. 1, pp. 53–60, Feb. 2010.
- [13] W. B. Handler, K. M. Gilbert, H. Peng, and B. A. Chronik, "Simulation of scattering and attenuation of 511 keV photons in a combined PET/field-cycled MRI system," *Phys. Med. Biol.*, vol. 51, no. 10, pp. 2479–2491, May 2006.
- [14] M. Bergeron, J. Cadorette, J.-F. Beaudoin, M. D. Lepage, G. Robert, V. Selivanov, M.-A. Tetrault, N. Viscogliosi, et al., "Performance evaluation of the LabPET APD-based digital PET scanner," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 1, pp. 10–16, Feb. 2009.
- [15] M. S. Judenhofer, H. F. Wehrl, D. F. Newport, C. Catania, S. B. Siegel, M. Becker, A. Thielscher, M. Kneilling, et al., "Simultaneous PET-MRI: A new approach for functional and morphological imaging," *Nat. Med.*, vol. 14, no. 4, pp. 459–465, Apr. 2008.
- [16] B. Ravindranath, S. Junnarkar, M. Purschke, S. Stoll, X. Hong, P.-J. Huang, S. H. Maramraju, P. Vaska, et al., "Results from a simultaneous PET-MRI breast scanner [abstract]," *J. Nucl. Med.*, vol. 52, p. 432, 2011.
- [17] T. W. Deller, A. M. Grant, M. M. Khalighi, S. H. Maramraju, G. Delso, and C. S. Levin, "PET NEMA performance measurements for a SiPM-based time-of-flight PET/MR system," in *2014 IEEE Nuclear Science Symp. and Medical Imaging Conf. (NSS/MIC)*, Seattle, WA.
- [18] T. Frach, G. Prescher, C. Degenhardt, R. de Gruyter, A. Schmitz, and R. Ballizany, "The digital silicon photomultiplier—Principle of operation and intrinsic detector performance," in *2009 IEEE Nuclear Science Symp. Conf. Record (NSS/MIC)*, Orlando, FL, pp. 1959–1965.
- [19] E. Roncali and S. R. Cherry, "Application of silicon photomultipliers to positron emission tomography," *Ann. Biomed. Eng.*, vol. 39, no. 4, pp. 1358–1377, Apr. 2011.
- [20] S. Moehrs, A. Del Guerra, D. J. Herbert, and M. A. Mandelkern, "A detector head design for small-animal PET with silicon photomultipliers (SiPM)," *Phys. Med. Biol.*, vol. 51, no. 5, pp. 1113–1127, Mar. 2006.
- [21] G. K. von Schulthess, F. P. Kuhn, P. Kaufmann, and P. Veit-Haibach, "Clinical positron emission tomography/magnetic resonance imaging applications," *Semin. Nucl. Med.*, vol. 43, no. 1, pp. 3–10, Jan. 2013.
- [22] P. Vaska and T. Cao, "The state of instrumentation for combined positron emission tomography and magnetic resonance imaging," *Semin. Nucl. Med.*, vol. 43, no. 1, pp. 11–18, Jan. 2013.
- [23] H. Zaidi, "A pivotal time for hybrid PET/MR imaging technology," *J. Am. Coll. Radiol.*, vol. 10, no. 11, pp. 878–881, Nov. 2013.
- [24] H. Zaidi, N. Ojha, M. Morich, J. Griesmer, Z. Hu, P. Maniawski, O. Ratib, D. Izquierdo-Garcia, et al., "Design and performance evaluation of a whole-body

- Ingenuity TF PET-MRI system," *Phys. Med. Biol.*, vol. 56, no. 10, pp. 3091–3106, May 2011.
- [25] P. Veit-Haibach, F. P. Kuhn, F. Wiesinger, G. Delso, and G. von Schulthess, "PET-MR imaging using a tri-modality PET/CT-MR system with a dedicated shuttle in clinical routine," *MAGMA*, vol. 26, no. 1, pp. 25–35, Feb. 2013.
- [26] Z. H. Cho, Y. D. Son, H. K. Kim, K. N. Kim, S. H. Oh, J. Y. Han, I. K. Hong, and Y. B. Kim, "A fusion PET-MRI system with a high-resolution research tomograph-PET and ultra-high field 7.0 T-MRI for the molecular-genetic imaging of the brain," *Proteomics*, vol. 8, no. 6, pp. 1302–1323, Mar. 2008.
- [27] H. P. Schlemmer, B. J. Pichler, M. Schmand, Z. Burbar, C. Michel, R. Ladebeck, K. Jattke, D. Townsend, et al., "Simultaneous MR/PET imaging of the human brain: Feasibility study," *Radiology*, vol. 248, no. 3, pp. 1028–1035, Sept. 2008.
- [28] A. Boss, A. Kolb, M. Hofmann, S. Bisdas, T. Nagele, U. Ernemann, L. Stegger, C. Rossi, et al., "Diffusion tensor imaging in a human PET/MR hybrid system," *Invest. Radiol.*, vol. 45, no. 5, pp. 270–274, May 2010.
- [29] G. Delso, S. Fürst, B. Jakoby, R. Ladebeck, C. Ganter, S. G. Nekolla, M. Schwaiger, and S. I. Ziegler, "Performance measurements of the Siemens mMR integrated whole-body PET/MR scanner," *J. Nucl. Med.*, vol. 52, no. 12, pp. 1914–1922, Dec. 2011.
- [30] A. Kolb, H. F. Wehrl, M. Hofmann, M. S. Judenhofer, L. Eriksson, R. Ladebeck, M. P. Lichy, L. Byars, et al., "Technical performance evaluation of a human brain PET/MRI system," *Eur. Radiol.*, vol. 22, no. 8, pp. 1776–1788, Aug. 2012.
- [31] A. Martinez-Möller, M. Eiber, S. G. Nekolla, M. Souvatzoglou, A. Drzezga, S. Ziegler, E. J. Rummeny, M. Schwaiger, et al., "Workflow and scan protocol considerations for integrated whole-body PET/MRI in oncology," *J. Nucl. Med.*, vol. 53, no. 9, pp. 1415–1426, Sept. 2012.
- [32] A. Mehranian and H. Zaidi, "Impact of time-of-flight PET on quantification errors in MRI-based attenuation correction," *J. Nucl. Med.*, vol. 56, no. 4, pp. 635–641, Apr. 2015.
- [33] D. R. Schaart, S. Seifert, R. Vinke, H. T. van Dam, P. Dendooven, H. Löhner, and F. J. Beekman, "LaBr(3):Ce and SiPMs for time-of-flight PET: Achieving 100 ps coincidence resolving time," *Phys. Med. Biol.*, vol. 55, no. 7, pp. N179–N189, Apr. 2010.
- [34] M. V. Nemallapudi, S. Gundacker, P. Lecoq, E. Auffray, A. Ferri, A. Gola, and C. Piemonte, "Sub-100 ps coincidence time resolution for positron emission tomography with LSO:Ce codoped with Ca," *Phys. Med. Biol.*, vol. 60, no. 12, pp. 4635–4649, June 2015.
- [35] M. S. Judenhofer and S. R. Cherry, "Applications for preclinical PET/MRI," *Semin. Nucl. Med.*, vol. 43, no. 1, pp. 19–29, Jan. 2013.
- [36] H. Zaidi, "Is MRI-guided attenuation correction a viable option for dual-modality PET/MR imaging?," *Radiology*, vol. 244, no. 3, pp. 639–642, Sept. 2007.
- [37] A. Mehranian, H. Arabi, and H. Zaidi, "Magnetic resonance imaging-guided attenuation correction in PET/MRI: Challenges, solutions and opportunities," *Med. Phys.*, vol. 43, no. 3, pp. 1130–1155, Mar. 2016.
- [38] Y. Berker, J. Franke, A. Salomon, M. Palmowski, H. C. Donker, Y. Temur, F. M. Mottaghy, C. Kuhl, et al., "MRI-based attenuation correction for hybrid PET/MRI systems: A 4-class tissue segmentation technique using a combined ultra-short-echo-time/Dixon MRI sequence," *J. Nucl. Med.*, vol. 53, no. 5, pp. 796–804, May 2012.
- [39] G. Delso, F. Wiesinger, L. Sacolick, S. Kaushik, D. Shanbhag, M. Hüllner, and P. Veit-Haibach, "Clinical evaluation of zero echo time MRI for the segmentation of the skull," *J. Nucl. Med.*, vol. 56, no. 3, pp. 417–422, Mar. 2015.
- [40] A. Martinez-Möller, M. Souvatzoglou, G. Delso, R. A. Bundschuh, C. Chefd'hotel, S. I. Ziegler, N. Navab, M. Schwaiger, et al., "Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: Evaluation with PET/CT data," *J. Nucl. Med.*, vol. 50, no. 4, pp. 520–526, Apr. 2009.
- [41] J. Ouyang, S. Y. Chun, Y. Petibon, A. A. Bonab, N. Alpert, and G. El Fakhri, "Bias atlases for segmentation-based PET attenuation correction using PET-CT and MR," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 5, pp. 3373–3382, Oct. 2013.
- [42] A. Akbarzadeh, M. R. Ay, A. Ahmadian, N. Riahi Alam, and H. Zaidi, "MRI-guided attenuation correction in whole-body PET/MR: Assessment of the effect of bone attenuation," *Ann. Nucl. Med.*, vol. 27, no. 2, pp. 152–162, Feb. 2013.
- [43] M. C. Aznar, R. Sersar, J. Saabye, C. N. Ladefoged, F. L. Andersen, J. H. Rasmussen, J. Löfgren, and T. Beyer, "Whole-body PET/MRI: The effect of bone attenuation during MR-based attenuation correction in oncology imaging," *Eur. J. Radiol.*, vol. 83, no. 7, pp. 1177–1183, July 2014.
- [44] H. Arabi, O. Rager, A. Alem, A. Varoquaux, M. Becker, and H. Zaidi, "Clinical assessment of MR-guided 3-class and 4-class attenuation correction in PET/MR," *Mol. Imaging Biol.*, vol. 17, no. 2, pp. 264–276, Apr. 2015.
- [45] D. H. Paulus, H. H. Quick, C. Geppert, M. Fenchel, Y. Zhan, G. Hermosillo, D. Faul, F. Boada, et al., "Whole-body PET/MR imaging: Quantitative evaluation of a novel model-based MR attenuation correction method including bone," *J. Nucl. Med.*, vol. 56, no. 7, pp. 1061–1066, July 2015.
- [46] M. Hofmann, I. Bezrukov, F. Mantlik, P. Aschoff, F. Steinke, T. Beyer, B. J. Pichler, and B. Scholköpf, "MRI-based attenuation correction for whole-body PET/MRI: Quantitative evaluation of segmentation- and atlas-based methods," *J. Nucl. Med.*, vol. 52, no. 9, pp. 1392–1399, Sept. 2011.
- [47] E. Schreibmann, J. A. Nye, D. M. Schuster, D. R. Martin, J. Votaw, and T. Fox, "MR-based attenuation correction for hybrid PET-MR brain imaging systems using deformable image registration," *Med. Phys.*, vol. 37, no. 5, pp. 2101–2109, May 2010.
- [48] N. Burgos, M. J. Cardoso, K. Thielemans, M. Modat, S. Pedemonte, J. Dickson, A. Barnes, R. Ahmed, et al., "Attenuation correction synthesis for hybrid PET-MR scanners: Application to brain studies," *IEEE Trans. Med. Imaging*, vol. 33, no. 12, pp. 2332–2341, Dec. 2014.
- [49] M. R. Juttukonda, B. G. Mersereau, Y. Chen, Y. Su, B. G. Rubin, T. L. Benzinger, D. S. Lalush, and H. An, "MR-based attenuation correction for PET/MRI neurological studies with continuous-valued attenuation coefficients for bone through a conversion from R2* to CT-Hounsfield units," *NeuroImage*, vol. 112, pp. 160–168, May 2015.
- [50] H. Arabi and H. Zaidi, "Fast atlas-based MRI-guided PET attenuation map generation in whole-body PET/MR imaging," in *2015 IEEE Nuclear Science Symp. and Medical Imaging Conf. (NSS/MIC)*, San Diego, CA, 2015.
- [51] A. Rezaei, M. Defrise, G. Bal, C. Michel, M. Conti, C. Watson, and J. Nuyts, "Simultaneous reconstruction of activity and attenuation in time-of-flight PET," *IEEE Trans. Med. Imaging*, vol. 31, no. 12, pp. 2224–2233, Dec. 2012.
- [52] A. Salomon, A. Goedicke, B. Schweizer, T. Aach, and V. Schulz, "Simultaneous reconstruction of activity and attenuation for PET/MR," *IEEE Trans. Med. Imaging*, vol. 30, no. 3, pp. 804–813, Mar. 2011.
- [53] A. Mehranian and H. Zaidi, "Joint estimation of activity and attenuation in whole-body TOF PET/MRI using constrained Gaussian mixture models," *IEEE Trans. Med. Imaging*, vol. 34, no. 9, pp. 1808–1821, Sept. 2015.
- [54] A. Mehranian and H. Zaidi, "Clinical assessment of emission- and segmentation-based MRI-guided attenuation correction in whole body TOF PET/MRI," *J. Nucl. Med.*, vol. 56, no. 6, pp. 877–883, June 2015.
- [55] A. J. Reader and H. Zaidi, "Advances in PET image reconstruction," *PET Clinics*, vol. 2, no. 2, pp. 173–190, Apr. 2007.
- [56] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imaging*, vol. 9, no. 1, pp. 84–93, Mar. 1990.
- [57] B. Bai, Q. Li, and R. M. Leahy, "Magnetic resonance-guided positron emission tomography image reconstruction," *Semin. Nucl. Med.*, vol. 43, no. 1, pp. 30–44, Jan. 2013.
- [58] K. Baete, J. Nuyts, W. Van Paesschen, P. Suetens, and P. Dupont, "Anatomical-based FDG-PET reconstruction for the detection of hypo-metabolic regions in epilepsy," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 510–519, Apr. 2004.
- [59] J. Tang and A. Rahmim, "Bayesian PET image reconstruction incorporating anato-functional joint entropy," *Phys. Med. Biol.*, vol. 54, no. 23, pp. 7063–7075, Dec. 2009.
- [60] K. Erlandsson, I. Buvat, P. H. Pretorius, B. A. Thomas, and B. F. Hutton, "A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology," *Phys. Med. Biol.*, vol. 57, no. 21, pp. R119–R159, Nov. 2012.
- [61] M. Quarantelli, K. Berkouk, A. Prinster, B. Landeau, C. Svarer, L. Balkay, B. Alfano, A. Brunetti, et al., "Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction," *J. Nucl. Med.*, vol. 45, no. 2, pp. 192–201, Feb. 2004.
- [62] H. Matsuda, T. Ohnishi, T. Asada, Z. J. Li, H. Kanetaka, E. Imabayashi, F. Tanaka, and S. Nakano, "Correction for partial-volume effects on brain perfusion SPECT in healthy men," *J. Nucl. Med.*, vol. 44, no. 8, pp. 1243–1252, Aug. 2003.
- [63] H. Zaidi, T. Ruest, F. Schoenahl, and M.-L. Montandon, "Comparative evaluation of statistical brain MR image segmentation algorithms and their impact on partial volume effect correction in PET," *NeuroImage*, vol. 32, no. 4, pp. 1591–1607, Oct. 2006.
- [64] V. Frouin, C. Comtat, A. Reilhac, and M.-C. Grégoire, "Correction of partial volume effect for PET striatal imaging: Fast implementation and study of robustness," *J. Nucl. Med.*, vol. 43, no. 12, pp. 1715–1726, Dec. 2002.
- [65] M. Shidahara, C. Tsoumpas, A. Hammers, N. Bousson, D. Visvikis, T. Suhara, I. Kanno, and F. E. Turkheimer, "Functional and structural synergy for resolution recovery and partial volume correction in brain PET," *NeuroImage*, vol. 44, no. 2, pp. 340–348, Jan. 2009.
- [66] C. Catana, "Motion correction options in PET/MRI," *Semin. Nucl. Med.*, vol. 45, no. 3, pp. 212–223, May 2015.
- [67] A. J. van der Kouwe, T. Benner, and A. M. Dale, "Real-time rigid body motion correction and shimming using cloverleaf navigators," *Magn. Reson. Med.*, vol. 56, no. 5, pp. 1019–1032, Nov. 2006.
- [68] C. Catana, T. Benner, A. van der Kouwe, L. Byars, M. Hamm, D. B. Chonde, C. J. Michel, G. El Fakhri, et al., "MRI-assisted PET motion correction for neurologic studies in an integrated MR-PET scanner," *J. Nucl. Med.*, vol. 52, no. 1, pp. 154–161, Jan. 2011.
- [69] C. Tsoumpas, J. E. Mackewn, P. Halsted, A. P. King, C. Buerger, J. J. Totman, T. Schaeffter, and P. K. Marsden, "Simultaneous PET-MR acquisition and

- MR-derived motion fields for correction of non-rigid motion in PET," *Ann. Nucl. Med.*, vol. 24, no. 10, pp. 745–750, Dec. 2010.
- [70] P. Kellman, C. Chef'd'hotel, C. H. Lorenz, C. Mancini, A. E. Arai, and E. R. McVeigh, "Fully automatic, retrospective enhancement of real-time acquired cardiac cine MR images using image-based navigators and respiratory motion-corrected averaging," *Magn. Reson. Med.*, vol. 59, no. 4, pp. 771–778, Apr. 2008.
- [71] A. H. Davarpanah, Y. P. Chen, A. Kino, C. T. Farrelly, A. N. Keeling, J. J. Sheehan, A. B. Ragin, P. J. Weale, et al., "Accelerated two- and three-dimensional cine MR imaging of the heart by using a 32-channel coil," *Radiology*, vol. 254, no. 1, pp. 98–108, Jan. 2010.
- [72] B. Guerin, S. Cho, S. Y. Chun, X. Zhu, N. M. Alpert, G. El Fakhri, T. Reese, and C. Catana, "Nonrigid PET motion compensation in the lower abdomen using simultaneous tagged-MRI and PET imaging," *Med. Phys.*, vol. 38, no. 6, pp. 3025–3038, June 2011.
- [73] S. Y. Chun, T. G. Reese, J. Ouyang, B. Guerin, C. Catana, X. Zhu, N. M. Alpert, and G. El Fakhri, "MRI-based nonrigid motion correction in simultaneous PET/MRI," *J. Nucl. Med.*, vol. 53, no. 8, pp. 1284–1291, Aug. 2012.
- [74] A. P. King, C. Buerger, C. Tsoumpas, P. K. Marsden, and T. Schaeffter, "Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator," *Med. Image Anal.*, vol. 16, no. 1, pp. 252–264, Jan. 2012.
- [75] H. Fayad, H. Schmidt, C. Wuerslin, and D. Visvikis, "Reconstruction-incorporated respiratory motion correction in clinical simultaneous PET/MR imaging for oncology applications," *J. Nucl. Med.*, vol. 56, no. 6, pp. 884–889, June 2015.
- [76] A. Varoquaux, O. Rager, A. Poncet, B. M. Delattre, O. Ratib, C. D. Becker, P. Dulguerov, N. Dulguerov, et al., "Detection and quantification of focal uptake in head and neck tumours: 18F-FDG PET/MR versus PET/CT," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 41, no. 3, pp. 462–475, Mar. 2014.
- [77] M. I. Vargas, M. Becker, V. Garibotto, S. Heinzer, P. Loubeyre, J. Gariani, K. Lovblad, J. P. Vallée, et al., "Approaches for the optimization of MR protocols in clinical hybrid PET/MRI studies," *MAGMA*, vol. 26, no. 1, pp. 57–69, Feb. 2013.
- [78] A. Drzegza, M. Souvatzoglou, M. Eiber, A. Beer, S. Ziegler, S. Furst, A. Martinez-Möller, S. G. Nekolla, et al., "First clinical experience with integrated whole-body MR/PET. Comparison to PET/CT in patients with oncological diagnoses," *J. Nucl. Med.*, vol. 53, no. 6, pp. 845–855, June 2012.
- [79] K. Kubiessa, S. Purz, M. Gawlitza, A. Kuhn, J. Fuchs, K. G. Steinhoff, A. Boehm, O. Sabri, et al., "Initial clinical results of simultaneous 18F-FDG PET/MRI in comparison to 18F-FDG PET/CT in patients with head and neck cancer," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 41, no. 4, pp. 639–648, Apr. 2014.
- [80] I. Platzek, B. Beuthien-Baumann, M. Schneider, V. Gudziol, H. H. Kitzler, J. Maus, G. Schramm, M. Popp, et al., "FDG PET/MR for lymph node staging in head and neck cancer," *Eur. J. Radiol.*, vol. 83, no. 7, pp. 1163–1168, July 2014.
- [81] S. Taneja, A. Jena, R. Goel, R. Sarin, and S. Kaul, "Simultaneous whole-body 18F-FDG PET-MRI in primary staging of breast cancer: A pilot study," *Eur. J. Radiol.*, vol. 83, no. 12, pp. 2231–2239, Dec. 2014.
- [82] T. de Perrot, O. Rager, M. Scheffler, M. Lord, M. Pusztaszeri, C. Iselin, O. Ratib, and J. P. Vallée, et al., "Potential of hybrid F-fluorocholine PET/MRI for prostate cancer imaging," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 41, no. 9, pp. 1744–1755, Sept. 2014.
- [83] B. S. Purohit, P. Dulguerov, K. Burkhardt, and M. Becker, "Dedifferentiated laryngeal chondrosarcoma: Combined morphologic and functional imaging with positron-emission tomography/magnetic resonance imaging," *Laryngoscope*, vol. 124, no. 7, pp. E274–E277, July 2014.
- [84] K. J. Beiderwellen, T. D. Poeppel, V. Hartung-Knemeyer, C. Buchbender, H. Kuehl, A. Bockisch, and T. C. Lauenstein, "Simultaneous ⁶⁸Ga-DOTATOC PET/MRI in patients with gastroenteropancreatic neuroendocrine tumors: Initial results," *Invest. Radiol.*, vol. 48, no. 5, pp. 273–279, May 2013.
- [85] B. S. Purohit, A. Ailianou, N. Dulguerov, C. D. Becker, O. Ratib, and M. Becker, "FDG-PET/CT pitfalls in oncological head and neck imaging," *Insights Imaging*, vol. 5, no. 5, pp. 585–602, Oct. 2014.
- [86] M. Becker and H. Zaidi, "Imaging in head and neck squamous cell carcinoma: The potential role of PET/MRI," *Br. J. Radiol.*, vol. 87, no. 1036, 20130677, Apr. 2014.
- [87] D. L. Bailey, G. Antoch, P. Bartenstein, H. Barthel, A. J. Beer, S. Bisdas, D. A. Bluemke, R. Boellaard, et al., "Combined PET/MR: The real work has just started. Summary report of the third international workshop on PET/MR imaging; February 17–21, 2014, Tübingen, Germany," *Mol. Imaging Biol.* vol. 17, no. 3, pp. 297–312, June 2015.
- [88] S. Purz, O. Sabri, A. Viehweger, H. Barthel, R. Kluge, I. Sorge, and F. W. Hirsch, "Potential pediatric applications of PET/MR," *J. Nucl. Med.*, vol. 55, pp. 32S–39S, Apr. 2014.
- [89] J. Gu, T. Chan, J. Zhang, A. Y. Leung, Y. L. Kwong, and P. L. Khong, "Whole-body diffusion-weighted imaging: The added value to whole-body MRI at initial diagnosis of lymphoma," *AJR Am. J. Roentgenol.*, vol. 197, no. 3, pp. W384–W391, Sept. 2011.
- [90] C. Lin, A. Luciani, E. Itti, C. Haioun, V. Safar, M. Meignan, and A. Rahmouni, "Whole-body diffusion magnetic resonance imaging in the assessment of lymphoma," *Cancer Imaging*, vol. 12, no. 2, pp. 403–408, Sept. 2012.
- [91] S. Punwani, S. A. Taylor, Z. Z. Saad, A. Bainbridge, A. Groves, S. Daw, A. Shankar, S. Halligan, et al., "Diffusion-weighted MRI of lymphoma: Prognostic utility and implications for PET/MRI?," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 40, no. 3, pp. 373–385, Feb. 2013.
- [92] A. Varoquaux, O. Rager, K. O. Lovblad, K. Masterson, P. Dulguerov, O. Ratib, C. D. Becker, and M. Becker, "Functional imaging of head and neck squamous cell carcinoma with diffusion-weighted MRI and FDG PET/CT: Quantitative analysis of ADC and SUV," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 40, no. 6, pp. 842–852, June 2013.
- [93] F. P. Kuhn, M. Hüllner, C. E. Mader, N. Kastrinidis, G. F. Huber, G. K. von Schulthess, S. Kollias, and P. Veit-Haibach, "Contrast-enhanced PET/MR imaging versus contrast-enhanced PET/CT in head and neck cancer: How much MR information is needed?," *J. Nucl. Med.*, vol. 55, no. 4, pp. 551–558, Apr. 2014.
- [94] M. Comoretto, L. Balestreri, E. Borsatti, M. Cimitan, G. Franchin, and M. Lise, "Detection and restaging of residual and/or recurrent nasopharyngeal carcinoma after chemotherapy and radiation therapy: Comparison of MR imaging and FDG PET/CT," *Radiology*, vol. 249, no. 1, pp. 203–211, Oct. 2008.
- [95] S. Punwani, K. K. Cheung, N. Skipper, N. Bell, A. Bainbridge, S. A. Taylor, A. M. Groves, S. F. Hain, et al., "Dynamic contrast-enhanced MRI improves accuracy for detecting focal splenic involvement in children and adolescents with Hodgkin disease," *Pediatr. Radiol.*, vol. 43, no. 8, pp. 941–949, Aug. 2013.
- [96] J. F. Schäfer, S. Gatidis, H. Schmidt, B. Guckel, I. Bezrukov, C. A. Pfannenberger, M. Reimold, M. Ebinger, et al., "Simultaneous whole-body PET/MR imaging in comparison to PET/CT in pediatric oncology: Initial results," *Radiology*, vol. 273, no. 1, pp. 220–231, Oct. 2014.
- [97] C. Martinez-Rios, R. F. Muzic, Jr., F. P. DiFilippo, L. Hu, C. Rubbert, and K. A. Herrmann, "Artifacts and diagnostic pitfalls in positron emission tomography-magnetic resonance imaging," *Semin. Roentgenol.*, vol. 49, no. 3, pp. 255–270, June 2014.
- [98] G. Delso, A. Martinez-Möller, R. A. Bundschuh, S. G. Nekolla, and S. I. Ziegler, "The effect of limited MR field of view in MR/PET attenuation correction," *Med. Phys.*, vol. 37, no. 6, pp. 2804–2812, June 2010.
- [99] U. Attenberger, C. Catana, H. Chandarana, O. A. Catalano, K. Friedman, S. A. Schonberg, J. Thrall, M. Salvatore, et al., "Whole-body FDG PET-MR oncologic imaging: Pitfalls in clinical interpretation related to inaccurate MR-based attenuation correction," *Abdom. Imaging*, vol. 40, no. 6, pp. 1374–1386, Aug. 2015.
- [100] C. B. Brendle, H. Schmidt, S. Fleischer, U. H. Braeuning, C. A. Pfannenberger, and N. F. Schwenzer, "Simultaneously acquired MR/PET images compared with sequential MR/PET and PET/CT: Alignment quality," *Radiology*, vol. 268, no. 1, pp. 190–199, July 2013.
- [101] N. F. Osman and J. L. Prince, "Regenerating MR tagged images using harmonic phase (HARP) methods," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 8, pp. 1428–1433, Aug. 2004.
- [102] R. Manber, K. Thielemans, B. F. Hutton, A. Barnes, S. Ourselin, S. Arridge, C. O'Meara, S. Wan, et al., "Practical PET respiratory motion correction in clinical PET/MR," *J. Nucl. Med.*, vol. 56, no. 6, pp. 890–896, June 2015.
- [103] R. Sutter, E. J. Ulbrich, V. Jellus, M. Nittka, and C. W. Pfirrmann, "Reduction of metal artifacts in patients with total hip arthroplasty with slice-encoding metal artifact correction and view-angle tilting MR imaging," *Radiology*, vol. 265, no. 1, pp. 204–214, Oct. 2012.
- [104] C. N. Ladefoged, F. L. Andersen, S. H. Keller, T. Beyer, I. Law, L. Højgaard, S. Darkner, and F. Lauze, "Automatic correction of dental artifacts in PET/MRI," *J. Med. Imaging (Bellingham)*, vol. 2, no. 2, 024009, Apr. 2015.
- [105] C. Lois, I. Bezrukov, H. Schmidt, N. Schwenzer, M. Werner, J. Kupferschläger, and T. Beyer, "Effect of MR contrast agents on quantitative accuracy of PET in combined whole-body PET/MR imaging," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 39, no. 11, pp. 1756–1766, Nov. 2012.
- [106] A. Varoquaux, O. Rager, P. Dulguerov, K. Burkhardt, A. Ailianou, and M. Becker, "Diffusion-weighted and PET/MR imaging after radiation therapy for malignant head and neck tumors," *RadioGraphics*, vol. 35, no. 5, pp. 1502–1527, Sept.–Oct. 2015.
- [107] P. Heusch, C. Sproll, C. Buchbender, E. Rieser, J. Terjung, C. Antke, I. Boeck, S. Macht, et al., "Diagnostic accuracy of ultrasound, ¹⁸F-FDG-PET/CT, and fused ¹⁸F-FDG-PET-MR images with DWI for the detection of cervical lymph node metastases of HNSCC," *Clin. Oral Investig.*, vol. 18, no. 3, pp. 969–978, Apr. 2014.
- [108] H. Chandarana, L. Heacock, R. Rakheja, L. R. Demello, J. Bonavita, T. K. Block, C. Geppert, J. S. Babb, et al., "Pulmonary nodules in patients with primary malignancy: Comparison of hybrid PET/MR and PET/CT imaging," *Radiology*, vol. 268, no. 3, pp. 874–881, Sept. 2013.



Xun Chen, Z. Jane Wang,
and Martin J. McKeown

Joint Blind Source Separation for Neurophysiological Data Analysis

Multiset and multimodal methods

Conventional blind source separation (BSS) methods have become widely adopted tools for neurophysiological data analysis. However, the increasing availability of multiset and multimodal neurophysiological data has posed new challenges for BSS methods originally designed to analyze one data set at a time. Concomitantly, there is growing recognition that joint analysis of neurophysiological data has the potential to substantially enhance our understanding of brain function by extracting information from

Digital Object Identifier 10.1109/MSP.2016.2521870
Date of publication: 27 April 2016

complementary modalities and synergistically combining the results. Therefore, joint data analysis methods represent both a challenge and an opportunity for the neurophysiological signal processing community that attempts to enhance understanding of normal brain function and the pathophysiology of many brain diseases. Over the past decade, various joint blind source separation (JBSS) methods have been proposed to simultaneously accommodate multiple data sets. In this article, we provide an overview and taxonomy of representative JBSS methods. We show, through illustrative numerical simulations, that different statistical assumptions and tradeoffs underlie different JBSS methods, affecting which method should be ideally chosen for a given application. We then discuss several real-world neurophysiological applications from both multiset and multimodal perspectives, highlighting the benefits of the JBSS methods as effective and promising tools for neurophysiological data analysis. Finally, we discuss remaining challenges for future JBSS development.

Introduction

Studying the electrical or metabolic activity of the brain in action—i.e., neurophysiological data analysis—is vital for studying human brain function, as it allows the investigation of how the brain is able to perceive, process, and act upon information. Traditionally, neuroscience has been considered “data rich, theory poor,” and substantial effort has been placed in finding ways to model and otherwise summarize the vast amounts of data generated by newer and more accurate brain monitoring technologies. One standard approach is approximation—for example, in an electroencephalogram (EEG), the instantaneous electric potential distribution across many electrodes on the scalp can be approximated by a relatively small number of dipoles judiciously localized within the head. Another approach is factorization of the data into individual components that can then be individually examined. Such methods include principal component analysis (PCA), nonnegative matrix factorization (NMF), and independent component analysis (ICA). ICA is typically considered a BSS approach, as underlying sources are recovered from observed signals with minimal or no prior information about the sources or the mixing process. ICA has been widely adopted for analyzing neurophysiological data [1]. For instance, McKeown et al. applied ICA to functional magnetic resonance imaging (fMRI) data and derived one independent spatial component whose time course closely matched the alternation pattern between experimental and control tasks [2]. Jung et al. employed ICA to decompose EEG data and effectively removed contamination from a wide variety of artifactual sources in EEG recordings [3]. Although it has been shown in many studies that these conventional BSS approaches are powerful, their inability to simultaneously model multiple data sets has limited their broader usage, especially within

situations where several types of simultaneously recorded neurophysiological data are available.

The emergence of new brain imaging techniques has provided researchers with new opportunities to probe brain function from different perspectives. Several new techniques have been developed (Figure 1) between 1988 and 2014, filling in the gaps in the spatiotemporal plane between existing techniques. Perhaps more importantly, collecting multiple types

of neurophysiological data from the same individual has become realistic. Concomitantly, there has been increased recognition that information from a single brain mapping modality provides an unacceptably incomplete picture of brain functioning (for a review, see [4]). As has recently been surmised, many current studies in neuroscience are performed by individual labs and their immediate collaborators, and thus studies

tend to be restricted to one modality or one place in the spatiotemporal plane of technologies used to probe brain function [5]. An example of the limitations of such an approach can be seen in the field of neurodegenerative disease—for example, in the study of Alzheimer’s disease (AD) or Parkinson’s disease (PD). By the time a person becomes clinically symptomatic with such a disease, the disease may already be too advanced; thus, there is a strong desire to identify early biomarkers. Yet finding a single biomarker for different neurodegenerative diseases has proven to be elusive, leading to the realization that biomarkers combining results from different modalities will be necessary [75].

In recent years, joint JBSS methods have emerged to meaningfully integrate data from different brain monitoring modalities [14], [23]. Here, the term *joint* means that multiple ensembles of underlying sources are simultaneously recovered from multiple sets of observed signals based on assumptions about dependencies among the data sets. The crucial difference between BSS and JBSS is reflected by the fact that BSS only examines each data set individually, while JBSS generalizes BSS to include similar dependence relations across multiple data sets [16]. The concept of JBSS in biomedical engineering was comprehensively proposed by Adali and Calhoun et al. [17], [22], [23], [44], though a few ad hoc methods had already been developed before and applied to clinical applications.

It is anticipated that synergistically combining neuroimaging analyses will provide an unprecedented opportunity for understanding normal brain function and the pathophysiology of many brain diseases. Since each brain imaging modality is an indirect reflection of underlying neural activity at a specific spatiotemporal scale, each single modality provides a different aspect of brain function [6]. For instance, fMRI measures the hemodynamic response related to neural activity with fairly good spatial resolution but relatively poor temporal resolution. Conversely, EEG measures brain electrical activity with excellent temporal resolution but low spatial resolution. A more precise spatiotemporal study of human brain function would incorporate integration of both EEG

The emergence of new brain imaging techniques has provided researchers with new opportunities to probe brain function from different perspectives.

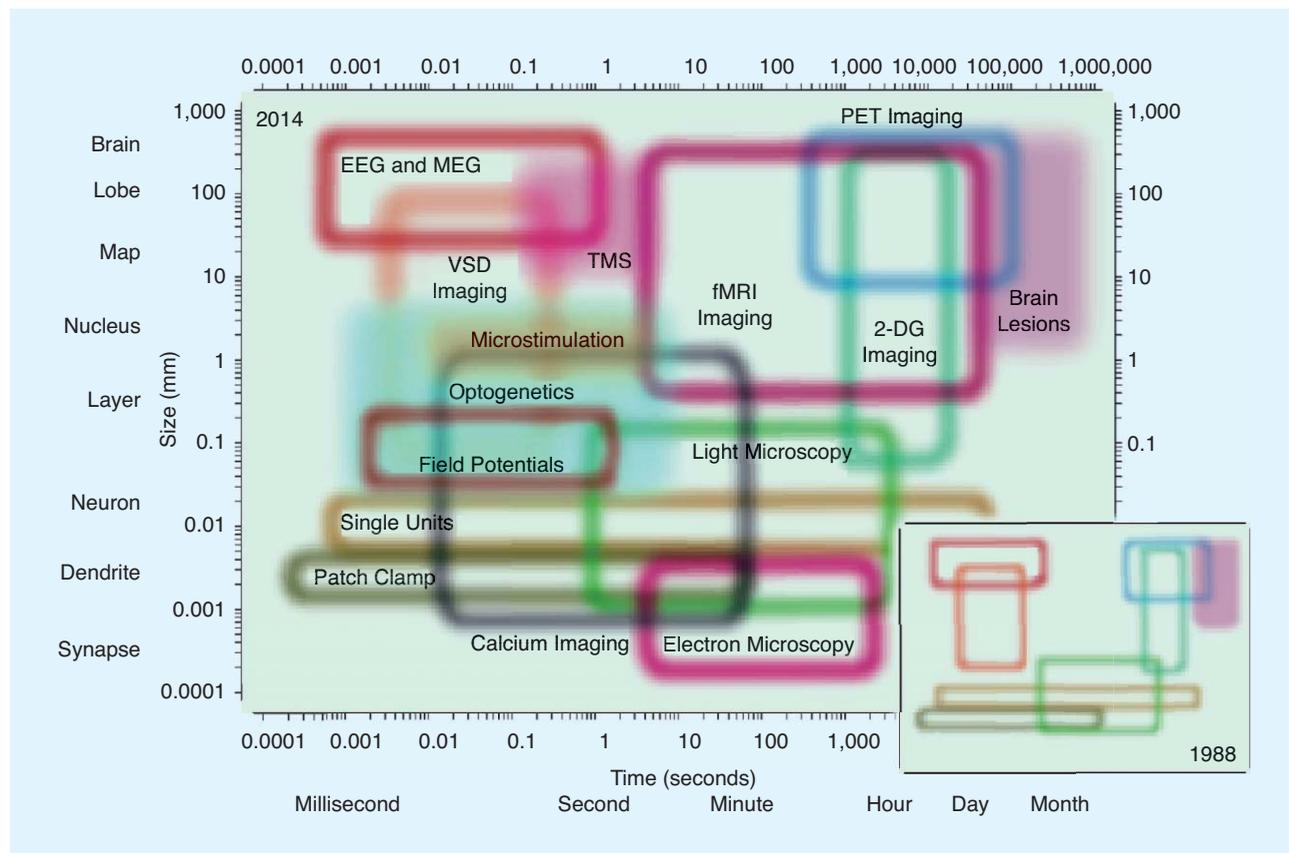


FIGURE 1. The spatiotemporal domain of neuroscience and the main methods available for the study of the nervous system in 2014. Each colored region represents the spatial and temporal resolution range for one specific method available for the study of brain function. Open regions represent measurement techniques; filled regions represent perturbation techniques. The inset shows a cartoon rendition of the methods available in 1988, which is notable for large gaps where no available methods existed. MEG: magnetoencephalography; PET: positron emission tomography; VSD: voltage-sensitive dye; TMS: transcranial magnetic stimulation; 2-DG: 2-deoxyglucose. (Figure used with permission from [5].)

and fMRI signals. In clinical applications such as epilepsy, high temporal resolution is needed to capture the temporal dynamics of epileptic activity, while high spatial resolution is essential to determine the seizure focus for subsequent possible surgical resection. By exploiting complementary information from EEG and fMRI, unprecedented spatiotemporal accuracy in neuroimaging can be achieved [7]–[9]. Similar examples include, but are not limited to, data fusion studies of fMRI and MEG [10], fMRI and PET, EEG and MEG [11], and fMRI and genetic data [12].

The original JBSS methods were likely canonical correlation analysis (CCA) [60] and partial least squares (PLS) [21]. Both methods utilize second-order statistics (SOS), with CCA emphasizing the role of correlations among data sets and PLS examining covariance information. Since, in general, real neurophysiological data do not strictly follow multivariate Gaussian distributions (that would have higher order statistical moments zero), it is often insufficient to consider only up to SOS (i.e., correlation and covariance) for obtaining a unique JBSS model. Higher-order statistics (HOS) can therefore be employed to enhance the accuracy of estimated sources. Another limitation with traditional implementations of CCA and PLS is that they accommodate only bisets (here we term

two data sets as a *biset* if they are from the same modality and as *bimodal* if they are from two different modalities). However, in many applications, more than two data sets are available, and a better understanding can be achieved from jointly analyzing multiple data sets together [22].

Over the past decade, JBSS methods have been developed to solve two major categories of neurophysiological data analysis challenges. The first challenge is the capacity to simultaneously handle multiple data sets from the same type of neurophysiological data (termed *multiset*)—for example, determining group inferences by combining multiple fMRI data sets from several subjects [22], [26]. The second challenge is the ability to jointly model multiple data sets from distinct types of multimodal neurophysiological data—for example, corticomuscular coupling analysis by integrating EEG, electromyography (EMG), and kinematic (KIN) data [27], [36]. In this article, we provide a taxonomy of the representative JBSS methods characterized by two features: 1) whether or not the method utilizes SOS or HOS; and 2) whether the method is designed for biset/bimodal or multiset/multimodal data. We comprehensively describe these JBSS methods in a comparative manner and offer illustrative numerical simulations. We then

discuss six real-world neurophysiological applications from both the multiset and multimodal perspectives and highlight the benefits of employing JBSS in these applications. Finally, we provide some suggestions on how to select an appropriate JBSS method for a given research question, and we discuss the potential challenges that remain for future JBSS development.

JBSS

In this section, we will provide an overview of representative JBSS methods. We first formulate the JBSS problem and describe relevant notations in detail. We then divide the JBSS methods into four categories (i.e., [SOS, HOS] × [biset/bimodal, multiset/multimodal]) and investigate the theoretical underpinnings of the different JBSS methods. For several fundamental methods, we perform a thorough comparison based on numerical simulations and illustrate their applicability.

A problem formulation and notations

Suppose a total of M data sets are available, which can be collected from the same modality or from multiple modalities. Let $\mathbf{X}^{[m]}$ denote the m th data set with the size $P_m \times N_m$, where generally $P_m < N_m$ is assumed, P_m means the number of observations (e.g., the number of EEG channels), and N_m means the number of variables (e.g., the number of time points). It is worth noting that the exact meanings of observations and variables depend on specific applications. For instance, in group fMRI analysis [23], observations usually indicate the subjects or time courses while variables represent the voxels or features in fMRI data. $\mathbf{X}^{[m]}$ can be expressed by its corresponding column vectors as $\mathbf{X}^{[m]} = [\mathbf{x}_{(1)}^{[m]}, \mathbf{x}_{(2)}^{[m]}, \dots, \mathbf{x}_{(N_m)}^{[m]}]$, ($1 \leq m \leq M$). Here, $\mathbf{x}_{(n)}^{[m]}$, for $1 \leq n \leq N_m$, with size $P_m \times 1$ is the n th realization of the random column vector $\mathbf{x}^{[m]} = [x_1^{[m]}, x_2^{[m]}, \dots, x_{P_m}^{[m]}]^T$, where the superscript T denotes the transpose operation. Each data set is assumed to be a linear mixture of L underlying uncorrelated (or independent) sources:

$$\mathbf{x}^{[m]} = \mathbf{A}^{[m]} \mathbf{s}^{[m]} \quad \text{or} \quad \mathbf{X}^{[m]} = \mathbf{A}^{[m]} \mathbf{S}^{[m]}, \quad 1 \leq m \leq M, \quad (1)$$

where $\mathbf{s}^{[m]}$ is a random source vector with size $L \times 1$, i.e., $\mathbf{s}^{[m]} = [s_1^{[m]}, s_2^{[m]}, \dots, s_L^{[m]}]^T$; and $\mathbf{A}^{[m]}$ is a mixing matrix or modulation profile with size $P_m \times L$, i.e., $\mathbf{A}^{[m]} = [\mathbf{a}_{(1)}^{[m]}, \mathbf{a}_{(2)}^{[m]}, \dots, \mathbf{a}_{(L)}^{[m]}]$. $\mathbf{S}^{[m]}$ is a source matrix with size $L \times N_m$, which can also be expressed by its corresponding column vectors as $\mathbf{S}^{[m]} = [\mathbf{s}_{(1)}^{[m]}, \mathbf{s}_{(2)}^{[m]}, \dots, \mathbf{s}_{(N_m)}^{[m]}]$, for $1 \leq m \leq M$. Here, for $1 \leq n \leq N_m$, $\mathbf{s}_{(n)}^{[m]}$ with size $L \times 1$ is the n th realization of the random column vector $\mathbf{s}^{[m]}$.

Two common formulations of JBSS are shown in Figure 2. In the first formulation, it is assumed that all N_m s are equal to N . This formulation employs the concept of a source component vector (SCV) defined across multiple data sets [15]. The l th SCV, $\mathbf{s}_l = [s_l^{[1]}, s_l^{[2]}, \dots, s_l^{[M]}]^T$ ($1 \leq l \leq L$), is a random vector uncorrelated (or independent) from all other SCVs, and the components $s_l^{[m]}$ within each SCV, defined as source components, are maximally dependent. The l th source component matrix (SCM) can be defined as N realizations of the l th SCV

\mathbf{s}_l [16], i.e., $\mathbf{S}_l = [\mathbf{s}_{l(1)}, \mathbf{s}_{l(2)}, \dots, \mathbf{s}_{l(N)}]$. This formulation is suitable for jointly analyzing the data sets with the same number of variables, e.g., multisubject fMRI data [23], multitrail EEG data [24], [62], and concurrently measured EEG and EMG data [28]. The rows of each SCM may represent the common information contained among the data sets. In the second formulation, it is assumed that all P_m s are equal to P . To better explain the difference between the two formulations, here we define a new concept, the profile component vector (PCV), analogous to the above SCV. Here, the l th PCV, $\mathbf{a}_l = [a_l^{[1]}, a_l^{[2]}, \dots, a_l^{[M]}]^T$ ($1 \leq l \leq L$), is a random vector uncorrelated with (or independent of) all other PCVs, and the components $a_l^{[m]}$ within each PCV, defined as profile components, are maximally dependent. The l th profile component matrix (PCM) can be defined as P realizations of the l th PCV \mathbf{a}_l , i.e., $\mathbf{A}_l = [\mathbf{a}_{l(1)}, \mathbf{a}_{l(2)}, \dots, \mathbf{a}_{l(P)}]^T$ with size $P \times M$. This formulation is suitable for fusing the data sets with the same number of observations but different numbers of variables, e.g., fusing fMRI, structural MRI (sMRI), and EEG data [23], as well as fusing fMRI and single nucleotide polymorphism (SNP) data [41]. The columns of each PCM then exhibit similar changing patterns across multiple modalities.

Before applying JBSS methods, dimension reduction is usually performed on the data to reduce the computational cost and to avoid overfitting. However, since any dimension reduction method will in general result in loss of information, it is important to perform this preprocessing step of JBSS carefully. A typical choice of dimension reduction is PCA [14], [18], [23]. PCA decomposes each individual data set into a set of uncorrelated principal components (PCs) ordered by the variance of each component, and it achieves dimension reduction by removing insignificant PCs that are assumed to represent noise. A potential complication is that the components of one data set $\mathbf{X}^{[m]}$ that account for most of the variance in $\mathbf{X}^{[m]}$ and the components of another data set $\mathbf{X}^{[n]}$ ($m \neq n$) that account for most of the variance in $\mathbf{X}^{[n]}$ may not necessarily correspond with the components that account for most of the correlation between $\mathbf{X}^{[m]}$ and $\mathbf{X}^{[n]}$ [66]. To address this issue, a number of methods have been designed to preprocess multiple data sets jointly, such as generalized singular value decomposition [63], [64], combining PCA/PLS with CCA [29], [65], [66], and modeling overall covariation information [27], [67]. Heuristics (such as selecting the number of components explaining a predefined fraction of variance in the data) or information-theoretical criteria are then generally used to determine the optimal number of components to retain [66], [68]. Under the JBSS framework, since each component in one data set has its dependent counterpart in another data set, the numbers of components in all data sets are normally assumed to be the same. The key idea of those methods is to make the components carry as much variation as possible within each data set and yet still remain as correlated as possible across multiple data sets.

A taxonomy of JBSS methods

Historically, JBSS methods have evolved from SOS to HOS formulations and from biset/bimodal to multiset/

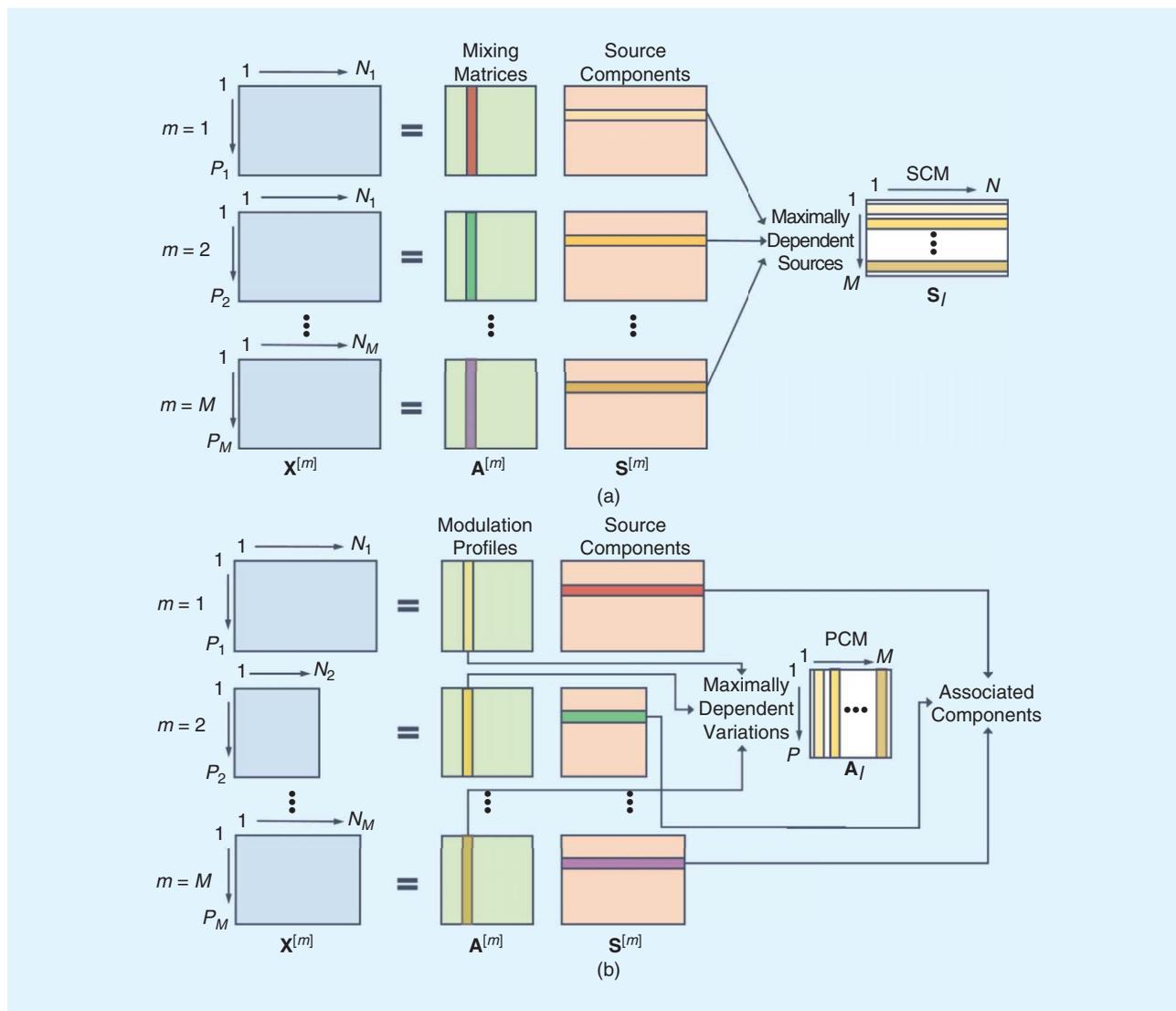


FIGURE 2. Two popular formulations of JBSS. (a) The common variation information across multiple data sets is assumed to exist in the horizontal dimension of source matrices. (b) The common variation information is assumed to exist in the vertical dimension of mixing matrices.

multimodal implementations. We can thus divide JBSS methods into four categories as shown in Figure 3. Below we elaborate representative methods within the four categories, demonstrating the advantages and disadvantages of different methods and providing guidance for their use in different situations.

BISET/BIMODAL using SOS

CCA and PLS are two classic JBSS methods that utilize SOS (i.e., covariance and correlation) and only process two data sets at once. Various extensions to these methods, including multiway data and/or sparsity constraints, have led to several variants of CCA and PLS (still within this category), including multiway PLS (N-PLS) [35], multiblock PLS (mb-PLS) [36], PLS+CCA [29], sparse CCA [30], ensemble empirical mode decomposition-CCA (EEMD-CCA) [31], [32], and multiway CCA [33], described below.

CCA

CCA aims to seek two sets of basis vectors, one for $\mathbf{X}^{[1]}$ and the other for $\mathbf{X}^{[2]}$ (i.e., $M = 2$), such that the correlations between the projections of the variables onto these basis vectors are mutually maximized [37]. Through a deflationary procedure, several pairs of basis vectors can be derived to ensure that the corresponding canonical variates between two data sets are maximally correlated and the canonical variates within each data set are mutually uncorrelated. When the canonical variates denote the sources as formulated in Figure 2(a), CCA can be utilized to extract highly correlated sources between two data sets. Nevertheless, in certain cases, such as the case that the dimension of observations $P_m = 1$, additional techniques such as EEMD, whereby a single channel is decomposed into different progressively smoother waveforms, are needed to assist in the joint analysis. EEMD-CCA methods have been proposed by Sweeney et al. and Chen et al. to remove

motion and muscular artifacts in functional near-infrared spectroscopy [31] and single-channel EEG data [32]. They first decomposed the single-channel data into multichannel data by EEMD, then applied CCA to both the obtained multichannel data and a time-delayed version of the multichannel data. They then kept the highly autocorrelated sources and discarded the sources with low autocorrelation. Zhang et al. extended the standard CCA method in a steady-state visual evoked potential (SSVEP) brain-computer interface (BCI) application [33], by jointly analyzing a three-way EEG tensor (channel \times time \times trial) and a reference data set of sine-cosine (harmonic \times time).

CCA can also be formulated as in Figure 2(b), where the goal is to find the covarying modulation profiles between two data sets. In this case, the modulation profiles, instead of the sources, become the canonical variates. This formulation is particularly useful for data fusion between two different modalities. Additional constraints, such as sparseness, can be incorporated into the cost function to improve the interpretability of the results. For example, Avants et al. introduced sparse CCA by imposing an L_1 penalty on the weight vectors in applications of fusing diffusion tensor imaging (DTI) and T1-weighted structural imaging data collected from healthy subjects and subjects with dementia. The sparsity constraint promotes the exclusion of voxels in each modality that are of little importance in predicting voxels in the other modality [30].

PLS

PLS explores the covariation between predictor variables and response variables (i.e., $M = 2$) and tries to find a new set of latent variables (LVs) that maximally relate them [34]. In other words, the covariance between pairs of LVs should be maximized. The PLS problem can be solved by performing an eigendecomposition to derive the first pair of weight vectors and thus LVs [38]. To obtain subsequent weights, the algorithm can be repeated with deflated $\mathbf{X}^{[1]}$ and $\mathbf{X}^{[2]}$ matrices. Refer to [29] for details of the calculation procedure. The conventional implementation of PLS is limited in that it can only simultaneously handle two flat-view (i.e., two-way) data matrices. Yet many practical applications may involve a multidimensional (i.e., multiway) data array, e.g., a time-varying EEG spectrum with spatial, spectral, and temporal information. A multiway PLS approach has been proposed that finds correlations between time-varying EEG spectra and fMRI time courses from a single subject by fitting multilinear models simultaneously for both data sets with a constraint of maximizing the covariance between corresponding temporal signatures of the EEG and fMRI [35]. Similarly, a multiblock PLS framework to accommodate concurrently recorded EEG and EMG data from multiple subjects can be achieved by incorporating a hierarchical structure, including individual data blocks at a sublevel and the sublevel information integration at a superlevel [36].

PLS+CCA

PLS and CCA have complementary objectives, and it is possible to combine them. The goal of PLS is to construct LVs

| | Biset/Bimodal | Multiset/Multimodal |
|-----|---|---|
| SOS | 1 PLS, CCA, Sparse CCA, Multiway CCA, EEMD-CCA, N-PLS, mb-PLS, PLS+CCA | 3 MCCA, EEMD-MCCA, JDIAG-SOS, JMSF |
| HOS | 2 pICA, IC-PLS | 4 Group ICA, jICA, Linked ICA, MCCA+jICA, Three-Way pICA, IVA |

FIGURE 3. The four categories of JBSS methods. The first category includes methods handling two data sets and exploiting SOS; the second category includes methods handling two data sets and exploiting HOS; the third category includes methods handling multiple data sets and exploiting SOS; and the fourth category includes methods handling multiple data sets and exploiting HOS.

that best explain the variation of one data set and still be well correlated to the corresponding LVs in the other data set. In other words, the first priority of PLS is to find the LVs that can explain a significant proportion of variance in each data set, and the second priority is to find LVs with relatively high correlation coefficients between the two data sets. In contrast, the only objective of CCA is to maximize correlation coefficients between the corresponding LVs extracted from both data sets. From this angle, PLS has the advantage that the LVs may contain major variation information for individual data sets, while the ones extracted by CCA may be trivial—for example, a small amount of common noise corrupting both data sets. PLS can also handle high dimensional and collinear data, which is often the case in real-world medical applications, while applying CCA directly to such data may be ill-conditioned.

However, a potential limitation of PLS is that higher covariance between two corresponding LVs may merely result from the larger variance of individual LVs. This may not necessarily imply strong correlations between them. To overcome this, CCA is a powerful tool to ensure that the corresponding LVs extracted from the two data sets are highly correlated irrespective of their scales. To take advantage of the benefits of both CCA and PLS, Chen et al. proposed combining them as a two-step method, termed PLS+CCA [29]. In the first step, PLS is performed to extract LVs that explain the variance in individual data sets yet are well correlated to the LVs in the other data set, thus removing trivial and irrelevant information across data sets. In addition, this step can also prevent ill-conditioning sometimes observed when CCA is applied directly to raw data. In the second step, CCA is

applied to the PLS-derived LVs. After these two steps, it is ensured that the extracted components are maximally correlated across data sets while still explaining the variation within each individual data set.

BISET/BIMODAL using HOS

Both PLS and CCA exploit SOS and can only extract uncorrelated LVs. This lack of uniqueness may impede interpretations of extracted LVs in real applications [39], which can sometimes be circumvented by exploiting HOS. If the data are not strictly multivariate Gaussian, SOS (i.e., correlation and covariance) will be insufficient for obtaining a unique LV model. ICA attempts to find mutually statistically independent sources that are linearly superimposed to create multivariate data. Since independence is a much stronger condition than uncorrelatedness, algorithms for ICA, which typically employ HOS criteria related to information theory and/or non-Gaussianity, can obtain unique solutions. ICA has been empirically shown to be very useful in many biomedical applications [1], suggesting a use for incorporating ICA into existing JBSS methods. The second category of JBSS methods, represented by the parallel ICA (pICA) [41] and IC-PLS [28], were designed for this purpose.

ICA has been empirically shown to be very useful in many biomedical applications, suggesting a use for incorporating ICA into existing JBSS methods.

pICA

pICA was developed to identify maximally independent components from each of two modalities ($M = 2$) and connections between them through enhancing intrinsic interrelationships [40]. PICA largely relaxes the rigid assumption of sharing the same mixing matrix in the joint ICA (jICA) model ($M \geq 2$) [26], which is also designed to fuse multimodal data (but will be subsequently described in the section “Multiset/Multimodal Using HOS”). It maximizes the independence within each modality using an entropy-based cost function while also identifying intermodality correlations through adding a squared correlation term. In the pICA model, there are a total of three terms that need to be optimized simultaneously—two of them relate to maximizing the independence of sources for the two modalities separately, while the third term is the determination of the relationship between the two modalities. During optimization, adaptively adjusting the learning rates is critically important for balancing the three aspects in the cost function [41].

pICA is formulated as in Figure 2(b). The vertical dimensions P_m of the data sets $\mathbf{X}^{[m]}$ may represent, for example, the number of healthy subjects and patients with disease. The assumption is that the patterns of intersubject modulation across two modalities are similar or covarying. This could be reflected by the correlation between the corresponding columns of modulation profiles $\mathbf{A}^{[m]}$. The associated components or sources $\mathbf{S}^{[m]}$ from different modalities may provide the interpretation from a different view of brain function or structure. Therefore, pICA is particularly suited

for data fusion problems especially in the medical realm. For instance, it has been employed to link fMRI and genetic data [41], and EEG and SNP data [42].

IC-PLS

IC-PLS is designed to combine the advantages of PLS and ICA [28]. The goal of IC-PLS is to extract maximally independent sources from two data sets while keeping the corresponding sources correlated across the two data sets ($M = 2$). Thus, the following conditions should be satisfied simultaneously: first, the covariance between the corresponding LVs across the two data sets should be maximized; second, the independence (e.g., non-Gaussianity) of the extracted LVs within each data set should be maximized. This leads to a multi-objective optimization problem, encapsulating three maximization objectives. A solution using an approximate Newton iteration approach has been suggested [28]. Similar to pICA, care must be taken during weight adjustment to balance the three terms during cost function optimization. IC-PLS is formulated as in Figure 2(a). It emphasizes the role of the source components $\mathbf{S}^{[m]}$,

but pays little attention to the mixing matrices $\mathbf{A}^{[m]}$. It has been applied to corticomuscular coupling analysis, extracting maximally independent source pairs from concurrent EEG and EMG data in order of relevance [28].

MULTISET/MULTIMODAL using SOS

The previously described methods were designed for analysis of biset or bimodal data. However, more than two data sets from different modalities (i.e., multimodal) or from the same modality (i.e., multiset) are frequently available. Thus a third category of JBSS methods exploiting SOS have been timely developed, including multiset CCA (MCCA) [20], joint multimodal statistical framework (JMSF) [27], EEMD-MCCA [25], and joint diagonalization of second-order cumulant matrices (JDIAG-SOS) [44]. These concepts are expanded upon below.

MULTISET CCA

MCCA ($M \geq 2$) extends the theory of CCA ($M = 2$) [60] to more than two random vectors and identifies canonical variates that summarize the correlation structure among multiple random vectors by linear transformations [20]. Unlike CCA, where correlation between two canonical variates is maximized, MCCA aims to optimize an objective function to make the canonical variates achieve the maximum overall correlation [20]. Recently, it has been shown that MCCA can be used to achieve JBSS [22], allowing for jointly analyzing multiset data. In such an approach, MCCA is implemented in multiple stages, such that one group of canonical variates is obtained at each stage through optimizing the objective function with respect to a set of transformation weight vectors [22]. Thus the l th group of canonical variates from M data sets $\mathbf{X}^{[m]}$ ($m = 1, 2, \dots, M$) is defined as $\mathbf{w}_l^{[m]T} \mathbf{X}^{[m]}$

$[l = 1, 2, \dots, L, L = \min(\text{rank}(\mathbf{X}^{[m]})]$, and they represent the M rows of the l th SCM, as shown in Figure 2(a) (the first formulation). At the first stage, one of the five given cost functions in [20], e.g., maximizing the summation of all squared mutual correlation coefficients (SSQCOR) between the first group of canonical variates, can be directly utilized. At the second and higher stages, the estimated canonical variates are constrained to be uncorrelated to the ones estimated in the previous stages using the same cost function. A deflationary procedure can be employed to find the weight vectors. Note that if each data set is whitened in advance, the demixing matrix for each data set becomes an orthonormal matrix. In this case, removing the already extracted sources in the deflationary procedure (i.e., the uncorrelatedness constraint) can be converted to the orthogonal constraint, indicating the current demixing vector to be orthogonal to the previously obtained ones. Imposing the orthogonal constraint on the demixing vectors has the benefit that each joint source extraction stage can be solved as a constrained optimization problem [20], [22]. Details about the model formulation and the implementation of MCCA can be found in [22].

MCCA is an effective and efficient way to decompose individual data sets into uncorrelated canonical variates while keeping the corresponding canonical variates across data sets maximally correlated. It has been applied to fMRI data collected from a number of subjects when performing a visuomotor task, resulting in a meaningful group-level consensus [23]. It has also been adopted to jointly analyze EEG data collected from SSVEP-based BCI experiments [24]. In some special cases, when only single-channel data sets are available, additional techniques such as EEMD are needed to assist the joint analysis. For example, Chen et al. proposed an EEMD-MCCA method for muscular artifact removal in a single-channel of EEG and demonstrated its advantage over the EEMD-CCA using simulations [25].

MCCA for multimodal data fusion

MCCA has also been frequently employed for fusing information from multiple complementary modalities [22], [23], [43], [45]. Unlike the formulation in Figure 2(a), MCCA for multimodal data fusion models the covariation along the vertical dimension of modulation profiles $\mathbf{A}^{[m]}$, as shown in Figure 2(b), and treats the corresponding columns of $\mathbf{A}^{[m]}$ as a group of canonical variates. The l th group of canonical variates from M data sets $\mathbf{X}^{[m]}$ is defined as $\mathbf{X}^{[m]}\mathbf{w}_l^{[m]}$ and represents the M columns of the l th PCM, as shown in Figure 2(b) (the second formulation). Although MCCA for multimodal data fusion is different from MCCA for multi-set data analysis in terms of the formulation and thus leads to different models with different definitions of canonical variates, they both are based on Kettenring's original description of MCCA [20], i.e., the same cost function and the same optimization procedure [22], [23], [45]. After the modulation profiles $\mathbf{A}^{[m]}$ are obtained, the associated components $\mathbf{S}^{[m]}$ can be derived in the least-square sense as $\mathbf{S}^{[m]} = (\mathbf{A}^{[m]T}\mathbf{A}^{[m]})^{-1}\mathbf{A}^{[m]T}\mathbf{X}^{[m]}$.

An excellent example of the power of MCCA as a tool for multimodal data fusion is in the field of schizophrenia. Data from fMRI, sMRI, and EEG data were collected from patients diagnosed with schizophrenia and healthy controls when performing an auditory oddball task. The associated components linked by the covarying modulation profiles provided additional insight into connectivity across brain networks and changes due to disease. As expected, the fusion of multiple modalities was more informative than the fusion of only two modalities [45].

JMSF

JMSF is a two-step method simultaneously modeling multiple data sets ($M \geq 2$) as formulated in Figure 2(a). It is suitable for modeling multiple data sets with redundant information, and it has been applied to concurrent EEG, EMG, and behavioral data for corticomuscular coupling analysis [27]. In the first step, denoted as multi-LV extraction, the l th group of sublatent variables (subLVs) are defined in each data set as the linear combinations $\mathbf{w}_l^{[m]T}\mathbf{X}^{[m]}$. One common super latent variable (supLV), \mathbf{t}_g , is designed to relate the first group of subLVs. The subLVs in each data set carry associated variation information, and $(\mathbf{w}_l^{[m]T}\mathbf{X}^{[m]}\mathbf{t}_g)^2$ models the covariance information between the subLV $\mathbf{w}_l^{[m]T}\mathbf{X}^{[m]}$ and the supLV \mathbf{t}_g . The supLV \mathbf{t}_g relates all corresponding subLVs across M data sets simultaneously and plays a role as a link bridge. The weight vectors $\mathbf{w}_l^{[m]}$ can be obtained by solving the constrained optimization problem with Lagrange multipliers, and the subsequent weight vectors $\mathbf{w}_l^{[m]}$ ($l = 2, \dots, L$) can be derived using a deflationary procedure [27]. The extracted subLVs in this way will carry as much variation as possible within each data set, and while the corresponding subLVs will be correlated as closely as possible. This step is essentially a preliminary LV preparation for keeping as much of the relevant variance across multiple data sets as possible. In the second step, the extracted subLVs are treated as the inputs to MCCA, and the canonical variates with maximum correlation are finally recovered.

JDIAG-SOS

JDIAG-SOS was developed to jointly diagonalize multiple second-order cumulant matrices ($M \geq 2$), and is formulated as in Figure 2(a). An off-norm cost function has been designed to realize the joint diagonalization [44], and it can be solved using a gradient search and more efficiently by iteratively solving orthogonal Procrustes problems. Compared to MCCA, it avoids the limitation of ad hoc cost functions and the error accumulation of a deflationary procedure for cost optimization. However, it still imposes an orthogonality constraint on the estimated demixing matrices, as does MCCA. While JDIAG-SOS has been employed to separate the fetal heartbeat in electrocardiogram (ECG) [44], it has not yet been widely applied to neurophysiological data analytics.

MULTISET/MULTIMODAL using HOS

The fourth category of JBSS methods, which exploit HOS, has been developed under different statistical assumptions.

Representative methods within this category include group ICA [14], [69], [70], jICA [26], linked ICA [46], MCCA+jICA [47], [48], three-way pICA [49], and independent vector analysis (IVA) [17]. These are expanded upon below.

Group ICA

Group ICA was initially motivated by the need for drawing inferences about groups of subjects (e.g., multisubject fMRI data) and is formulated as in Figure 2(a). Several Group ICA approaches, each with different assumptions and data grouping strategies, have been developed ($M \geq 2$) [14], [69], [70]. The most popular data organization form is concatenating the data sets $\mathbf{X}^{[m]}$ along the vertical dimension (e.g., temporal concatenation in fMRI). This assumes that the data from each subject have been spatially transformed to a common template. A key assumption with the vertical concatenation is that each data set $\mathbf{X}^{[m]}$ has a unique mixing matrix $\mathbf{A}^{[m]}$, but all data sets share a common source subspace \mathbf{S} , within which the statistical independence of the sources is maximized. Group ICA has been widely applied to fMRI data for group analysis, and it is attractive because the unique time courses for each subject can be detected. An additional back-reconstruction also allows for capturing variations in subject-specific spatial maps [14]. Nevertheless, this approach requires the very restricting assumption that all subjects can be meaningfully spatially transformed to the same space so that each voxel can be considered comparable across subjects. To mitigate problems with subtle misregistration, spatial smoothing is often employed, but this degrades the spatial resolution of the data.

jICA

Unlike the group ICA's organization of data sets from one single modality (multiset), jICA has been developed to accommodate fusing of information from multiple modalities (multimodal) collected in the same set of subjects ($M \geq 2$) [26]. It is formulated as in Figure 2(b). In jICA, the data sets $\mathbf{X}^{[m]}$ are concatenated along the horizontal dimension (e.g., voxels/time courses) prior to the ICA analysis. jICA assumes that all modalities share the same modulation profile \mathbf{A} . When it is applied to multiple modalities collected from groups of subjects, the assumption is that the intersubject variations across these modalities are exactly the same. Although this assumption is fairly stringent, jICA has the advantage of providing a parsimonious way to link data sets from multiple modalities, and its utility has been demonstrated in a number of medical applications [14].

Linked ICA

Linked ICA is designed for discovering common features across multiple modalities based on a modular Bayesian framework ($M \geq 2$) [46]. These modalities can potentially have completely different units, noise level, voxel counts, spatial smoothnesses, and intensity distributions. Linked ICA can be configured to allow tensor ICA or spatially concatenated ICA decompositions, or a simultaneous combination of both [Figure 2(b)]. In linked ICA, each modality is

modeled using Bayesian tensor ICA, and all of the modalities share the same modulation profile (i.e., intersubject variation matrix or subject loading matrix). This fully probabilistic approach, implemented using variational Bayes, automatically determines the optimal weighting of each modality and can also detect single-modality structured components when present [46]. Linked ICA has been applied to morphological and diffusion sMRI data collected from AD patients and age-matched controls. One derived joint component had subject loadings strongly correlated with age (0.49) and pathology (0.30) [46].

MCCA+jICA

MCCA+jICA is a two-step method, combining MCCA and jICA ($M \geq 2$) [47], [48], which can be formulated either as in Figure 2(a) or (b). When the former formulation is adopted, it concentrates on the extraction of independent source components. Both MCCA and jICA have their individual advantages and disadvantages, and, fortunately, these two methods are complementary. As previously mentioned, MCCA is able to jointly extract the group of corresponding sources from each data set through maximizing the correlations among the extracted sources. However, MCCA requires a stringent assumption that correlation coefficients of the corresponding sources between multiple data sets be sufficiently distinct. MCCA may fail to separate sources whose correlation coefficients are equal or very close, as is frequently seen in biomedical data. Specifically, if L sources can be correspondingly extracted from each of M data sets, the following requirement must be met to successfully recover the sources by MCCA: $|r_{i,j}^{(\alpha)}| \neq |r_{i,j}^{(\beta)}|, (1 \leq \alpha < \beta \leq L, \forall i, j \in \{1, 2, \dots, M\})$. Here, $|r_{i,j}^{(\beta)}|$ represents the correlation coefficient between the β th source from the i th data set and the β th source from the j th data set. Therefore, the components extracted by MCCA should be regarded as incompletely decomposed sources, i.e., mixtures of the true underlying independent components, especially since MCCA only exploits SOS.

jICA maximizes the independence of joint sources of multiple data sets. However, in the jICA framework, all modalities are assumed to share the same mixing matrix—something that is not always easily satisfied in practice. By combining MCCA and jICA, MCCA first links multiple data sets via correlation and specifies the associated components across data sets; then, jICA is performed on the horizontally concatenated components to extract joint independent components. In this manner, MCCA makes the jICA step more reliable by providing a closer initial match via correlation, while jICA further decomposes the remaining mixtures in the associated components and relaxes the requirement of sufficiently distinct canonical correlation coefficients. MCCA+jICA demonstrated superior performance compared to MCCA and jICA alone in both simulations and real studies where EEG, EMG, and behavioral data (dynamic forces recorded from a pressure-responsive bulb) were concurrently collected from normal subjects and patients with PD [48]. When MCCA+jICA is formulated as in Figure 2(b), it prioritizes the maximization of

correlation between modulation profiles, allowing both highly and weakly connected modulations as well as joint independent components [47]. It is particularly suitable for jointly analyzing data sets from multiple modalities in the same set of subjects. For instance, it has been applied to fMRI and DTI data from healthy controls and people with schizophrenia patients and bipolar disorder [47].

Three-way pICA

Recall that pICA analyzes two data modalities concurrently and allows the integration of data from different contexts and feature spaces. This flexibility makes pICA particularly suitable for data fusion (e.g., MRI and genetic data). Three-way pICA extends the concept of pICA to include three modalities ($M = 3$), which is formulated as in Figure 2(b) [49]. It also maximizes independence within each modality using an entropy-based cost function, while identifying intermodality correlations through adding three squared correlation terms. Similar to pICA, an adaptive optimization procedure is employed. These online adjustments guarantee the convergence of the three-way pICA. Compared to MCCA for multimodal data fusion [23], [45], three-way pICA exploits HOS for obtaining statistically independent components, which likely leads to a more accurate and meaningful estimation. Compared to jICA and linked ICA, three-way pICA relaxes the strong assumption of the same modulation profile. Also, in contrast to mCCA+jICA, three-way pICA explicitly incorporates the information provided by all modalities in one comprehensive data decomposition. It has been applied to fMRI, sMRI, and SNP data for investigating genetic effects on alcohol dependence and performed better than pICA and separate ICA (sICA) in identifying pairwise links between modalities and estimating independent components [49].

IVA

IVA is a generalization of ICA from one to multiple data sets ($M \geq 2$), originally designed to address the permutation problem in the frequency domain for the separation of acoustic sources [15]. IVA can be formulated within a general JBSS framework to ensure that the extracted source components are independent within each data set and maximally correlated across multiple data sets [17]. Specifically, the goal of IVA is to identify L independent SCVs s_l from M data sets \mathbf{X}_m . This can be achieved by minimizing the mutual information among the estimated SCVs \tilde{s}_l . It can be proven that minimizing the IVA cost function is equivalent to simultaneously minimizing the entropy of all components $\tilde{s}_l^{[m]}$ and maximizing the mutual information within each estimated SCV \tilde{s}_l . IVA can ultimately solve the problem of permutation ambiguity when applying BSS techniques to multiple data sets. IVA has been shown to achieve superior performance than previous techniques in simulation studies [17]. Although IVA was formulated as in Figure 2(a) in [17], it is quite straightforward to formulate IVA as in Figure 2(b), resembling the relationship between MCCA for multiset

data analysis and MCCA for multimodal data fusion. IVA generalizes MCCA to the case where both SOS and HOS are taken into account and where the demixing matrix is not constrained to be orthogonal [16]. In a recent study, IVA was used to fuse EEG, functional, and sMRI in the manner of Figure 2(b) [71].

Different implementation algorithms of IVA involve the assumption of specific SCV distributions. The most widely used ones include IVA-L [15], which assumes that each SCV follows a multivariate Laplace distribution that is isotropic and possesses no second-order correlation, and IVA-G [17], which assumes each SCV is multivariate Gaussian distributed. In applications like speech recognition [15], the second-order information across data sets may be minimal. However, in most neurophysiological applications, a second-order dependence across data sets is likely. IVA has been utilized in a number of applications, such as group fMRI analysis [16] and concurrent multidimensional EEG and unidimensional KIN data analysis [19]. Recently, IVA-GL, using the IVA-G solution to initialize the IVA-L algorithm, has been recommended for fMRI applications [18]. The implementation first takes into account full second-order dependence among entries of an SCV by IVA-G. The estimates of the demixing matrices are then employed to initialize IVA-L, and HOS are taken into account by assuming a Laplacian distribution for each entry within an SCV. However, since it is a two-step method, IVA-GL may not work very well if the SCVs are not Laplacian distributed. To summarize the different methods, Table 1 provides a comprehensive summary of all aforementioned JBSS methods in terms of different categories, motivations, optimization criteria, solutions, software, and related major works.

Numerical simulations

In this section, we provide numerical simulations to illustrate the applicability and the performance of several fundamental JBSS methods. Without loss of generality, we generate three data sets ($M = 3$) and use the JBSS methods related to the first formulation [shown in Figure 2(a)] for demonstration. The studied representative methods include jICA, MCCA, JDIAG-SOS, MCCA+jICA, IVA-G, IVA-GL, and sICA. In the following simulation, FastICA is employed as the ICA algorithm [61], and the SSQCOR cost function is used to implement MCCA due to its robustness [17], [22].

Data generation

The following six sources were generated and analyzed as in [48]:

$$\begin{aligned} s_1 &= \sin(0.015n) + \cos(0.005n) \\ s_2 &= 2\cos(0.08n)\sin(0.006n) \\ s_3 &= \text{ECG}, \quad s_4 = \text{EMG} \\ s_5 &= 1.5\cos(0.01n)\sin(0.5n) \\ s_6 &= 1.5\sin[0.025(n + 63)]\sin(0.2n), \end{aligned} \quad (2)$$

Table 1. A summary of the representative BSS methods.

| Methods | Categories | Motivations | Optimization criteria | Solutions and software | Reference |
|----------------|--|--|---|---|-----------------------|
| CCA | $M = 2$, SOS, Figure 2(a) and (b) | Determine correlated sources across two data sets without considering variation information. | Maximize the correlation between corresponding sources across two data sets. | Eigendecomposition; built within MATLAB | [30]–[33] |
| PLS | $M = 2$, SOS, Figure 2(a) | Determine correlated sources across two data sets considering variation information | Maximize the covariance between corresponding sources across two data sets. | Eigendecomposition, NIPALS, etc.; Built within MATLAB or contact authors | [34]–[36] |
| PLS+CCA | $M = 2$, SOS, Figure 2(a) | Take advantage of both PLS and CCA. | Apply PLS to remove insignificant sources first and then apply CCA. | Eigendecomposition and a deflationary procedure | [29] |
| pICA | $M = 2$, HOS, Figure 2(b) | Identify both independent sources and their dependent relation between two modalities. | Maximize the independence of sources within each set/modality and meanwhile maximize the correlation between corresponding modulation profiles across two data sets. | Gradient descent; http://mialab.mrn.org/software/#fusion | [40]–[42] |
| IC-PLS | $M = 2$, HOS, Figure 2(a) | Extract independent sources within each data set while keeping the corresponding sources correlated across two data sets. | Maximize the independence of sources within each set/modality and meanwhile maximize the covariance between corresponding sources across two data sets. | Approximate Newton iteration; contact authors | [28] |
| MCCA | $M \geq 2$, SOS, Figure 2(a) and (b) | Detect highly correlated sources across multiple data sets or common modulation profiles across modalities with associated components. | Maximize the overall correlation between the corresponding sources or modulation profiles across multiple data sets. | Lagrangian derivative iterative maximization; http://mlsp.umbc.edu/codes/mcca.m | [22]–[25], [43], [45] |
| JMSF | $M \geq 2$, SOS, Figure 2(a) | Extract highly correlated sources across multiple data sets considering variation information. | Maximize a linked overall covariance between the corresponding sources across multiple data sets first, and then apply MCCA. | Lagrangian method with a deflationary procedure; contact authors | [27] |
| JDIAG-SOS | $M \geq 2$, SOS, Figure 2(a) | Detect correlated sources across multiple data sets without deflationary error accumulation. | Jointly diagonalize multiple second-order cumulant matrices by minimizing an off-normal cost function. | Iteratively solving orthogonal procrustes problems; http://mlsp.umbc.edu/codes/lbss-sos.m | [44] |
| Group ICA | $M \geq 2$, HOS, Figure 2(a) | Draw inferences about groups of subjects. | Concatenate multiple data sets along the vertical dimension and maximize the independence of sources with the assumption that all data sets share a common source subspace. | Typical ICA algorithms; http://mialab.mrn.org/software/#gica | [14] |
| jICA | $M \geq 2$, HOS, Figure 2(b) | Identify covarying modulation profiles across subjects among modalities with their linked components. | Concatenate multiple data sets along the horizontal dimension and maximize the independence of joint sources with the assumption that all data sets share the same mixing matrix. | Typical ICA algorithms; http://mialab.mrn.org/software/#fusion | [26] |
| Linked ICA | $M \geq 2$, HOS, Figure 2(b) | Discover common features across multiple modalities. | Model each modality (group) using Bayesian tensor ICA with the assumption that all modalities share the same modulation profile. | Variational Bayes; contact authors | [46] |
| MCCA+jICA | $M \geq 2$, HOS, Figure 2(a) and (b) | Achieve more accurate joint source separation. | Apply MCCA first to make the jICA job more reliable by providing a closer initial match via correlation, and then jICA further decomposes the remaining mixtures in sources. | Straightforward combination | [47], [48] |
| Three-way pICA | $M = 3$, HOS, Figure 2(b) | Identify both independent sources and their connections across three modalities. | Similar to pICA; but generalize the pICA concept by including three modalities. | Gradient descent; contact authors | [49] |
| IVA | $M \geq 2$, HOS, Figure 2(a) and (b) | Estimate both independent and self-dependent SCVs. | Minimize the mutual information among SCVs. | Gradient descent; http://mlsp.umbc.edu/resources | [15]–[19] |

where n denotes the sample index vector, valued from 1 to 1,000, and $s_i(i = 1, 2, \dots, 6)$ represents six simulated sources. Note that here s_i is a row vector. Three mixed data sets $\mathbf{X}^{[1]}$, $\mathbf{X}^{[2]}$, and $\mathbf{X}^{[3]}$ were generated as follows, with each column denoting one observation in their respective data set:

$$\mathbf{X}^{[m]} = \mathbf{A}^{[m]} \cdot \mathbf{S}^{[m]}, \quad m = 1, 2, 3, \quad (3)$$

where $\mathbf{S}^{[1]} = [s_1; s_3; s_2; s_4]$, $\mathbf{S}^{[2]} = [s_1; s_3; s_2^*; s_5]$, and $\mathbf{S}^{[3]} = [s_1; s_3; s_2^{**}; s_6]$ with the size $4 \times 1,000$. s_2^* and s_2^{**} denote the five-point and 10-point delayed versions of s_2 , respectively. This is slightly different from the source configuration of the simulation in [48]. The underlying sources in each data set are shown in Figure 4(a). The sources s_1, s_2 , and s_3 exist in all data sets, representing common information across them. The

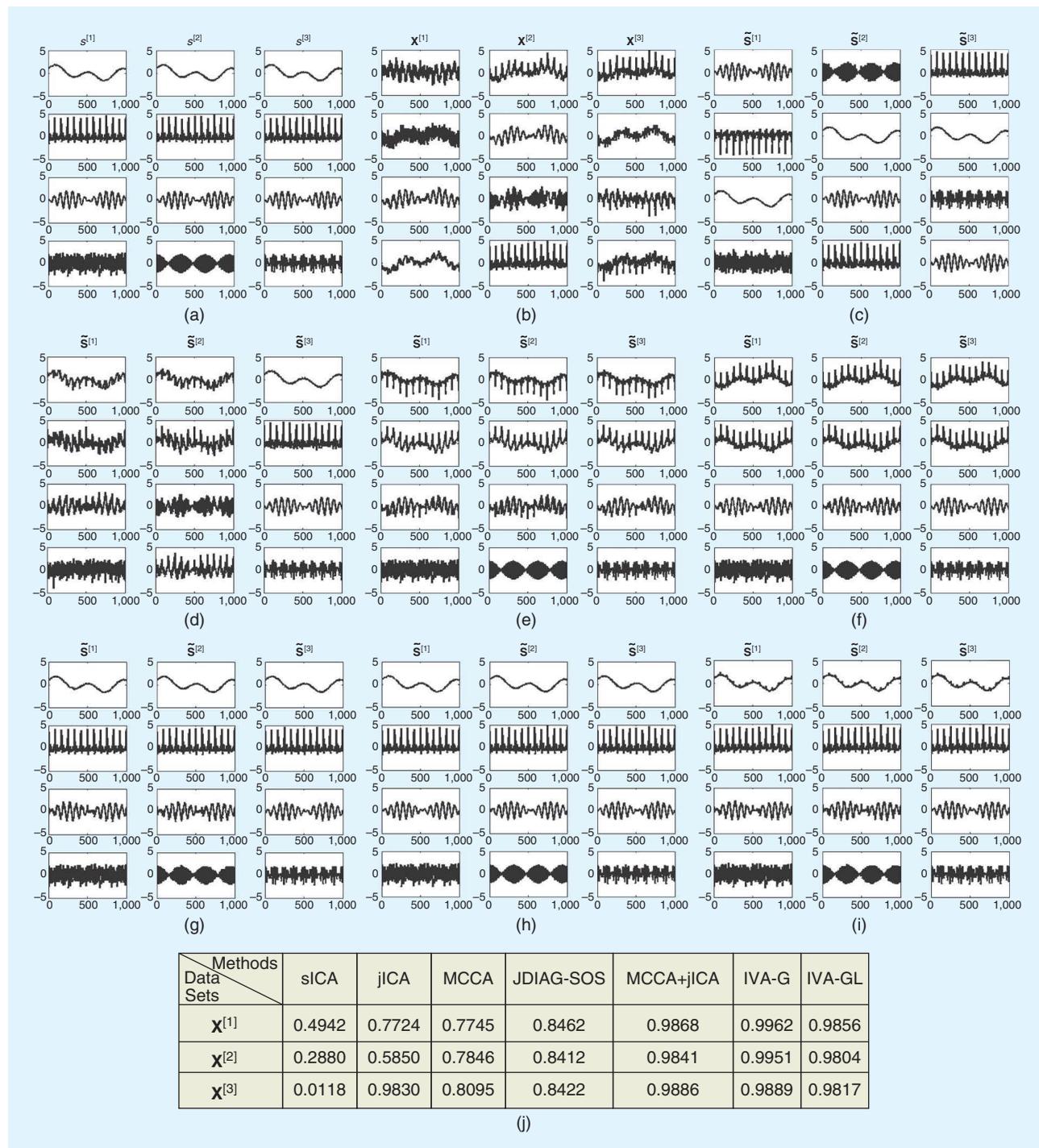


FIGURE 4. The original sources in the three mixed data sets and the recovered sources by employing seven different methods: (a) true sources, (b) mixed data sets, (c) sICA, (d) jICA, (e) MCCA, (f) JDIAG-SOS, (g) MCCA+jICA, (h) IVA-G, and (i) IVA-GL. (j) A reference for performance comparisons among the different methods.

source s_4 only contributes to $\mathbf{X}^{[1]}$, s_5 only to $\mathbf{X}^{[2]}$, and s_6 only to $\mathbf{X}^{[3]}$, representing unique information within each data set. $\mathbf{A}^{[m]}$ s denote randomly generated mixing matrices with the size 4×4 . Random Gaussian noise with a mean of 0 and a standard deviation of 0.05 was independently generated and added to each source matrix $\mathbf{S}^{[m]}$ before generating the mixed data sets [shown in Figure 4(b)].

Simulation results

The source separation results using the seven methods are shown in Figure 4(c)–(i). A quantitative performance evaluation based on average correlation coefficient (ACC) was also made for reference [Figure 4(j)]. In this simulation, ACC is defined as the average of the overall correlation between the corresponding original and estimated sources $\text{acc}(m) = \sum_{i=1}^4 \text{Corr}[\mathbf{S}^{[m]}(i,:), \hat{\mathbf{S}}^{[m]}(i,:)]/4$, ($m = 1, 2, 3$).

The results demonstrated that there were distinct differences between methods. As anticipated, sICA was capable of recovering the original sources accurately in all data sets [Figure 4(c)] but was unable to meaningfully relate the three data sets or rank the sources appropriately. Although a source alignment method could be subsequently used to order the sources, this may introduce ambiguity, especially when the estimated number of sources is high. For jICA, the first jointly extracted sources, although perhaps less accurate, were at least qualitatively similar to each other [Figure 4(d)]. However, the remaining extracted sources seemed to be uninformative, likely due to the stringent assumption that all data sets share the same mixing matrix. The sources extracted by MCCA were at least automatically ordered in terms of their overall correlation coefficient values among the data sets [Figure 4(e)]. However, compared to the original sources, the extracted sources were distorted, suggesting 1) performance of MCCA may suffer when correlation coefficients between corresponding sources are equal or very close [22] and 2) uncorrelatedness may not be a sufficiently rigorous criterion to accurately recover the underlying sources. JDIAG-SOS had an obvious improvement for source recovery. The final two sources across data sets were accurately identified and aligned [Figure 4(f)], as reflected by the increase of ACC [Figure 4(j)]. However, the first two sources were still mixed together, again, likely due to the inadequacy of SOS.

When MCCA+jICA was employed, the performance was improved [Figure 4(g)]. The sources within all data sets were accurately identified and ordered with the focus entirely on sources common across the three data sets. MCCA+jICA mitigated the deficiencies of both MCCA and jICA and could separate sources accurately and link them correctly using less stringent assumptions. Nevertheless, MCCA+jICA is a two-step method. Incorporating source dependency and HOS into a unified framework, such as IVA, may further benefit the source identification. By using IVA-G, the values of ACC were further increased in comparison to those of MCCA+jICA, indicating better source separation. This can be seen in Figure 4(h), in which the third sources within all data sets were more precisely recovered in contrast to Figure 4(g). The

performance of IVA-GL was unsatisfactory as the first sources were contaminated by the second ones [Figure 4(i)], probably due to the incorrect assumption of the Laplace distribution.

Example applications

In this section, we demonstrate how JBSS methods can be used to solve two practical problems in real-world applications: simultaneously handling multiple data sets from the same type of neurophysiological data, and jointly modeling multiple data sets from several distinct types of neurophysiological data. By describing the following representative applications, we demonstrate how JBSS can be used to explore information from multiset, multimodal neurophysiological data and achieve superior performance compared to traditional approaches. In addition, we hope that these examples further inspire researchers to investigate JBSS for other potential neurophysiological applications.

Case studies from the multiset perspective

In the field of neurophysiology, analyzing individual data sets separately and then subsequently integrating the results is the traditional way for multiset data analysis. However, this is suboptimal, since it does not allow for direct interactions among multiple data sets. JBSS provides a straightforward way to jointly analyze multiple sets of the same type of data and exploit the intrinsic data dependency information. In the following sections, three successful applications will be introduced to demonstrate the benefit of using JBSS, including EEG denoising, SSVEP-based BCI, and group fMRI data analysis.

EEG denoising

Following its first application to EEG decomposition [3], ICA has been extensively investigated in a number of EEG denoising applications. ICA has been established as a standard tool for removing several types of artifacts in EEG, such as the ECG and the electrooculogram. However, ICA has been shown inadequate for removing some types of artifacts, such as muscular artifact [50] and gradient artifact when simultaneous EEG/fMRI data are acquired [51]. We introduce recent work that suggests how JBSS can be used to address these issues.

Muscular artifact in EEG recordings

Muscle artifact removal from EEG recordings can be particularly challenging, as muscle electrical activity tends to be of high amplitude, has a wide frequency spectrum, and a potentially broad anatomical distribution over the scalp [50]. Since traditional ICA implementations only exploit spatial information while ignoring temporal structure, it is more suited for isolating artifacts with stereotyped scalp topographies into a single independent component. However, muscle artifacts usually include different muscles activated at different times, and thus they have nonstereotyped scalp topographies, so traditional ICA may not perform effectively in this situation. Recently, CCA, as a JBSS method, has been proposed as a more robust tool for muscle artifact removal in scalp EEG

recordings [37]. Let $\mathbf{X}^{[1]}$ be a multichannel EEG recording and $\mathbf{X}^{[2]}$ be its delayed version. Applying CCA to the two data sets leads to the sources that are maximally autocorrelated and still mutually uncorrelated. Since muscle artifacts broadly distribute in the frequency domain and approach a temporal structure of white noise compared to EEG signals of interest, these artifacts tend to have relatively low autocorrelation. CCA makes use of this distinguishable feature to separate muscle artifacts from the interested EEG signals, and CCA was shown to outperform ICA on simulated data [37]. Further clinical applications also support the usage of CCA because of its performance improvement [52].

Recently, several single-channel techniques have been proposed to remove artifacts in EEG recordings, e.g., EEMD-ICA [53] and EEMD-CCA [31], [32]. These applications are important for situations where minimal instrumentation is available, such as ambulatory healthcare. These single-channel techniques first decompose single-channel data into multichannel data by EEMD, then apply ICA or CCA to the obtained multichannel data, and finally exclude the sources related to artifacts in the reconstruction. Through examinations on both synthetic and real-world data, EEMD-CCA has been demonstrated to outperform EEMD-ICA for muscular artifact removal in EEG recordings [31], [32]. Figure 5 demonstrates an example of

removing muscle artifacts in ictal EEG signals using two different single-channel techniques. Ictal EEG signals are often severely contaminated with muscular artifacts, complicating the localization of the ictal onset—which is of great clinical importance. In Figure 5(a), a 10-second scalp ictal EEG recording from the BioSource database, contaminated with muscular artifacts and eye blinks, is shown (21 channels and 250-Hz sampling rate). Muscular artifacts can be observed between 0–3.9 seconds on channels F7, T3, T5, C3, and T1 and between 5–10 seconds on channels F8, T4, F4, C4, and P4. Figure 5(a) shows the ictal EEG signals after removing muscle artifacts by applying EEMD-CCA to each channel. Note that muscular artifacts are effectively removed compared to the original EEG (black). The ictal activity in each of the T2, F8, T4, and T6 channels is still well preserved. The ictal activity in F8 and T4, which originally was blurred by muscular artifacts, becomes visible after using EEMD-CCA. The decomposition results by applying EEMD-ICA and EEMD-CCA to channel C3 are presented in Figure 5(c) and (d), respectively. While EEMD-CCA is able to isolate muscle activity into the bottom two components (indexes 10 and 11), EEMD-ICA is unable to effectively separate muscular artifacts from brain activity [e.g., as shown in IC9 in Figure 5(c)], demonstrating the advantage of using JBSS in this particular application.

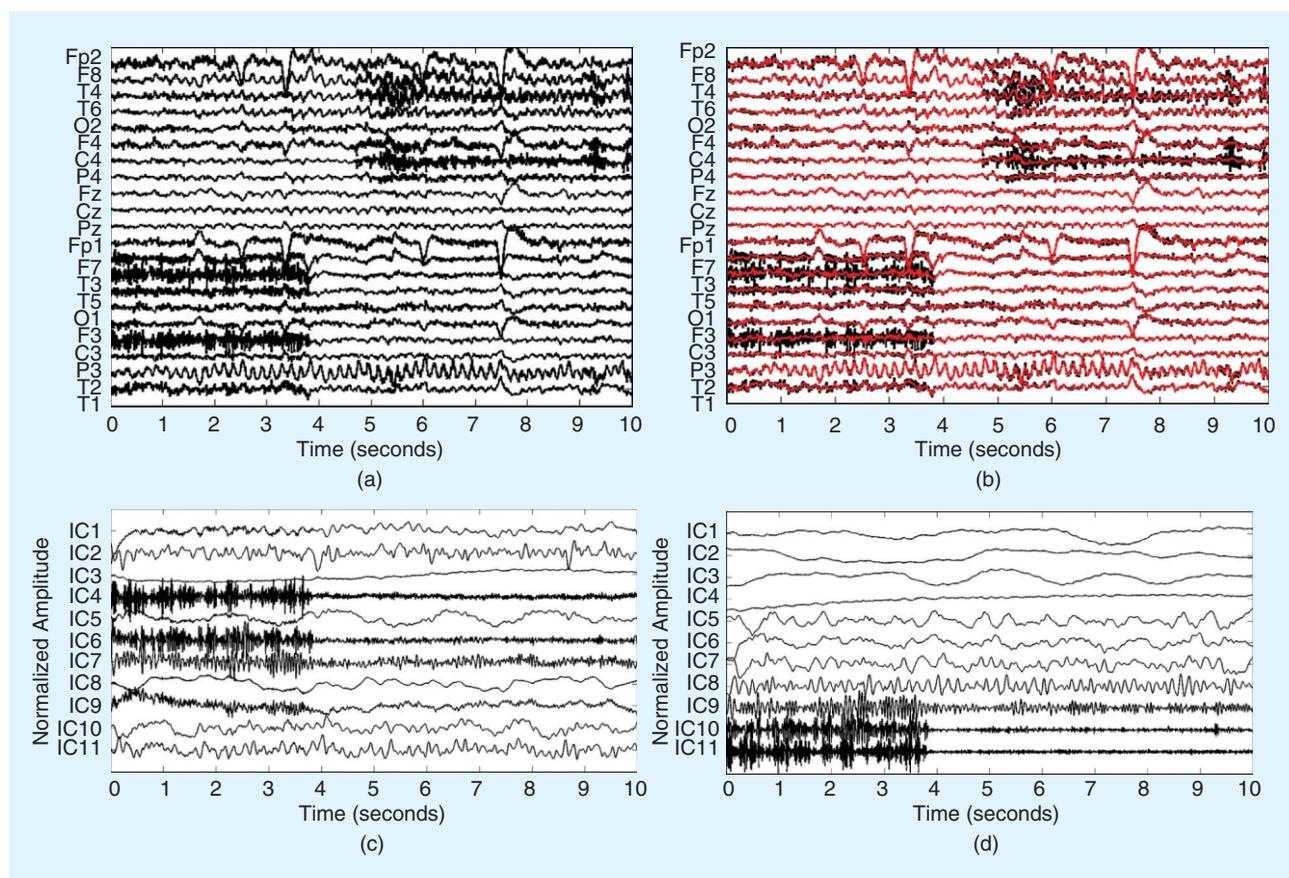


FIGURE 5. (a) An original 10-second scalp ictal EEG recording, (b) cleaned ictal EEG signals using EEMD-CCA (red) compared with the original EEG recordings (black), (c) independent components (ICs) of the Channel-C3 EEG signal obtained using EEMD-ICA, and (d) canonical variates (CVs) of the Channel-C3 EEG signal obtained using EEMD-CCA. (Figure adapted from [32] with permission.)

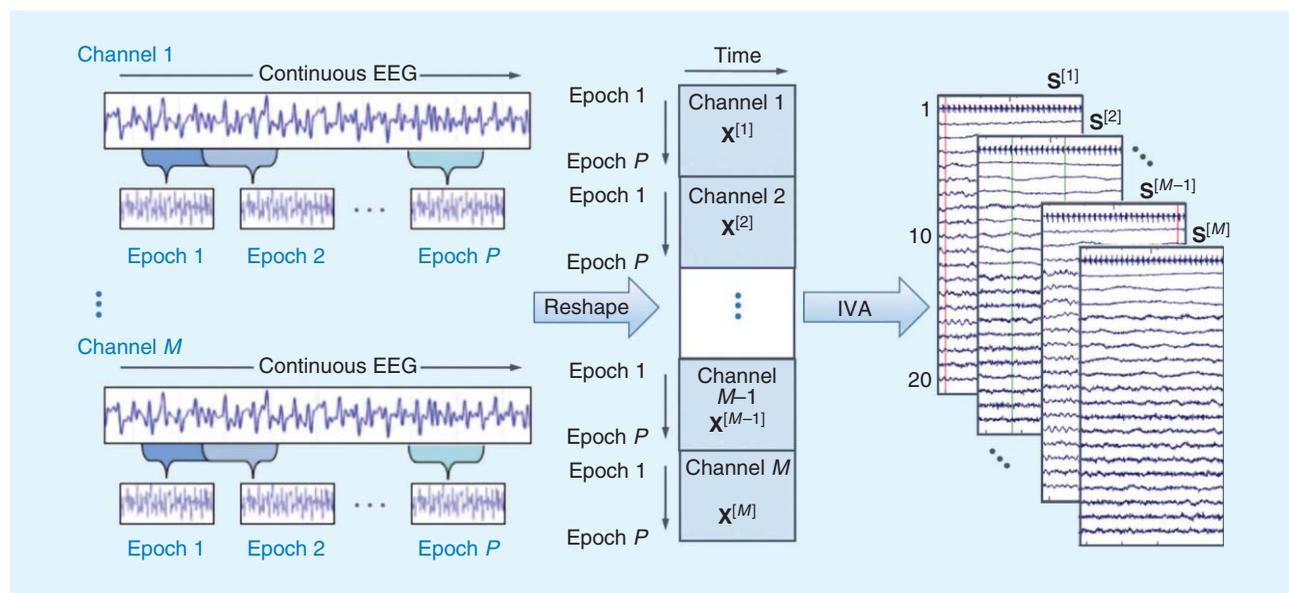


FIGURE 6. An illustration of separating continuous EEG data into epochs for JBSS and the decomposition results of applying IVA. (Figure adapted from [51] with permission.)

Gradient artifact

Combining the advantages of EEG and fMRI for high temporal and spatial resolution assessment of brain function is an attractive approach. During concurrent and continuous EEG-fMRI acquisition, rapidly switching magnetic field gradients and the radio frequency pulses needed to acquire fMRI images results in large artifacts in the EEG signals. A number of methods for gradient artifact removal have been proposed, including average artifact subtraction (AAS), PCA- or ICA-based template subtraction, and adaptive or spatial filtering. Among these methods, AAS is robust and relatively simple. Recently, IVA, a JBSS method, has been proposed as a means to remove gradient artifact and could potentially achieve better performance than that of AAS [51]. Unlike AAS, which just takes the average over epochs to create a gradient artifact template, IVA exploits the dependency of gradient artifacts across channels and reshapes multichannel EEG signals into multiple data sets, as shown in Figure 6. From continuous EEG signals, P (here $P = 20$) epochs are extracted in each single channel, and the duration of each epoch is equal to a repeat time (2 seconds, with a sampling rate of 1,024 Hz). All epochs from a single channel are reshaped into a data set $\mathbf{X}^{[m]}$, and M (here $M = 30$) channels provide M data sets for the IVA algorithm. From each data set, IVA provides $P = 20$ estimated sources with length $N = 2,048$. As the gradient artifact is present in all channels and all epochs of the raw EEG signals, it is anticipated that after applying IVA, the first components (i.e., the first SCV) extracted from the M data sets will correspond to the gradient artifacts contained in the M channels. The first component from each data set (Figure 6) is quasiperiodic, highly suggesting that it is related to gradient artifact, which is confirmed by subsequent frequency analysis. EEG signals with the artifact minimized can be obtained by removing the components related to gradient artifacts during the back

reconstruction. With such an approach, IVA has been demonstrated to consistently outperform AAS on both simulated and real-world EEG/fMRI data [51].

SSVEP-based BCI

BCI is a communication system that translates electrical brain activity (typically measured by EEG) into computer commands, and, hence, it provides a communication channel for severely disabled people. SSVEP-based BCI has been increasingly studied due to its reduced training requirement and higher ultimate information transfer rates compared to those of other BCI techniques. In an SSVEP application, different regions of the computer screen, corresponding to different potential choices for communication purposes, flicker at different frequencies and/or phases. If subjects want to select a particular choice, they direct their attention to the appropriate region on the computer screen. An SSVEP is then detectable over occipital scalp regions at the same frequency as the flicker frequency, and its harmonics are also detectable [54]. SSVEP-based BCI is designed to detect the subject's desired commands by recognizing the SSVEP-induced frequencies in the EEG signals. However, SSVEP responses are likely to be contaminated by background EEG and other noise. Therefore, identifying the frequency components with high accuracy is a particularly challenging and important issue for designing a reliable SSVEP-based BCI, and a number of methods have been proposed for this purpose. The CCA-based recognition method has significantly better recognition performance than traditional power spectral density analysis [55]. By using CCA, correlations are maximized between the multichannel EEG and reference signals (e.g., the sine-cosine waves in [55]) at each of the used stimulus frequencies. The stimulus frequency that yields the maximal correlation coefficient is recognized as the target frequency. Although CCA works well in many

studies in SSVEP-based BCIs, directly using the sine-cosine waves as the reference signals may be problematic in practice due to subject-specific and intertrial effects.

MCCA was introduced as a natural way to extract SSVEP common features from multiple trials of EEG signals recorded at the same stimulus frequency f_r ($r = 1, 2, \dots, R$), as shown in Figure 7(a). For each subject focusing on a particular screen area associated with a specific stimulus frequency, M trials of EEG are collected, which form M data sets. In this particular example, each trial of EEG includes $P = 30$ channels and $N = 250$ time points sampled at 250 Hz. As SSVEP responses relevant to the same f_r are present in all trials of the raw EEG signals, it is anticipated that the first components (i.e., the first SCM in the red box) from applying MCCA to the M data sets more accurately represent reference signals at f_r than the original, pure sine-cosine waves do. Since $M = 19$ data sets are available, the optimized reference signals at each frequency, denoted by each red box, have the dimension 19×250 . Then, given a new EEG trial, the CCA-based method is utilized to recognize the target frequency. Figure 7(b) shows examples of the two types of reference signals, including the sine-cosine waves (top) and the data-driven reference signals optimized by MCCA (bottom). The procedure of the reference signal optimization in the MCCA-based method is completely based on the training data. An experimental study [24] has demonstrated that the MCCA-based method improved the recognition accuracy when compared with previous methods relying on the sine-cosine wave type of reference signals, and thus it provided a new promising tool for frequency recognition in SSVEP-based BCIs.

Group fMRI inference

Detection of activation-related signal changes in fMRI data is a challenging issue due to the relatively low image contrast-to-noise ratio of the blood oxygenation level dependent (BOLD) fMRI signal, head movement, and undesired physiological sources of variability [16]. McKeown et al. first applied ICA to fMRI data analysis and demonstrated that a number of spatially independent sources with specific temporal characteristics were present in fMRI data [2]. ICA has subsequently been successfully employed in a number of fMRI applications. However, ICA can only be performed on each data set separately, and it does not make use of the statistical dependence across data sets. Making inferences on groups of subjects, e.g., control versus patient, is often the goal of fMRI studies. Group ICA has been a popular method for jointly analyzing multisubject fMRI data. Nevertheless, the assumption of the common subspace for all subjects in the group ICA may limit its ability to capture individual subject variability that may be seen in spatial maps.

Unlike group ICA, IVA does not limit the solution space and does not require back-reconstruction for spatial components, but rather it directly estimates a demixing matrix for each data set (i.e., each subject) simultaneously. Figure 8(a) illustrates the detailed procedure of utilizing IVA to extract group-level spatial maps. Recently, IVA has been investigated in studies of the group fMRI analysis [18]. IVA, as a more general JBSS framework, is able to better capture intersubject variability, which is of crucial importance for group comparison studies. Therefore, IVA typically produces

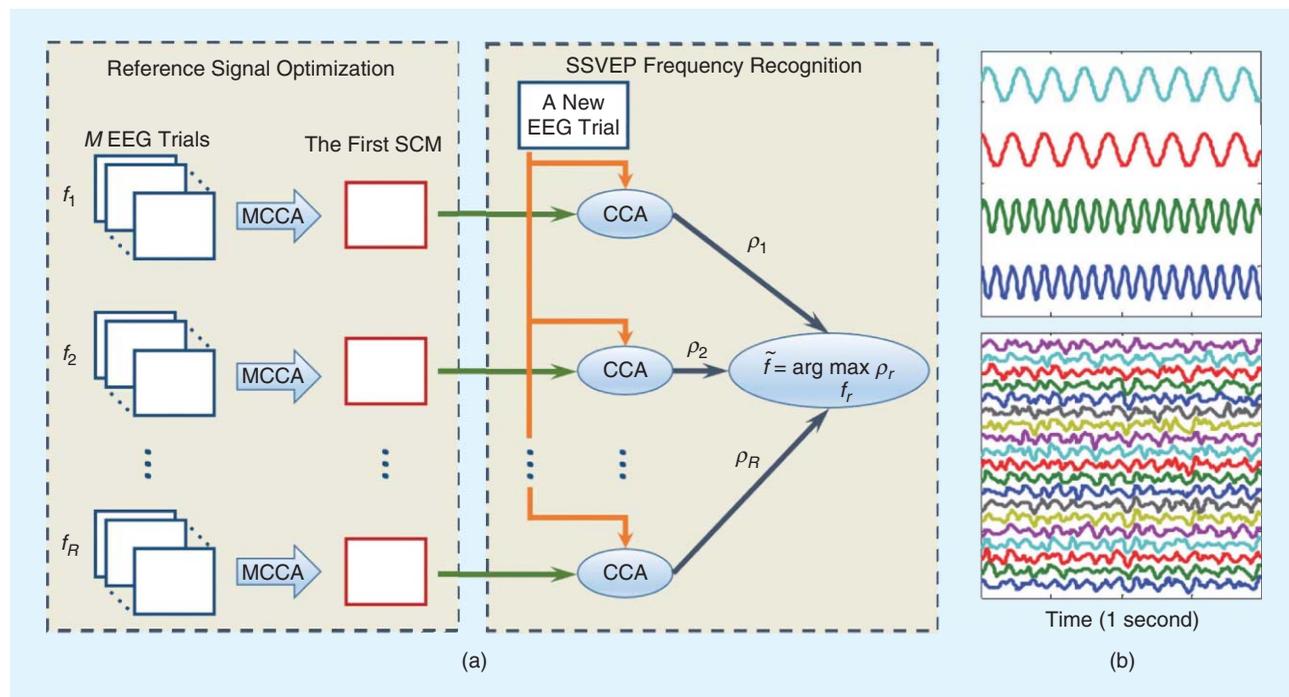


FIGURE 7. (a) An illustration of the MCCA-based method for the SSVEP frequency recognition and (b) examples of the reference signals that are sine-cosine waves (top) and that are optimized by MCCA (bottom). (Figure adapted from [24] with permission.)

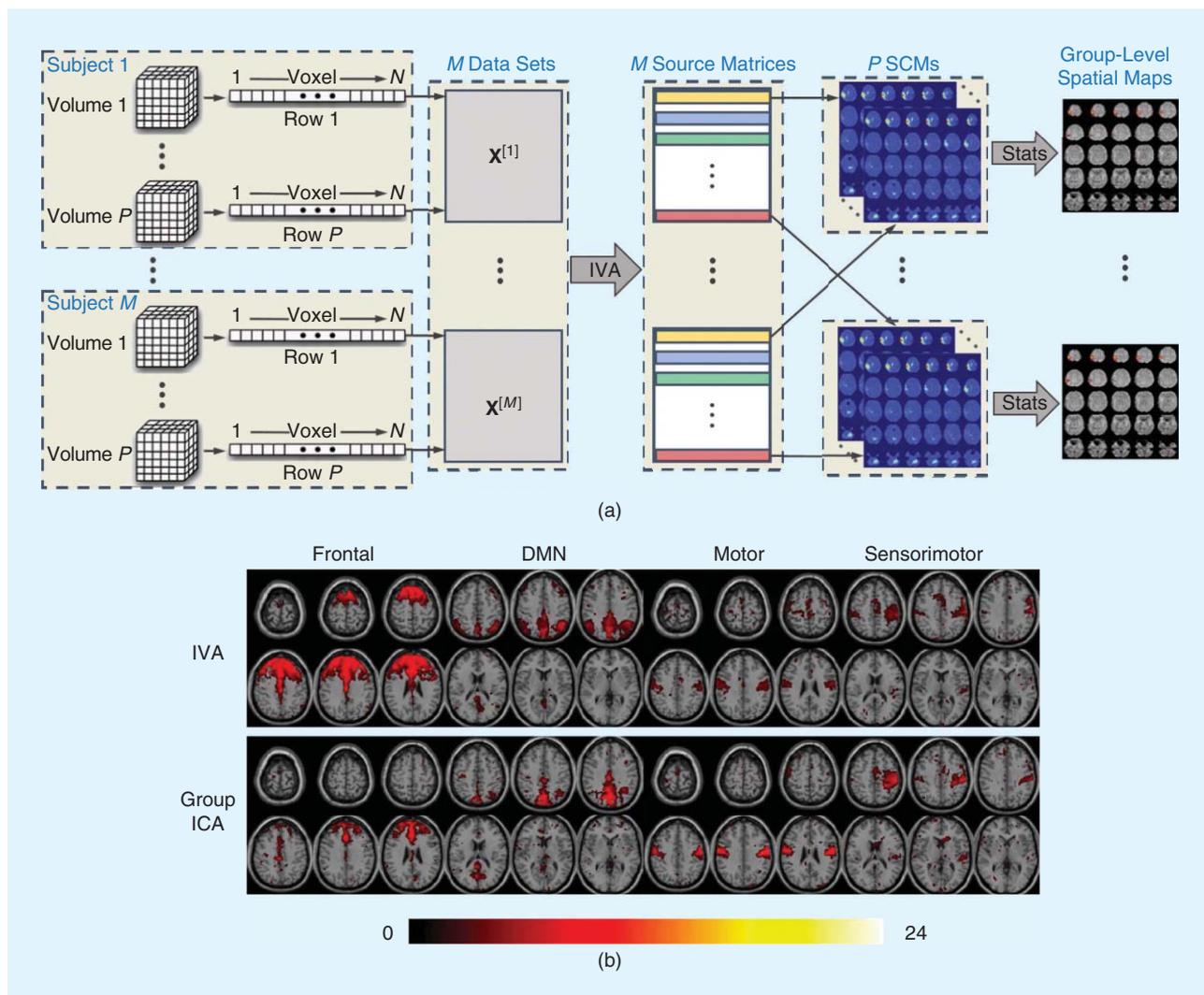


FIGURE 8. (a) An illustration of IVA for multisubject fMRI analysis and (b) examples of estimated spatial maps derived from IVA and the group ICA decompositions. (Figure adapted from [18] with permission.)

components with more activated voxels and better definitions of components, as shown in Figure 8(b) [e.g., the case for the default mode network (DMN)].

Case studies from the multimodal perspective

A current trend in neurophysiological studies is to collect measurements from a variety of modalities, e.g., EEG data, fMRI data, sMRI data, genetic data, and so on. Here we introduce three specific applications to demonstrate the strengths of JBSS in multimodal data fusion, including EEG-fMRI fusion, corticomuscular coupling analysis, and brain imaging genetics.

Joint EEG-fMRI analysis

Given that currently no single brain imaging tool can provide an optimal combination of spatial and temporal resolution, it is particularly attractive to fuse EEG signals (with excellent temporal resolution) and fMRI data (with excellent spatial resolution) when both modalities are concurrently available.

Conventional methods first analyze each modality separately (e.g., via decomposition) and then explore the relationship of the components (e.g., via correlating). A method that is able to simultaneously model the EEG and fMRI data is preferable. Multiway PLS (N-PLS), a JBSS method, provides a natural way to identify EEG spectral/spatial atoms having the maximal temporal covariance with fMRI spatial signatures [35]. Here, each EEG atom is the outer product of spatial, spectral, and temporal signatures, and each fMRI atom is the product of spatial and temporal signatures. To temporally align the EEG data with the fMRI signal, the EEG data recorded during each TR interval are used to estimate an EEG spectrum. The decomposition is constrained to maximize the covariance between corresponding temporal signatures of the EEG and the fMRI data, as shown in Figure 9. Subsequent statistical testing has shown one atom associated with the alpha band having significant temporal correlation with the fMRI signal [35]. Brain regions participating in the control of alpha rhythm included the parietooccipital cortex, thalamus, and insula

inferred from the spatial map of fMRI, while the EEG only included the parietooccipital electrodes. This indicates that joint analysis of EEG and fMRI data meaningfully extends the spatiotemporal resolution and sensitivity of each modality.

Coupling analysis between EEG, EMG, and KIN data
Corticomuscular coupling analysis, i.e., studying simultaneous cortical and muscular activity typically during sustained isometric muscle contraction, is a key technique to assess functional interactions in the motor control system. The most common analysis method to compare the simultaneously measured EEG and EMG signals is the magnitude-squared coherence (MSC), which is a normalized measure of correlation between two signals in the frequency domain. Despite MSC's popularity, several implicit assumptions of MSC potentially limit its usefulness in practical scenarios, including 1) difficulty in robustly assessing group inference, 2) only handling two modalities simultaneously, 3) the biologically implausible assumption of pair-wise interactions (i.e., emphasizing the role of individual loci in the brain), and 4) directly applying the analysis to raw EEG and EMG data, leading to a low coherence value since only a small fraction of ongoing EEG activity is related to motor control.

To address the above limits of MSC, efficient multimodal data-driven methods are needed for corticomuscular coupling analysis. For instance, when studying altered corticomuscular activity in PD, with concurrent EEG, EMG, and KIN data collected from patients with PD and control subjects during a dynamic force tracking task, a more data-driven way of estimating common underlying sources to achieve group-level spatial consensus is required. MCCA+jICA is a method to meet the requirement from the JBSS point of view [48]. As shown in Figure 10, with the assumption that all subjects share common group patterns in the temporal dimension, all subjects' data sets are first correspondingly concatenated along the feature dimension. In this case, the EEG features could be interchannel correlation coefficients, band-limited energy values, and/or coherence values; the EMG feature could be the

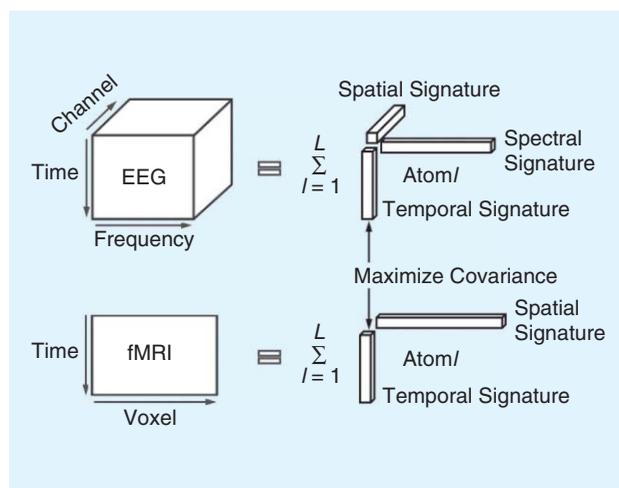


FIGURE 9. An illustration of applying N-PLS to the time-varying EEG spectrum data and the fMRI signals for simultaneously extracting L atoms by constraining the temporal signatures to have maximal covariance. (Figure adapted from [35] with permission.)

amplitude. Then multi-LV extraction is utilized for dimension reduction, and MCCA+jICA is subsequently employed for joint source extraction. The source in the EEG features, which is most correlated with the corresponding sources in EMG and KIN, is then adopted to do regression on the original EEG features for each subject. Finally, statistical tests are applied to the weights in each group for generating spatial patterns [48]. The demonstrated results enhanced occipital connectivity in PD subjects, consistent with the fact that PD subjects rely excessively on visual information to counteract the deficiency of being able to generate internal commands from their impaired basal ganglia [48].

Genetic data fused with brain imaging data

Imaging genetics combines genetic information and neuroimaging data of the same subjects with the aim to discover neural mechanisms of psychiatric disorders. It provides a unique way to investigate genetic influence on the variation of brain

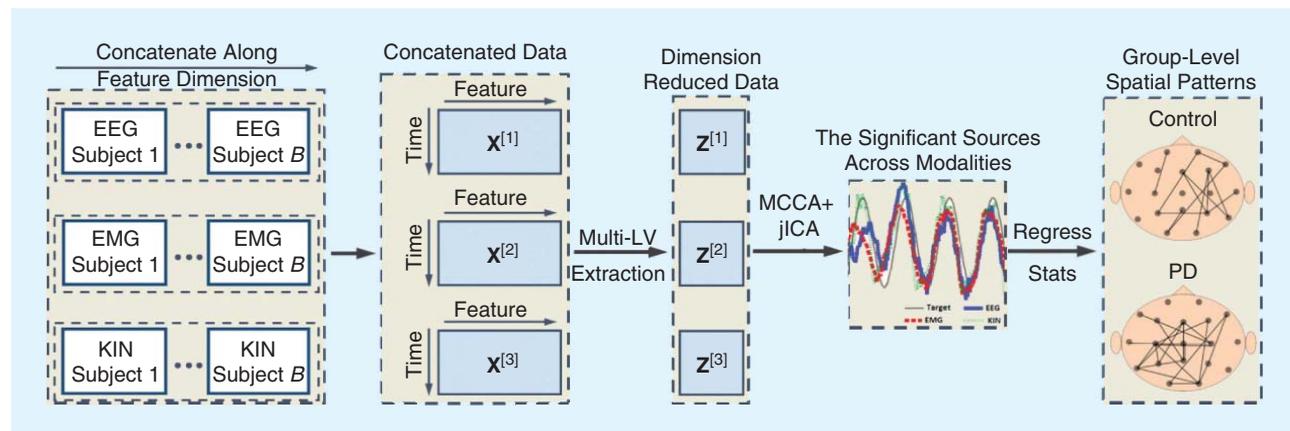


FIGURE 10. An illustration of applying MCCA+jICA to the time-varying features of EEG, EMG, and KIN data for inferring group-level spatial patterns. (Figure adapted from [48] with permission.)

attributes. Traditional methods either do not explicitly incorporate the information provided by all modalities in one comprehensive analysis, or they have not yet been widely applied to more than two modalities [56]. Under the situation that genetic (e.g., SNP), structural (e.g., sMRI), and functional (e.g., fMRI) data are all available, the need for leveraging valuable information from those complementary modalities is evident. Recently, three-way pICA has been introduced as a useful JBSS tool in identifying pairwise links between modalities and estimating associated independent components within each modality [49]. A recent study investigated genetic effects on alcohol dependence [49], and, as illustrated in Figure 11, subjects with alcohol dependence underwent functional and sMRI scans and also provided saliva for SNP generation. The three types of data were concatenated along the subject dimension. Then three-way pICA was applied to the sMRI, fMRI, and SNP data, extracting intersubject covariation patterns (i.e., modulation profiles) across the three modalities and their associated independent components. Through statistical tests, significant links between a SNP component, a functional component, and a gray matter component were identified: the SNP component implicated mental disorder associated genes including *BDNF*, *GRIN2B*, and *NRG1*; the functional component demonstrated increased activation in the precuneus area; and the structural component involved part of the DMN and the caudate.

Discussion

In this article, four categories of JBSS methods (i.e., [SOS versus HOS] \times [biset/bimodal versus multiset/multimodal]) have been introduced in terms of their motivations, formulations, characteristics, and feasible applications, and they have been summarized in Table 1. For a given neurophysiological application, since a rich set of JBSS methods is available,

several key points should be taken into account for choosing an appropriate JBSS method. It is straightforward to choose bimodal versus multimodal categories by counting the number of data sets that need to be jointly analyzed, i.e., determining $M = 2$ or $M > 2$. In most cases, the multi-set/multimodal methods can still be used if $M = 2$ (e.g., jICA and IVA). Nevertheless, the group ICA may fail to function well if the number M is too low to achieve statistically significant results. Depending on the statistical properties of the specific data of interest, one needs to determine what type of statistics to use, e.g., choosing SOS or HOS. In practice, if the data follow Gaussian distributions, exploiting SOS is sufficient to recover the underlying components and is also more computationally efficient (e.g., MCCA). Otherwise, HOS is generally preferred for ultimate source separation.

When SOS are employed, one needs to consider which measure, correlation or covariance, should be used. This depends on the problem itself: PLS and its variants exploit covariance information, while CCA-based methods and JDIAG-SOS exploit correlation information. JDIAG-SOS is more generalized than MCCA, since the former is based on the joint diagonalization of multiple cumulant matrices while the latter solves ad hoc cost functions by a deflationary procedure; but they both impose orthogonality constraints on demixing matrices, which may limit the solution space examined. When HOS are utilized, one needs to consider the assumptions of the different methods. Group ICA and jICA simply use temporal and spatial concatenations by assuming a common source space and the same mixing matrix respectively. Linked ICA is able to incorporate multiple modalities with similar spatial properties into one modality group and use tensor decomposition by assuming the same mixing matrix across different modality groups; modalities in the same group share the same source space.

Imaging genetics combines genetic information and neuroimaging data of the same subjects with the aim to discover neural mechanisms of psychiatric disorders.

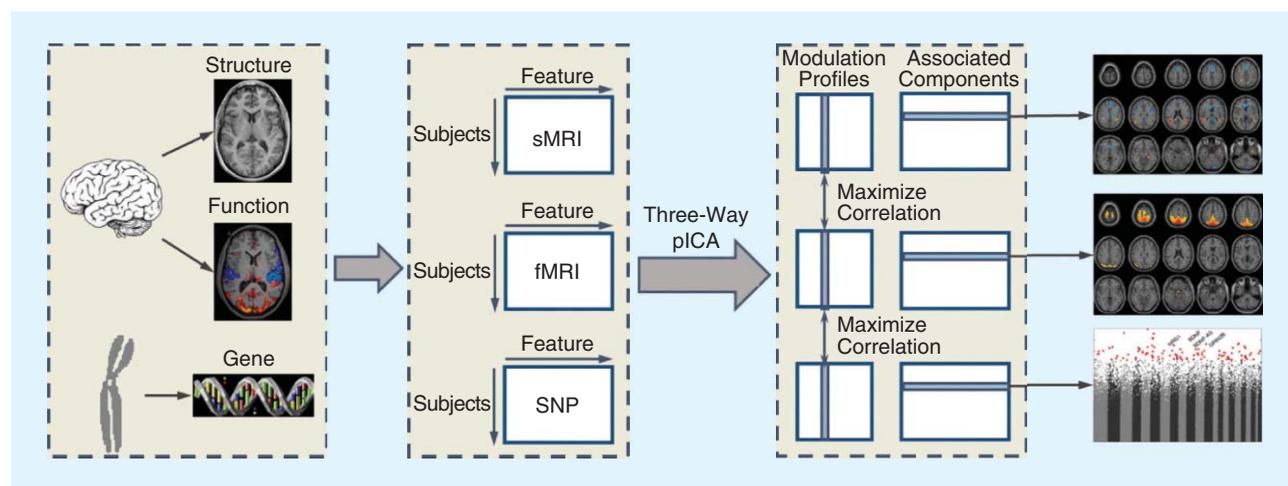


FIGURE 11. An illustration of employing the three-way pICA to analyze links between the genetic, brain structural, and brain functional data. (Figure adapted from [49] with permission.)

Thus, linked ICA is a method encapsulating the ideas from both group ICA and jICA. MCCA+jICA may relax the stringent assumption of jICA and lead to improved source separation performance. IC-PLS considers both independence and covariance aspects and focuses on source extraction. pICA flexibly identifies both independent sources and their links across modalities. IVA, using mutual information rate minimization, extracts independent sources and maximally correlates them across multiple data sets. IC-PLS, pICA, and IVA do not have any stringent assumption for source spaces and mixing matrices, indicating that each data set has its unique mixing and source matrices.

A number of recent studies demonstrate that IVA is superior to others for JBSS as formulated in Figure 2(a) [17], [18]. Compared with CCA, PLS, and IC-PLS, IVA is able to jointly model more than two data sets. In contrast to MCCA and JDIAG-SOS, IVA incorporates statistics higher than two and also allows a general nonorthogonal demixing matrix. Unlike other ICA-based JBSS methods, IVA does not constrain mixing or source matrices to be the same. However, just as ICA and the methods using ICA, IVA does need a predefined probability distribution such as IVA-G using Gaussian distribution and IVA-L employing Laplace distribution. IVA-G takes into account second-order correlation across data sets while IVA-L does not. In medical applications, the source dependency in second-order is reasonable. It has been previously shown that for non-Gaussian sources, IVA-G can be superior to identifying sources in each data set independently using a BSS algorithm followed by an optimal source alignment algorithm. This is a key motivation for JBSS and the general IVA formulation. IVA-G can even outperform IVA-L for non-Gaussian sources as long as sufficient second-order dependency across data sets exists [17]. Lately, IVA based on multivariate generalized Gaussian distribution has also been developed and can fit more general situations [57]. It is worth emphasizing that while most JBSS methods take HOS and source dependence across data sets into account, they do not take sample dependence into account. To our knowledge, methods recently proposed in [44], [72] are the only solutions that exploit sample dependence in addition to HOS and source dependence. Finally, the unique information contained in each individual data set may be important. On one hand, under the JBSS framework, after the common information has been extracted, the unique information in each data set can still be explored by using the methods such as orthogonal signal correction [27], [74]; on the other hand, JBSS methods have the ability to capture the variability of the source components within the same SCV, indicating that though the source components within the same SCV are dependent, they possess some unique information [18].

No single study has investigated all possible JBSS methods and demonstrated their strength and weakness

A current trend in neurophysiological studies is to collect measurements from a variety of modalities.

thoroughly for one specific neurophysiological problem. Most studies have only applied one or two methods to real-world data and obtained preliminary results. Inferring the validity of such results is difficult, as often the “ground truth” is not currently known and must be indirectly inferred. This emphasizes the need for strengthening the collaboration between clinical neurophysiologists and researchers in biomedical engineering and signal processing, particularly in the era of “big data” [5]. There also exist several challenges for the development of JBSS methods in the field of neurophysiology.

Current available JBSS methods generally assume given data sets are temporally stationary. However, most neurophysiological signals are known to be highly nonstationary, especially for brain imaging techniques with high temporal resolution, such as EEG recordings. Aforementioned JBSS methods assume that the number of underlying sources in each data set is equal to or less than that of observations (i.e., $L \leq P_m$), actually leading to (over)determined problems. Nevertheless, most real-world JBSS problems are in fact underdetermined, especially in this era of ambulatory healthcare systems, which utilize minimal instrumentation (e.g., even single-channel). In current JBSS frameworks, only latent components and their dependency across data sets have been considered, while identifying latent subspaces and finding the relations among subspaces across data sets may provide novel neurophysiological interpretations with different potential applications [58], [59]. While this review article has been largely focused on statistical JBSS methods in the field of neurophysiology, we note that there are deterministic approaches in other fields (e.g., speech recognition [73]).

In conclusion, JBSS is a promising research direction to combine multiple sets of neurophysiological data from the same modality or multiple modalities. It has been successfully employed in many medical applications and helped medical researchers derive interesting results, which are otherwise unobtainable using traditional BSS methods. This article has provided a comprehensive overview for these JBSS methods and related applications. It is expected that JBSS will continue to draw increasing attention from the neurophysiological community and will be widely applied and further developed for solving novel real-world problems.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants 81571760 and 61501164), the Research and Development Foundation of Hefei Research Institutes (Grant IMICZ2015111), the Natural Sciences and Engineering Research Council of Canada, and the Canadian Institutes of Health Research.

Authors

Xun Chen (xun.chen@hfut.edu.cn) received his B.S. degree in electrical engineering from the University of Science and

Technology of China in 2009 and his Ph.D. degree in biomedical engineering from the University of British Columbia, Canada, in 2014, where he was a research scientist in the Department of Electrical and Computer Engineering. Currently, he is with the Intelligent Manufacturing Institute and the Department of Biomedical Engineering at the Hefei University of Technology, China, as a full professor. His research interests include the broad areas of statistical signal processing and machine learning in biomedical applications.

Z. Jane Wang (zjanew@ece.ubc.ca) received her B.S. degree from Tsinghua University, China, in 1996, and her M.S. and Ph.D. degrees from the University of Connecticut in 2000 and 2002, respectively, all in electrical engineering. She was a research associate in the Electrical and Computer Engineering Department at the University of Maryland, College Park. Since 2004, she has been with the Department Electrical and Computer Engineering at the University of British Columbia, Canada, and is currently a full professor. Her research interests include statistical signal processing theory and applications with a current focus on brain data analytics. She has published approximately 80 journal articles and more than 90 refereed conference papers.

Martin J. McKeown (martin.mckeown@ubc.ca) is the Pacific Parkinson's Research Institute and University of British Columbia chair in Parkinson's research, director at the Pacific Parkinson's Research Centre, full professor in the Department of Medicine, and adjunct professor in the Department of Electrical and Computer Engineering at the University of British Columbia, Canada. He completed his engineering physics, medicine, and neurology training at McMaster University, the University of Toronto, and the University of Western Ontario, respectively. He is a member of the Systems and Clinical Neurosciences—A Canadian Institutes of Health Research scientific peer review committee as well as a member of the Scientific Advisory Board of the Parkinson's Society of Canada. He has authored more than 100 peer-reviewed papers and book chapters.

References

- [1] R. Vigario and E. Oja, "BSS and ICA in neuroinformatics: from current practices to open challenges," *IEEE Rev. Biomed. Eng.*, vol. 1, pp. 50–61, Nov. 2008.
- [2] M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," *Hum. Brain Mapp.*, vol. 6, no. 3, pp. 160–188, Dec. 1998.
- [3] T. P. Jung, S. Makeig, C. Humphries, T. W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, Mar. 2000.
- [4] K. Uludağ and A. Roebroeck, "General overview on the merits of multimodal neuroimaging data fusion," *Neuroimage*, vol. 102, pp. 3–10, Nov. 2014.
- [5] T. J. Sejnowski, P. S. Churchland, and J. A. Movshon, "Putting big data to good use in neuroscience," *Nat. Neurosci.*, vol. 17, no. 11, pp. 1440–1441, Nov. 2014.
- [6] F. Biessmann, S. Plis, F. C. Meinecke, T. Eichele, and K. Müller, "Analysis of multimodal neuroimaging data," *IEEE Rev. Biomed. Eng.*, vol. 4, pp. 26–58, Oct. 2011.
- [7] J. R. Ives, S. Warach, F. Schmitt, R. R. Edelman, and D. L. Schomer, "Monitoring the patient's EEG during echo planar MRI," *Electroencephalogr. Clin. Neurophysiol.*, vol. 87, no. 6, pp. 417–420, Dec. 1993.
- [8] R. I. Goldman, J. M. Stern, J. Engel, and M. S. Cohen, "Acquiring simultaneous EEG and functional MRI," *Clin. Neurophysiol.*, vol. 111, no. 11, pp. 1974–1980, Nov. 2000.
- [9] P. A. Valdes-Sosa, J. M. Sanchez, R. C. Bornot, Y. Sotero, Y. Iturria-Medina, J. Aleman-Gomez, F. Bosch-Bayard, and T. Carbonell, Ozaki, "Model driven EEG/fMRI fusion of brain oscillations," *Hum. Brain Mapp.*, vol. 30, no. 9, pp. 2701–2721, Sept. 2009.
- [10] S. M. Plis, V. D. Calhoun, M. P. Weisend, T. Eichele, and T. Lane, "MEG and fMRI fusion for non-linear estimation of neural and BOLD signal changes," *Front. Neuroinform.*, vol. 4, pp. 114–114, Nov. 2010.
- [11] B. Horwitz and D. Poeppel, "How can EEG/MEG and fMRI/PET data be combined?" *Hum. Brain Mapp.*, vol. 17, no. 1, pp. 1–3, Sept. 2002.
- [12] H. Yang, J. Liu, J. Sui, G. Pearlson, and V. D. Calhoun, "A hybrid machine learning method for fusing fMRI and genetic data: Combining both improves classification of schizophrenia," *Front. Hum. Neurosci.*, vol. 4, pp. 192, Oct. 2010.
- [13] J. Sui, T. Adali, Q. Yu, J. Chen, and V. D. Calhoun, "A review of multivariate methods for multimodal fusion of brain imaging data," *J. Neurosci. Methods*, vol. 204, no. 1, pp. 68–81, Feb. 2012.
- [14] V. D. Calhoun, J. Liu, and T. Adali, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," *Neuroimage*, vol. 45, no. 1, pp. S163–S172, Mar. 2009.
- [15] T. Kim, T. Eltoft, T. W. Lee, "Independent vector analysis: An extension of ICA to multivariate components," *Independent Component Analysis and Blind Signal Separation*, vol. 3889, J. Rosca, D. Erdogmus, J. C. Principe, S. Haykin, Eds. Berlin: Springer-Heidelberg, 2006, pp. 165–172.
- [16] T. Adali, M. Anderson, and G. Fu, "Diversity in independent component and vector analyses: Identifiability, algorithms, and applications in medical imaging," *IEEE Signal Processing Mag.*, vol. 31, no. 3, pp. 18–33, May 2014.
- [17] M. Anderson, T. Adali, and X. L. Li, "Joint blind source separation with multivariate Gaussian model: algorithms and performance analysis," *IEEE Trans. Signal Processing*, vol. 60, no. 4, pp. 1672–1683, Apr. 2012.
- [18] J. Laney, K. Westlake, S. Ma, E. Woytowicz, V. D. Calhoun, and T. Adali, "Capturing subject variability in fMRI data: A graph-theoretical analysis of GICA vs. IVA," *J. Neurosci. Methods*, vol. 247, pp. 32–40, May 2015.
- [19] X. Chen, A. Liu, M. J. McKeown, H. Poizner, and Z. J. Wang, "An EEMD-IVA Framework for concurrent multidimensional EEG and unidimensional kinematic data analysis," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 7, pp. 2187–2198, Apr. 2014.
- [20] J. R. Kettenring, "Canonical analysis of several sets of variables," *Biometrika*, vol. 58, no. 3, pp. 433–451, Dec. 1971.
- [21] P. Geladi and B. Kowalski, "Partial least-squares regression: A tutorial," *Analytica Chimica Acta*, vol. 185, pp. 1–17, 1986.
- [22] Y. Li, T. Adali, W. Wang, and V. Calhoun, "Joint blind source separation by multiset canonical correlation analysis," *IEEE Trans. Signal Processing*, vol. 57, no. 10, pp. 3918–3929, Apr. 2009.
- [23] N. M. Correa, T. Adali, Y. Li, and V. D. Calhoun, "Canonical correlation analysis for data fusion and group inferences: Examining applications of medical imaging data," *IEEE Signal Processing Mag.*, vol. 27, no. 4, pp. 39–50, Aug. 2010.
- [24] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in ssvp-based bci using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 4, June 2014.
- [25] X. Chen, C. He, and H. Peng, "Removal of muscle artifacts from single-channel EEG based on ensemble empirical mode decomposition and multiset canonical correlation analysis," *J. Appl. Mathematics*, vol. 2014, Mar. 2014.
- [26] V. D. Calhoun, T. Adali, K. A. Kiehl, R. Astur, J. J. Pekar, and G. D. Pearlson, "A method for multitask fMRI data fusion applied to schizophrenia," *Hum. Brain Mapp.*, vol. 27, no. 7, pp. 598–610, July 2006.
- [27] X. Chen, X. Chen, R. K. Ward, and Z. J. Wang, "A joint multimodal group analysis framework for modeling corticomuscular activity," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1049–1059, Feb. 2013.
- [28] X. Chen, C. He, Z. J. Wang, and M. J. McKeown, "An IC-PLS Framework for Group Corticomuscular Coupling Analysis," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 2022–2033, July 2013.
- [29] X. Chen, A. Liu, Z. J. Wang, and H. Peng, "Corticomuscular activity modeling by combining partial least squares and canonical correlation analysis," *J. Appl. Mathematics*, vol. 2013, Mar. 2013.
- [30] B. B. Avants, P. A. Cook, L. Ungar, J. C. Gee, and M. Grossman, "Dementia induces correlated reductions in white matter integrity and cortical thickness: A multivariate neuroimaging study with sparse canonical correlation analysis," *Neuroimage*, vol. 50, no. 3, pp. 1004–1016, Apr. 2010.

- [31] K. T. Sweeney, S. F. McLoone, and T. E. Ward, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 97–105, Jan. 2013.
- [32] X. Chen, A. Liu, H. Peng, and R. K. Ward, "A preliminary study of muscular artifact cancellation in single-channel EEG," *Sensors*, vol. 14, no. 10, pp. 18370–18389, Oct. 2014.
- [33] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "L1-regularized multiway canonical correlation analysis for ssvep-based bci," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 6, pp. 887–896, Nov. 2013.
- [34] A. McIntosh, F. Bookstein, J. Haxby, and C. Grady, "Spatial pattern analysis of functional brain images using partial least squares," *Neuroimage*, vol. 3, no. 3, pp. 143–157, June 1996.
- [35] E. Martínez-Montes, P. A. Valdés-Sosa, F. Miwakeichi, R. I. Goldman, and M. S. Cohen, "Concurrent EEG/fMRI analysis by multiway partial least squares," *Neuroimage*, vol. 22, no. 3, pp. 1023–1034, July 2004.
- [36] J. Chiang, Z. J. Wang, and M. J. McKeown, "A multiblock PLS model of cortico-cortical and corticomuscular interactions in Parkinson's disease," *Neuroimage*, vol. 63, no. 3, pp. 1498–1509, Nov. 2012.
- [37] W. D. Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2583–2587, Dec. 2006.
- [38] F. Lindgren, P. Geladi, and S. Wold, "The kernel algorithm for PLS," *J. Chemometrics*, vol. 7, no. 1, pp. 45–59, Jan./Feb. 1993.
- [39] M. G. Gustafsson, "Independent component analysis yields chemically interpretable latent variables in multivariate regression," *J. Chem. Inf. Model.*, vol. 45, no. 5, pp. 1244–1255, Sept.–Oct. 2005.
- [40] J. Liu, V. D. Calhoun, "Parallel independent component analysis for multimodal analysis: application to fMRI and EEG data," in *Proc. IEEE Int. Symp. Biomedical Imaging Conf.*, Arlington, VA, 2007, pp. 1028–1031, 2007.
- [41] J. Liu, G. Pearlson, A. Windemuth, G. Ruano, N. I. Perrone-Bizzozero, and V. D. Calhoun, "Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA," *Hum. Brain Mapp.*, vol. 30, no. 1, pp. 241–255, Jan. 2009.
- [42] J. Liu, K. A. Kiehl, G. Pearlson, N. I. Perrone-Bizzozero, T. Eichele, and V. D. Calhoun, "Genetic determinants of target and novelty-related event-related potentials in the auditory oddball response," *Neuroimage*, vol. 46, no. 3, pp. 809–816, July 2009.
- [43] N. M. Correa, T. Eichele, T. Adali, Y. Li, and V. D. Calhoun, "Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI," *Neuroimage*, vol. 50, no. 4, pp. 1438–1445, Jan. 2010.
- [44] X. Li, T. Adali, and M. Anderson, "Joint blind source separation by generalized joint diagonalization of cumulant matrices," *Signal Processing*, vol. 91, no. 10, pp. 2314–2322, Oct. 2011.
- [45] N. M. Correa, Y. Li, T. Adali, V. D. Calhoun, "Fusion of fMRI, sMRI, and EEG data using canonical correlation analysis," *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, 2009, pp. 385–388.
- [46] A. R. Groves, C. F. Beckmann, S. M. Smith, and M. W. Woolrich, "Linked independent component analysis for multimodal data fusion," *Neuroimage*, vol. 54, no. 3, pp. 2198–2217, Feb. 2011.
- [47] J. Sui, G. Pearlson, A. Caprihan, T. Adali, K. A. Kiehl, J. Liu, J. Yamamoto, and V. D. Calhoun, "Discriminating schizophrenia and bipolar disorder by fusing fMRI and DTI in a multimodal CCA+ joint ICA model," *Neuroimage*, vol. 57, no. 3, pp. 839–855, Aug. 2011.
- [48] X. Chen, Z. J. Wang, and M. J. McKeown, "A three-step multimodal analysis framework for modeling corticomuscular activity with application to Parkinson's disease," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 4, pp. 1232–1241, July 2014.
- [49] V. M. Vergara, A. Ulloa, V. D. Calhoun, D. Boutte, J. Chen, and J. Liu, "A three-way parallel ICA approach to analyze links among genetics, brain structure and brain function," *Neuroimage*, vol. 98, pp. 386–394, Sept. 2014.
- [50] B. W. McMenamin, A. J. Shackman, L. L. Greischar, and R. J. Davidson, "Electromyogenic artifacts and electroencephalographic inferences revisited," *Neuroimage*, vol. 54, no. 1, pp. 4–9, Jan. 2011.
- [51] P. P. Acharjee, R. Phlypo, L. Wu, V. D. Calhoun, and T. Adali, "Independent vector analysis for gradient artifact removal in concurrent EEG-fMRI data," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 7, pp. 1750–1758, July 2015.
- [52] A. Vergult, W. D. Clercq, A. Palmi, B. Vanrumste, P. Dupont, S. Van Huffel, and W. Van Paesschen, "Improving the interpretation of ictal scalp EEG: BSS-CCA algorithm for muscle artifact removal," *Epilepsia*, vol. 48, no. 5, pp. 950–958, Mar. 2007.
- [53] B. Mijović, M. D. Vos, I. Gligorijević, J. Taelman, and S. V. Huffel, "Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 9, pp. 2188–2196, June 2010.
- [54] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002.
- [55] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2610–2614, June 2006.
- [56] J. Liu and V. D. Calhoun, "A review of multi variate analyses in imaging genetics," *Front. Neuroinform.*, vol. 8, pp. 1–11, Mar. 2014.
- [57] Z. Boukouvalas, G. S. Fu, T. Adali, "An efficient multivariate generalized Gaussian distribution estimator: Application to IVA," in *49th Annu. Conf. Information Sciences and Systems (CISS)*, Baltimore, MD, 2015, pp. 1–4.
- [58] R. F. Silva, S. M. Plis, T. Adali, V. D. Calhoun, "Multidataset independent subspace analysis extends independent vector analysis," in *IEEE Int. Conf. Image Processing (ICIP)*, Paris, France, 2014, pp. 2864–2868.
- [59] D. Lahat and C. Jutten. (2015, Jan.). Joint Independent Subspace Analysis Using Second-Order Statistics. HAL. France. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01132297v3>
- [60] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, Dec. 1936.
- [61] A. Hyvärinen and E. Oja, "Fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, no. 7, pp. 1483–1492, Oct. 1997.
- [62] D. Emge, F. Vialatte, G. Dreyfus, T. Adali, "Independent vector analysis for SSVEP signal enhancement," in *Proc. Conf. Information Sciences and Systems*, Baltimore, MD, 2015.
- [63] O. Alter, P. O. Brown, and D. Botstein, "Generalized singular value decomposition for comparative analysis of genome-scale expression data sets of two different organisms," *Proc. Natl. Acad. Sci. U S A*, vol. 100, no. 6, pp. 3351–3356, Mar. 2003.
- [64] S. P. Ponnampalli, M. A. Saunders, C. F. Van Loan, and O. Alter, "A higher-order generalized singular value decomposition for comparison of global mRNA expression from multiple organisms," *PLoS One*, vol. 6, no. 12, pp. e28072, Dec. 2011.
- [65] V. Tsatsishvili, F. Cong, P. Toivaiainen, T. Ristaniemi, "Combining PCA and multiset CCA for dimension reduction when group ICA is applied to decompose naturalistic fMRI data," in *Int. Joint Conf. Neural Networks (IJCNN)*, Anchorage, Alaska, 2015, pp. 1–6.
- [66] N. J. Roseveare and P. J. Schreier, "Model-order selection for analyzing correlation between two data sets using CCA with PCA preprocessing," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Brisbane, Australia, 2015, pp. 5684–5687.
- [67] T. Lofstedt and J. Trygg, "OnPLS—a novel multiblock method for the modeling of predictive and orthogonal variation," *J. Chemometrics*, vol. 25, no. 8, pp. 441–455, Aug. 2011.
- [68] T. Adali, Y. Levin-Schwartz, and V. D. Calhoun, "Multimodal data fusion using source separation: Two effective models based on ICA and IVA and their properties," *Proc. IEEE*, vol. 103, no. 9, pp. 1478–1493, Sept. 2015.
- [69] V. D. Calhoun, T. Adali, J. J. Pekar, and G. D. Pearlson, "A method for making group inferences from functional MRI data using independent component analysis," *Hum. Brain Mapp.*, vol. 14, no. 2, pp. 140–151, Nov. 2001.
- [70] V. D. Calhoun and T. Adali, "Multi subject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery," *IEEE Rev. Biomed. Eng.*, vol. 5, pp. 60–73, 2012.
- [71] Y. Levin-Schwartz, V. D. Calhoun, and T. Adali, "Data-driven fusion of EEG, functional and structural MRI: A comparison of two models," in *48th Annu. Conf. Information Sciences and Systems (CISS)*, Princeton, NJ, 2014, pp. 1–6.
- [72] G. S. Fu, R. Phlypo, M. Anderson, X. L. Li, and T. Adali, "Blind source separation by entropy rate minimization," *IEEE Trans. Signal Processing*, vol. 62, no. 16, pp. 4245–4255, June 2014.
- [73] K. Nathwani and R. M. Hegde, "Joint source separation and dereverberation using constrained spectral divergence optimization," *Signal Processing*, vol. 106, pp. 266–281, Jan. 2015.
- [74] T. Fearn, "On orthogonal signal correction," *Chemometrics and Intelligent Laboratory Syst.*, vol. 50, no. 1, pp. 47–52, Jan. 2000.
- [75] M. J. McKeown and G. M. Peavy, "Biomarkers in Parkinson disease: It's time to combine," *Neurology*, vol. 84, no. 24, pp. 2392–2393, June 2015.

TIPS & TRICKS

David Shiung, Ya-Yin Yang, and Chu-Sing Yang

Improving FIR Filters by Using Cascade Techniques

High-performance filtering is the holy grail for all digital signal processing (DSP) practitioners. However, high-performance filtering is almost synonymous with high implementation complexity and, thus, high cost. This article focuses on designing high-performance finite impulse response (FIR) filters with less complexity. We present novel methods for improving the frequency response of an FIR filter by cascading it with complementary comb filters (CCFs). In particular, we redesign a low-order FIR filter by intentionally inducing a ripple at the passband edge and using CCFs of various lengths to compensate for the ripple. We use the zeros of CCFs to reshape the frequency response of the low-order FIR filter, subsequently called the *prototype filter*. Thus, the composite filter approaches an ideal lowpass filter (LPF) more closely in both the transition band and stopband. Because a CCF contains only one adder and a few storage units, the cost is minimal. The composite filter also maintains a linear phase response because all CCFs have linear phase responses. Through shifting the LPF response to high-frequency bands, we obtain highpass filters and bandpass filters (BPFs) by using the same methodology.

State-of-the-art design

FIR filters can achieve linear phase responses and preserve the envelope of input signals in contrast to infinite impulse response (IIR) filters; however, their applications are often restricted by high computation complexity. An FIR filter

block diagram with five filter taps is shown in Figure 1. The filter output sequence $y(n)$ can be related with the filter input sequence $x(n)$ by using

$$y(n) = \sum_{m=0}^4 h(m)x(n-m). \quad (1)$$

The filter coefficients to be designed are denoted by $h(m)$, $m = 0, \dots, 4$. Numerous FIR filter design techniques exist, such as the window design [1], Parks–McClellan design [2], minimum-phase equiripple design [3], and interpolated equiripple design [3]. However, all of these designs approach an ideal LPF in various ways and are generally tradeoffs among the following design targets: 1) flatness or ripple(s) at the passband, 2) sharpness at the transition band, and 3) suppression at the stopband. However, an ideal LPF can almost always be approached if implementation complexity is not a concern. This indicates that, for a fixed filter tap number, when flatness is gained at the passband, the other two design targets are lost. In other words, a new design that surpasses the described state-of-the-art designs in all three design targets is highly desired. This article proposes an answer to this problem.

Cascaded filters

Improving FIR filters by using simple cascaded structures can approach an ideal filter in terms of the following design considerations: 1) flatness at the passband, 2) sharpness at the transition band, and 3) suppression at the stopband. The composite FIR filter consists of one prototype filter and one shaping filter, which is composed of cascaded CCFs of various lengths. As shown in Figure 2, the resultant transfer function is expressed as follows [1]:

$$H_{\text{cmp}}(z) = H_{\text{sha}}(z)H_{\text{pro}}(z). \quad (2)$$

$H_{\text{sha}}(z)$ and $H_{\text{pro}}(z)$ denote the transfer function of the shaping filter and prototype filter, respectively. $X(z)$ and $Y(z)$ denote the input and output of the system, respectively.

The transfer function of a CCF is defined as [4]

$$H_{\text{ccb}}(z) \triangleq 1 + z^{-K}, \quad (3)$$

where the filter order $K = 1, 2, 3, \dots$. The two filter coefficients of $H_{\text{ccb}}(z)$ are both ones, and we can realize a CCF by using only one adder and K memory units. In addition, the coefficients are

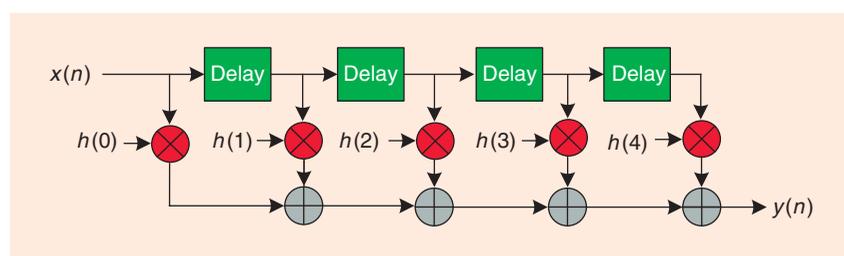


FIGURE 1. An FIR filter block diagram.

Digital Object Identifier 10.1109/MSP.2016.2519919
Date of publication: 27 April 2016

symmetric. Thus, $H_{ccb}(z)$ has a linear phase response [5]. Because the CCF is sufficiently simple, the computation complexity is limited despite numerous CCFs being used. There are K zeros for a CCF, and its zeros are used to reshape $H_{pro}(z)$. The zeros of $H_{ccb}(z)$ are obtained by solving $H_{ccb}(z) = 0$. Applying De Moivre's formula [6] to $z^K = -1$ yields [4]

$$z = \sqrt[K]{-1} \left[\cos\left(\frac{\pi + 2\pi n}{K}\right) + j \sin\left(\frac{\pi + 2\pi n}{K}\right) \right],$$

$$n = 0, 1, \dots, K - 1. \quad (4)$$

Comparing (4) with $z = e^{j\omega} = \cos(\omega) + j \sin(\omega)$, we can locate zeros of $H_{ccb}(\omega)$ as

$$\omega_n = \frac{\pi + 2\pi n}{K}, \quad n = 0, 1, \dots, K - 1. \quad (5)$$

Figure 3 displays the magnitude frequency responses of CCFs for $K = 1, 2, 3, 4$. In the following discussion, all of the filter frequency responses at $\omega = 0$ rad/sample are normalized to 0 dB to facilitate comparison. Although the CCFs are simple to implement, their frequency responses are not satisfactory for reshaping an LPF. We must elaborate on the combination of CCFs of various lengths to reshape the prototype filter.

Figure 4 shows an example of cascading CCFs of orders $K = 1, 2, 3, 4, 5$. Using a cascade of CCFs to reshape an LPF is markedly more effective than using a single CCF. Note that the shaping filter has no multiplier; it uses nine adders and 24 memory units. A by-product of the shaping filter is that it preserves the linear phase response for the composite filter. One trick for improving a prototype filter is to use cascaded CCFs to reshape it. Another method, called the *modified window design method*, is to deliberately design a prototype filter that has an overshoot at the passband edge so that the composite filter has a flat passband. We then detail the design of a lowpass prototype filter.

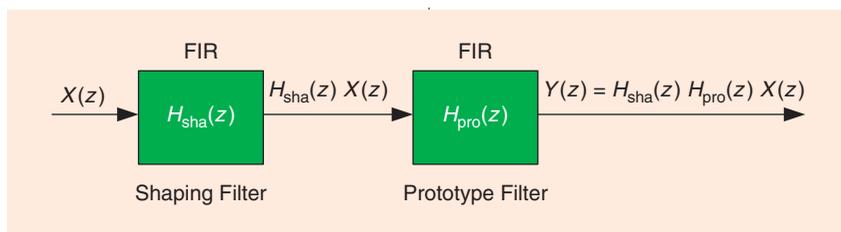


FIGURE 2. The structure of a composite FIR filter.

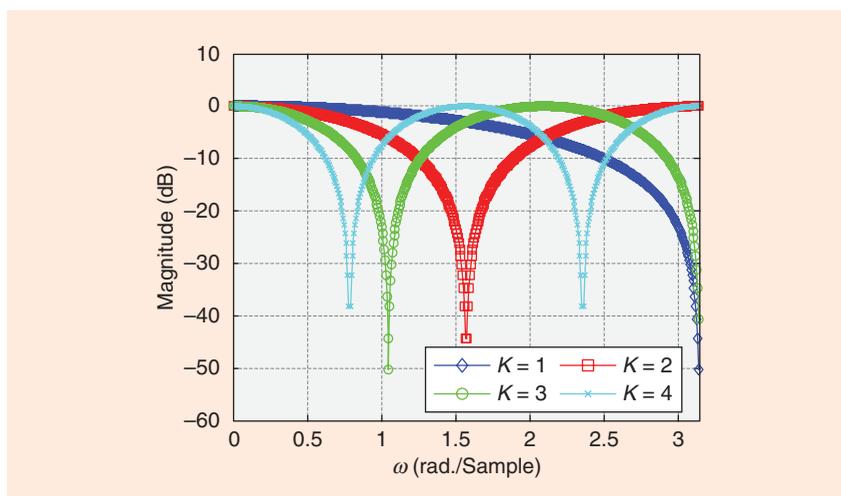


FIGURE 3. The magnitude frequency responses of CCFs for $K = 1, 2, 3, 4$.

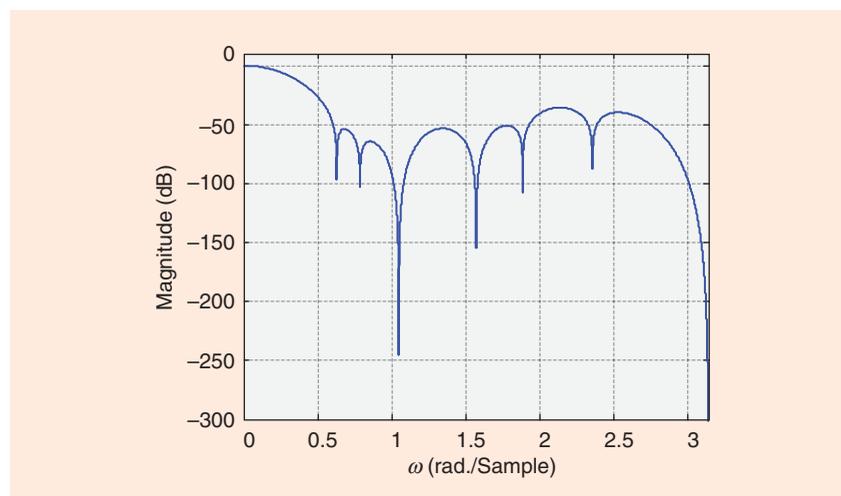


FIGURE 4. The magnitude frequency response of a cascade of CCFs. The transfer function is $H_{sha}(z) = (1 + z^{-1})^2(1 + z^{-2})^2(1 + z^{-3})^3(1 + z^{-4})(1 + z^{-5})$.

Modified window design method

The window method for FIR filter design begins with determining the cutoff frequency of a continuous LPF, which is plotted in a solid black line in Figure 5(a). The continuous LPF is assumed to be ideal with a unity gain at low frequencies and zero (infinite atten-

uation) beyond the cutoff frequency [1]. Let the sample rate of the discrete filter be f_s . The frequency response $H(f)$ on $[-f_s/2, f_s/2]$ is then simulated using a discrete frequency-domain representation $H(m)$ with N samples, which are plotted in black dots in Figure 5(b). The passband of $H(f)$ is equivalent to that

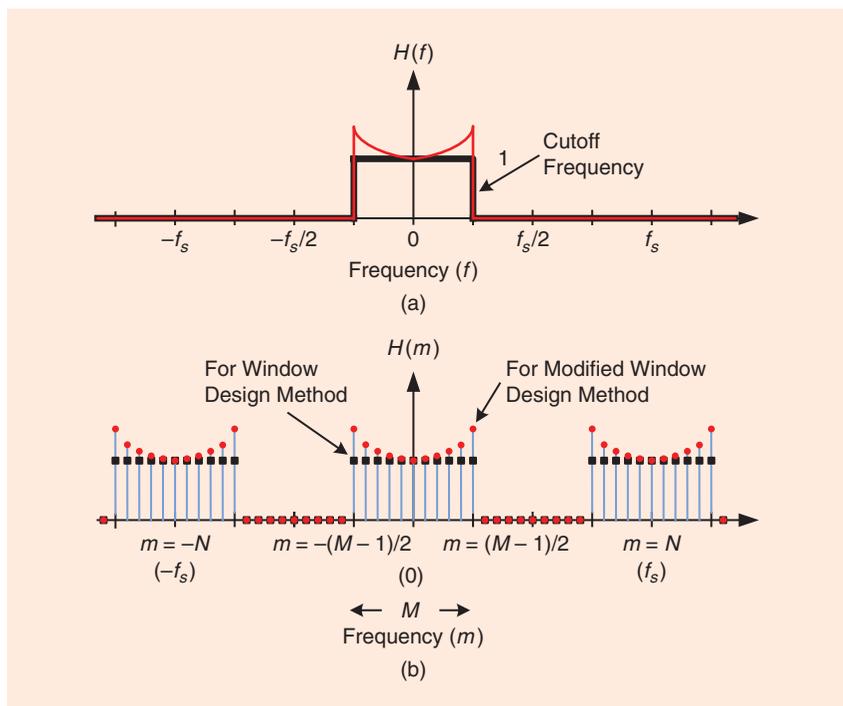


FIGURE 5. LPF frequency responses for the modified window design method and the window design method: (a) continuous frequency response $H(f)$; (b) periodic, discrete frequency response $H(m)$.

of the M frequency-domain samples of $H(m)$. Notice that $H(m)$ is periodic with a period identical to the sample rate f_s . We then apply the discrete frequency response $H(m)$ to the inverse discrete Fourier transform (IDFT) equation to get the time domain $h^\infty(n)$. The filter coefficients are obtained by calculating $h(n) = h^\infty(n) \cdot w(n)$, where $w(n)$ represents a window function to truncate $h^\infty(n)$.

In this article, to design the prototype filter with a deliberately generated ripple at the passband edge, we remodel a continuous LPF frequency response $H(f)$ by using an exponential function. This method is called the *modified window design method*.

In contrast to the window design method, the passband frequency response of the modified window design method is exponential and is zero

beyond the cutoff frequency. For the modified window design method, the discrete frequency response is simulated using $H(m) = e^{\gamma|m|}$, $-(M-1)/2 \leq m \leq (M-1)/2$. Notice that the parameter γ controls the shape of $H(m)$. When γ is set at zero, $H(m)$ is reduced to that for the traditional window design method. Applying $H(m)$ as a symmetric exponential function induces a ripple at the passband edge so that the composite filter has a flat passband. The time-domain filter coefficients $h^\infty(n)$ is obtained by applying the IDFT equation to $H(m)$ as

$$h^\infty(n) = \frac{1}{N} \sum_{m=-(M-1)/2}^{(M-1)/2} e^{\gamma|m|} e^{j2\pi mn/N}, \tag{6}$$

where n is an integer. If we set γ at zero, then the modified window design method is reduced to the traditional window design method. Reorganizing (6), we obtain the simplified representation of $h^\infty(n)$ as

$$h^\infty(n) = \frac{1}{N} \left[1 + 2 \sum_{m=1}^{(M-1)/2} e^{\gamma m} \cos\left(\frac{2\pi mn}{N}\right) \right]. \tag{7}$$

The filter coefficients are obtained by calculating

$$h(n) = h^\infty(n) \cdot w(n), \tag{8}$$

where $w(n)$ represents a window function to truncate $h^\infty(n)$. More coefficients can be used for $w(n)$ to narrow the filter transition region. As γ increases, the ripple induced at the passband becomes more substantial.

Both the window design method and modified method enhance the stopband attenuation through a smooth window function other than a rectangular window function. Some commonly used window functions include the Blackman, Chebyshev, and Kaiser window functions [1]. We adopt the Blackman window for illustration.

Figure 6 presents a comparison of the magnitude frequency responses for different choices of γ designed using the modified window design method. In contrast to the traditional

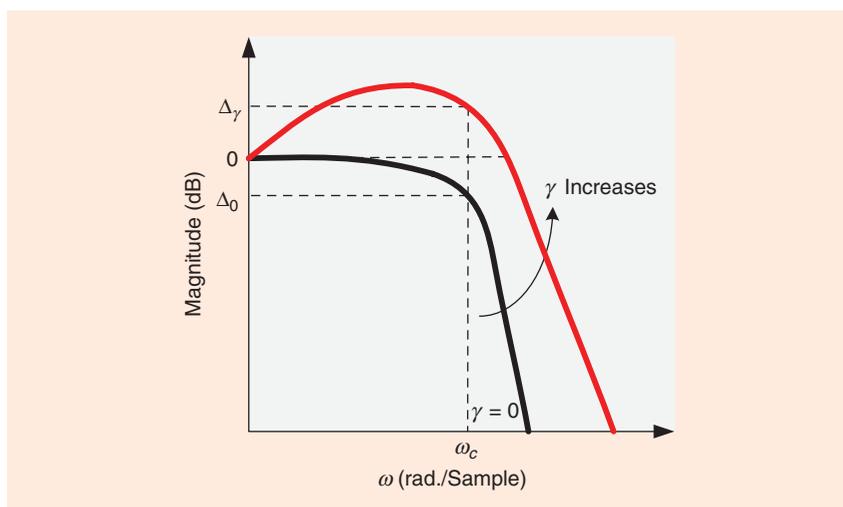


FIGURE 6. The comparison of the magnitude frequency responses of two lowpass FIR filters designed using the modified window design method. One filter is with $\gamma = 0$ and the other is with $\gamma > 0$.

window design method ($\gamma = 0$), the magnitude frequency response for the case $\gamma > 0$ tends to have a wider passband width. Let the magnitude variation at the cutoff frequency ω_c be Δ_γ and the magnitude variation for the case $\gamma = 0$ be Δ_0 . Using γ enables the prototype filter to cascade with the shaping filter. The composite filter is supposed to have an amount of magnitude attenuation at the cutoff frequency identical to that of the competing filters, which we call baseline filters in a subsequent section.

Figure 7 shows the relationship between the magnitude variation Δ_γ and γ at $\omega_c = 0.1\pi$ rad/sample. The lowpass FIR filter has 33 taps, and the Blackman window function is used. The value of Δ_γ increases from Δ_0 to +5.35 dB as γ increases from 0 to 0.28, and Δ_γ tends to saturate when γ is higher than 0.2.

Design procedures

Observing Figure 6 shows that the response of the prototype filter at ω_c plays a key role in meeting the design specifications. The main design specification regarding filter frequency response is the magnitude of the passband ripple, which affects how the shaping filter is designed. More precisely, if the targeted passband ripple of the composite filter has a magnitude of $-\Delta_0$ decibels, where $\Delta_0 < 0$, then the margin for cascading CCFs at the cutoff frequency ω_c is $\Delta_\gamma - \Delta_0$. Suppose there are a total of L types of CCFs whose passband widths are all higher than ω_c . We use k_i CCFs of order i , where $i = 1, 2, \dots, L$, to construct the shaping filter by cascading all of them. We can relate the transfer function of the composite filter as

$$|H_{\text{cmp}}(z)| = |H_{\text{pro}}(z)| \cdot |1 + z^{-1}|^{k_1} \cdot |1 + z^{-2}|^{k_2} \cdots |1 + z^{-L}|^{k_L}. \quad (9)$$

If k_i equals zero, then we use no CCF of order i .

Let δ_i denote the attenuation of an order- i CCF at ω_c with respect to its response at dc. Notice that the attenuation increases with respect to the increase of the filter order;

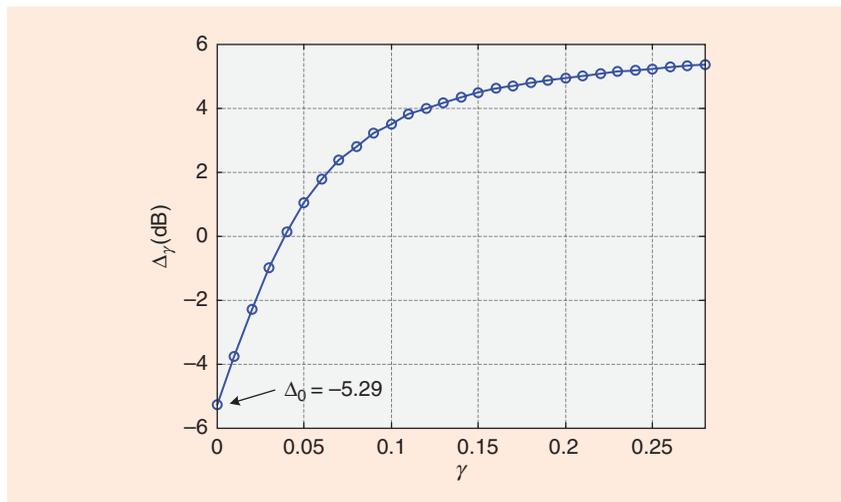


FIGURE 7. The relationship between the magnitude variation Δ_γ and γ at $\omega_c = 0.1\pi$ rad/sample.

$0 < \delta_1 < \delta_2 < \dots < \delta_L$. We can cascade some CCFs with the prototype filter so that the frequency response at ω_c approaches Δ_0 . More precisely, the filter design problem can be transferred to an integer programming problem in searching feasible sets (k_1, k_2, \dots, k_L) . We formulate the integer programming problem as in (10) shown in the box at the bottom of the page.

The indicator function $\mathbf{1}_A$ in (10) is 1 if A is true and 0 if A is false. Equation (10) searches feasible solutions that correspond to a combination of CCFs having as diverse responses as possible. Equation (11) constrains k_i , for $i = 1, \dots, L$, to be either natural numbers (positive integers, denoted as \mathbb{N}) or zeros. Equation (12) refines the candidate solutions so that the attenuation induced by the shaping filter at ω_c reaches the extreme. Alternatively, the attenuation reaches either the amount $\Delta_\gamma - \Delta_0$ or no less than $\Delta_\gamma - \Delta_0 - \delta_1$. Finally, (13) limits the complexity of the shaping filter to a maximum of η CCFs.

$$\max_{(k_1, k_2, \dots, k_L)} \sum_{i=1}^L \mathbf{1}_{\{k_i \neq 0\}} \quad (10)$$

subject to

$$k_1, k_2, \dots, k_L \in \mathbb{N} \cup \{0\}, \quad (11)$$

$$\Delta_\gamma - \Delta_0 - \delta_1 < k_1 \delta_1 + k_2 \delta_2 + \dots + k_L \delta_L \leq \Delta_\gamma - \Delta_0, \quad (12)$$

$$k_1 + k_2 + \dots + k_L \leq \eta. \quad (13)$$

In the next paragraph, we present our filter design procedures. The design specifications include 1) the magnitude of the passband ripple, $|\Delta_0|$ dB, 2) the tap number of the prototype filter, Q , and 3) the cutoff frequency, ω_c .

Step 1

Determine the maximal order L of the CCFs that constitute the shaping filter and calculate their amounts of attenuation at ω_c . For example, if ω_c is 0.1π rad/sample, from (5) we conclude that $L = 9$. We tabulate the attenuations δ_i for $1 \leq i \leq L$ in Table 1.

Step 2

Determine the relation between γ and Δ_γ at ω_c for a prototype filter designed using the modified window design method. This relation depends on the value of the cutoff frequency. When ω_c is 0.1π rad/sample and Q is 33, the results are shown in Figure 7.

Step 3

According to Figure 7, use one γ so that the margin $\Delta_\gamma - \Delta_0$ is sufficient for cascading a prototype filter. For an ordinary lowpass FIR filter, where ω_c lies between 0.1π and 0.4π rad/sample, set γ at 0.1. Thus, the resulting Δ_γ is +3.5 dB and the margin $\Delta_\gamma - \Delta_0$ becomes 8.79 dB.

Step 4

Obtain the transfer function of the prototype filter $H_{\text{pro}}(z)$ by using the

Table 1. The relation between CCF order and attenuation at ω_c .

| Filter order, i | First zero, ω_0 (rad/sample) | Attenuation, δ_i (dB) |
|-------------------|-------------------------------------|------------------------------|
| 1 | π | 0.1 |
| 2 | 0.5π | 0.44 |
| 3 | 0.33π | 1 |
| 4 | 0.25π | 1.84 |
| 5 | 0.2π | 3 |
| 6 | 0.167π | 4.62 |
| 7 | 0.143π | 6.86 |
| 8 | 0.125π | 10.2 |
| 9 | 0.111π | 16.1 |

Table 2. The number of candidate solutions satisfying (11)–(13).

| Constraint(s) | Number of candidate solutions |
|----------------------|-------------------------------|
| Equation (11) | 9,369 |
| Equations (11), (12) | 418 |
| Equations (11)–(13) | 44 |

modified window design method, where the parameter γ is obtained in Step 3.

Step 5

Determine the maximum number of CCFs to be used in the shaping filter. This parameter is denoted as η , which

constrains the complexity of the shaping filter. For example, we can set it at $\eta = 10$.

Step 6

Substitute the results obtained in Steps 1–5 into (10)–(13), and find all the feasible solutions. This optimization problem can be solved using the software MATLAB. Table 2 shows the number of feasible solutions that meet the constraints. Among the 44 candidate solutions, only four of them maximize (10): $(k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8, k_9) = (2, 2, 3, 1, 1, 0, 0, 0, 0), (1, 1, 2, 1, 0, 1, 0, 0, 0), (2, 3, 1, 1, 0, 1, 0, 0, 0), (4, 2, 1, 2, 1, 0, 0, 0, 0)$. These four solutions correspond to our composite filters 1–4.

Step 7

Construct the transfer function of the composite FIR filter by using (2). This completes our filter design procedures.

LPF design example

To emphasize the superiority of the proposed tricks over the old filter design methods mentioned in filter design textbooks, we compare our composite FIR filter with three widely used methods under the constraint of identical pass-

band flatness. These three baseline FIR filter designs are 1) the window design method [1], 3) the Parks–McClellan FIR filter design method [2], and 3) the cascaded integrator-comb (CIC) filter [7]. Please note that although the Butterworth and Chebyshev filters can easily outperform these three baseline FIR designs in magnitude response, they belong to IIR filters and cannot achieve linear phase responses. Therefore, we omit IIR filters from our comparison. For information on the cascading tricks for improving an IIR filter, readers can refer to the techniques mentioned in [4].

For the first solution, the frequency response of the shaping filter is shown in Figure 4, and the frequency response of the composite FIR filter is presented in Figure 8. Figure 8 also displays the three baseline filters for comparison. Each of the two baseline filters designed using the Parks–McClellan design method and the window design method has a number of taps (33) identical to that of our prototype filter. The only difference among them lies in their filter coefficients. The third baseline filter is an 11-tap CIC filter, which also meets the targeted passband specification. Clearly, composite filter 1 has the sharpest transition band and the highest suppression at the stopband compared with the three baseline filters, except for the response at $\omega = 0.58$ (rad/sample). However, the magnitude response of the composite filter at $\omega = 0.58$ (rad/sample) is approximately -50 dB. Compared with the three baseline filters, our composite FIR filter has a more favorable overall response. As shown in Figure 8, all four filters have linear phase responses at the passband.

Figure 9 shows a pole-zero diagram of the shaping filter of composite filter 1. Because the shaping filter consists of only FIR filters, there are only zeros and no pole on the z -plane. In particular, all of the zeros are on the unit circle because all of the filter coefficients of CCFs are ones. Substituting $z = e^{j\omega}$ for the filter zeros in Figure 9 indicates that all the zeros are located outside the passband of the prototype filter. Some frequencies possess two, three, or even more zeros. This implies that both the transition band and stopband of the

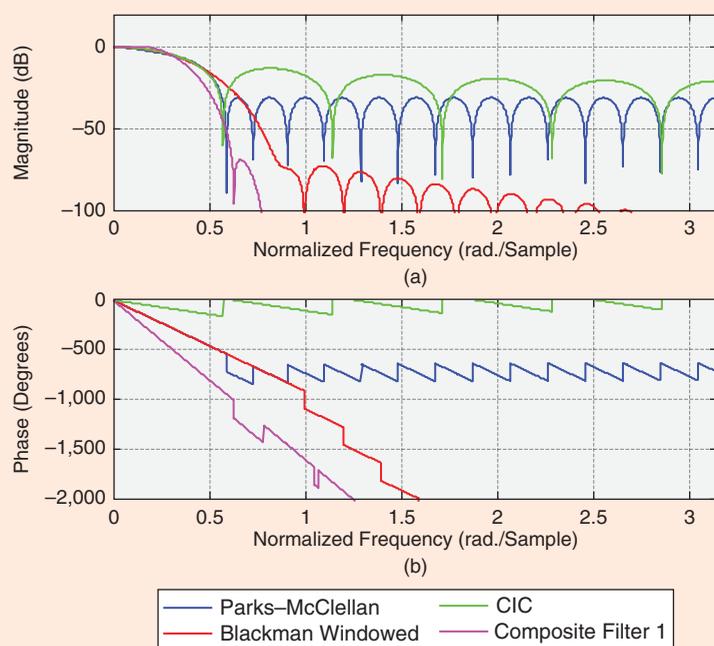


FIGURE 8. A comparison of the frequency responses of composite FIR filter 1 and the three baseline filters.

prototype filter are either sharpened or suppressed by the shaping filter.

Figures 10–12 present comparisons of our composite FIR filters (2, 3, and 4) with the three baseline filters. Except at a few isolated frequencies, our composite FIR filters outperform the three baseline filters at nearly all the frequency bands. For filters 1, 3, and 4, the improvement in magnitude responses can easily exceed 20 dB. Because the shaping filter consists of CCFs with symmetric coefficients, the composite filter preserves a linear phase response at its passband.

Bandpass and highpass FIR filter design

The design of a lowpass FIR filter can be used as the first step in designing a bandpass FIR filter or a highpass FIR filter [1]. This article presents tricks for designing an FIR LPF with an exceptionally sharp transition band and a highly suppressed stopband. We can shift its frequency response in the frequency domain by multiplying a sinusoid $s_{\text{shift}}(n)$ of an appropriate frequency with the composite filter's coefficients. The realization of a bandpass/highpass composite filter is presented in Figure 13.

When designing a highpass FIR filter, the frequency of the sinusoid is $f_s/2$ Hz. The sinusoid sequence is

$$s_{\text{shift1}}(n) = e^{j2\pi\left(\frac{f_s}{2}\right)(nT_s)} = e^{j\pi n} = (1, -1, 1, -1, 1, -1, \dots), \tag{14}$$

where T_s is the reciprocal of the sample rate f_s .

For designing a bandpass FIR filter, a suitable sinusoid is selected according to the passband center of the desired BPF. If the passband is located at, for example, $+f_s/4$ Hz, then the sinusoid sequence is

$$s_{\text{shift2}}(n) = e^{j2\pi\left(\frac{f_s}{4}\right)(nT_s)} = e^{j\pi n/2} = (1, j, -1, -j, 1, j, \dots). \tag{15}$$

However, when the passband of the LPF is separated into two parts, where

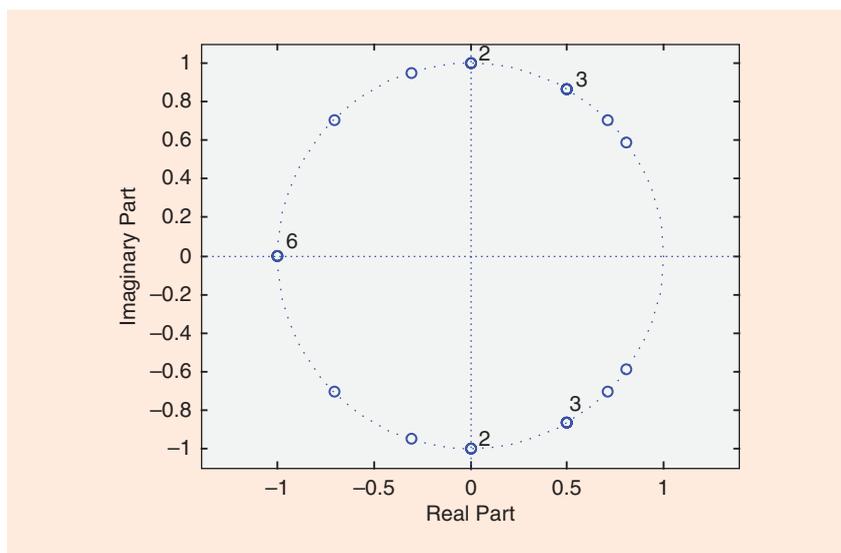


FIGURE 9. Pole-zero diagram of the shaping filter of composite FIR filter 1.

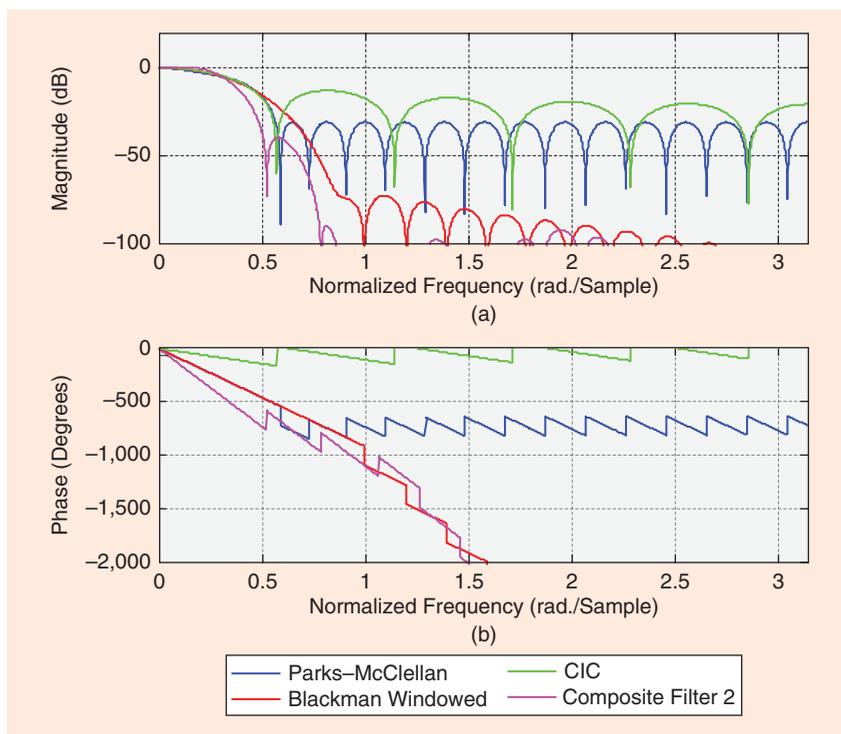


FIGURE 10. The comparison of the frequency responses of composite FIR filter 2 and the three baseline filters.

one part is located at $+f_s/4$ Hz and the other part is located at $-f_s/4$ Hz, the sinusoid sequence thus becomes

$$s_{\text{shift3}}(n) = \frac{1}{2} \left(e^{j2\pi\left(\frac{f_s}{4}\right)(nT_s)} + e^{j2\pi\left(-\frac{f_s}{4}\right)(nT_s)} \right) = \cos\left(2\pi \cdot \frac{f_s}{4} \cdot nT_s\right) = (1, 0, -1, 0, 1, 0, \dots). \tag{16}$$

Conclusions

This article presented tricks for designing composite filters with exceptionally sharp transition bands and highly suppressed stopbands. We first design a lowpass prototype FIR filter by using the modified window design method, whose passband width exceeds the design specifications. After solving a mathematic optimization problem that

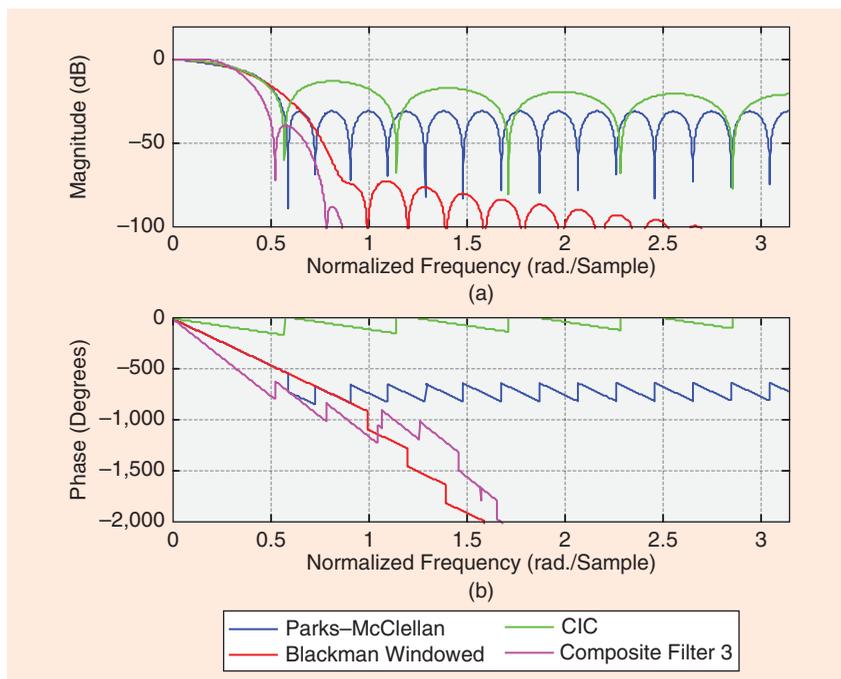


FIGURE 11. The comparison of the frequency responses of composite FIR filter 3 and the three baseline filters.

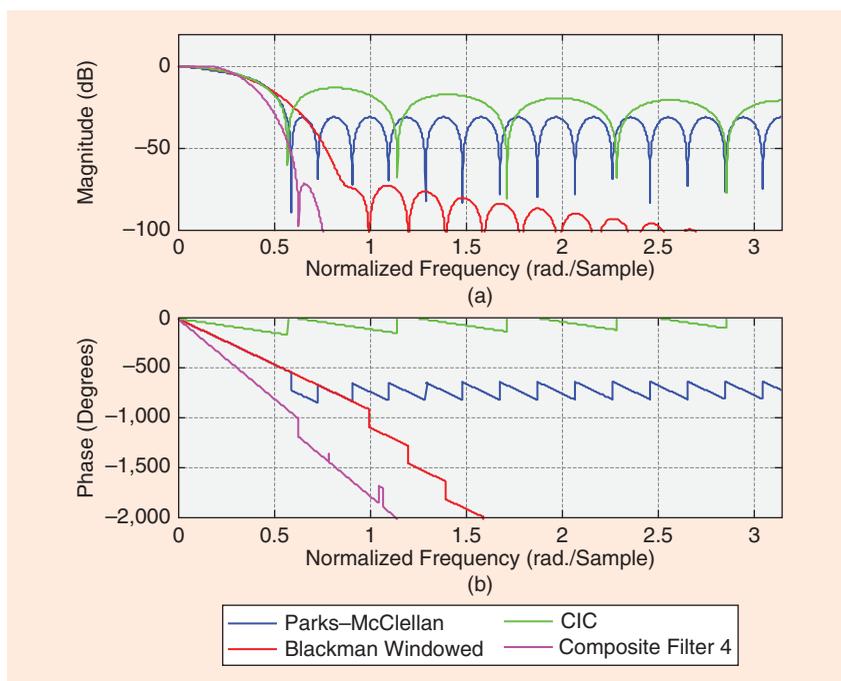


FIGURE 12. The comparison of the frequency responses of composite FIR filter 4 and the three baseline filters.

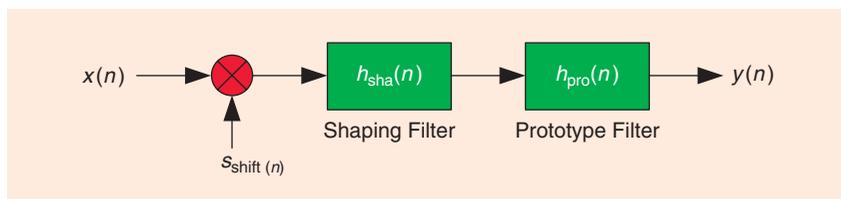


FIGURE 13. The filter structure of a bandpass/highpass composite filter.

accounts for the implementation complexity and design specifications, we obtain the shaping filter. The shaping filter consists of only CCFs of various lengths and is of negligible complexity. On the basis of the composite FIR LPF, we can obtain the design of bandpass and highpass FIR filters by multiplying the filter coefficients with a sinusoid. All of these composite filters preserve appealing linear phase responses at the passbands. In most cases, our composite filters can easily reach more than 20-dB magnitude improvement in the transition bands and at the stopbands, compared with competing baseline FIR filters.

Authors

David Shiung (davids@cc.ncue.edu.tw) is an associate professor at the National Changhua University of Education, Taiwan.

Ya-Yin Yang (ivy yang64@gmail.com) is a postdoctoral fellow at National Cheng Kung University, Tainan, Taiwan.

Chu-Sing Yang (csyang@mail.vee.ncku.edu.tw) is a professor at National Cheng Kung University, Tainan, Taiwan.

References

- [1] R. G. Lyons, *Understanding Digital Signal Processing*. Boston: Pearson Education, 2011.
- [2] J. H. McClellan, T. W. Parks, and L. R. Rabiner, "A computer program for designing optimum FIR linear phase digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, no. 6, pp. 506–526, Dec. 1973.
- [3] MathWorks. (2015, June). Designing low pass FIR filters. [Online]. Available: <http://www.mathworks.com/help/dsp/examples/designing-low-pass-fir-filters.html>
- [4] D. Shiung, Y. Y. Yang, and C. S. Yang, "Tricks for designing infinite impulse response filters with sharp transition bands" *IEEE Signal Processing Mag.*, vol. 33, no. 1, pp. 151–162, Jan. 2016.
- [5] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1999.
- [6] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1965.
- [7] E. B. Hogenauer, "An economical class of digital filters for decimation and interpolation," *IEEE Trans. Audio, Speech, Signal Processing*, vol. 29, no. 2, pp. 155–162, Apr. 1981.

Soo-Chang Pei, Bo-Yi Guo, and Wen-Yang Lu

Narrowband Notch Filter Using Feedback Structure

The design of a notch filter poses challenges in managing the trade-off between distortion in the passband and the notch band. The ideal notch filter has a 3-dB notch bandwidth (BW) equal to zero in the notch band and the unity gain in the passband. Therefore, the design objective is to shrink the BW as small as possible while preserving the gain as close to one as possible.

The task of a notch filter is to eliminate interference while preserving the target broadband signal.

The task of a notch filter is to eliminate interference while preserving the target broadband signal. These filters have many applications for narrowband interference removal in biomedical engineering [1], speech processing [2], image processing [3], [4], and communication [5].

In this article, we describe two ways to implement a notch filter. First, we present a simple notch filter and then illustrate why its performance is less than ideal. Next, we present a trick to apply a novel feedback structure to that notch filter, which improves its performance to nearly ideal. With the help of the feedback structure, we obtain an adjustable parameter. By tuning this parameter, different characteristics of the notch filter are obtained due to the change of its poles' position.

A simple notch filter

The simplest conventional notch filter is given by [6]

$$C_n(z) = \frac{1 - 2 \cos(\omega^*)z^{-1} + z^{-2}}{1 - 2\rho \cos(\omega^*)z^{-1} + \rho^2 z^{-2}}, \quad (1)$$

where ω^* is the notch frequency and ρ is the radius of the poles of $C_n(z)$. For stability, one needs to limit the radius to $0 \leq \rho < 1$. That is, poles should reside within the z -domain's unit circle. The magnitude responses of $C_n(z)$ with different poles' radii of $C_n(z)$ are shown in Figure 1. This notch filter attenuates the signals whose frequencies are in the vicinity of ω^* and allows other frequencies to pass.

The problem and the solution

We find that the notch filter shown in Figure 1 does not have narrow BW and unity gain. The BW is not narrow enough to effectively eliminate the undesired frequency components, and

the nonunity gain in the passband amplifies other frequency components of the input signal.

There is a way to cure this problem—apply a feedback structure to a notch filter. The general feedback transfer function is given by

$$H_N(z) = \frac{(1 + \alpha)C(z)}{1 + \alpha C(z)}, \quad (2)$$

where $\alpha \geq 0$ for stability and $C(z)$ represents any type of existing notch filter. Note that, if $\alpha = 0$, the feedback structure has no effect and the new filter is equal to the original one, i.e., $H_N(z) = C(z)$. The block diagram is illustrated in Figure 2(a). This method not only enhances the effectiveness of any existing notch filter but also gives an extremely narrow bandpass filter. By subtracting (2) from one, we have the extremely narrow bandpass filter written as

$$H_{BP}(z) = 1 - H_N(z) = \frac{1 - C(z)}{1 + \alpha C(z)}. \quad (3)$$

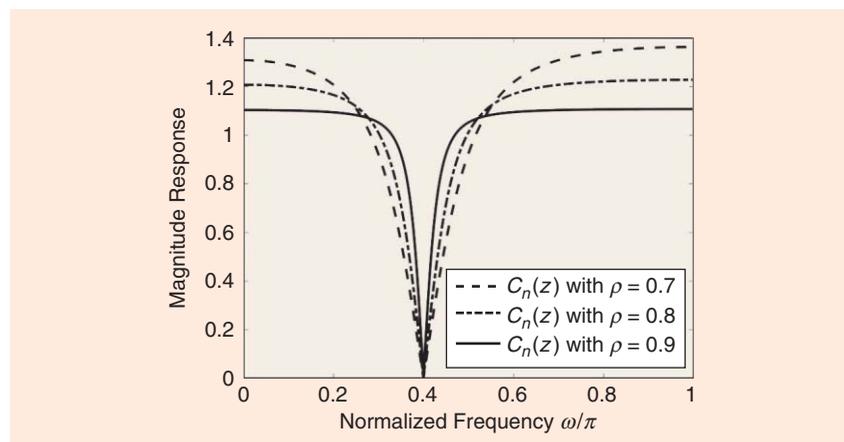


FIGURE 1. The magnitude responses of $C_n(z)$ with different poles' radii of $C_n(z)$.

Digital Object Identifier 10.1109/MSP.2016.2531578
Date of publication: 27 April 2016

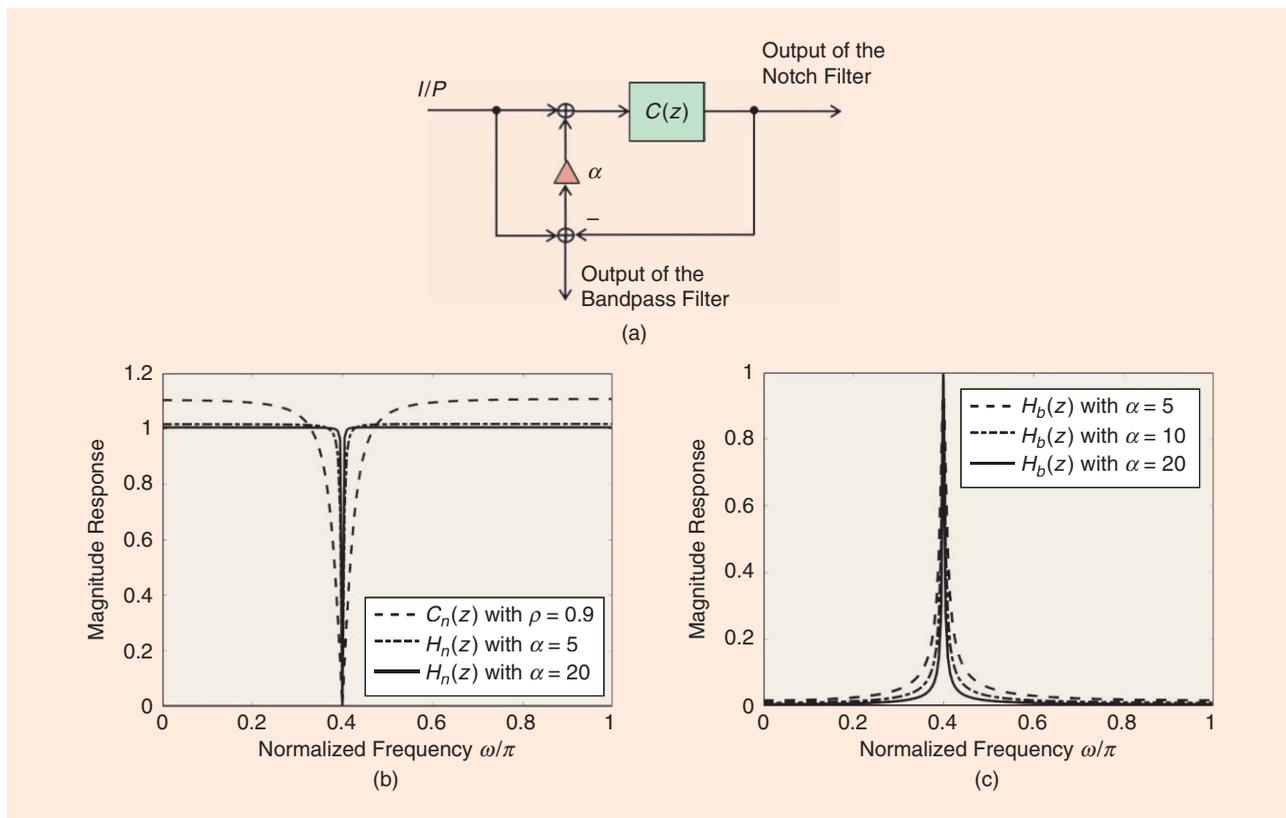


FIGURE 2. (a) The general feedback structure. (b) The magnitude responses of the original notch filter $C_n(z)$ and the improved notch filters $H_n(z)$ with $\alpha = 5$ and 20, when $\rho = 0.9$ and $\omega^* = 0.4\pi$. (c) The magnitude responses of the extremely narrow bandpass filter $H_b(z)$ with $\alpha = 5, 10,$ and 20, when $\rho = 0.9$ and $\omega^* = 0.4\pi$.

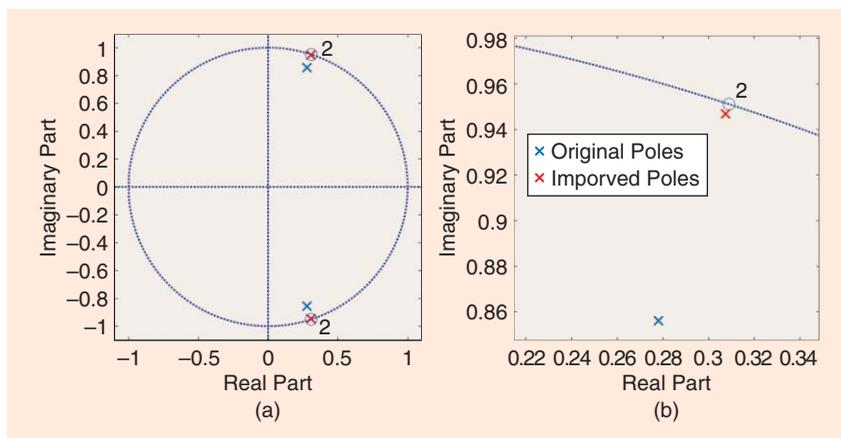


FIGURE 3. The pole-zero plot. (a) Poles and zeros of the original $C_n(z)$ with $\rho = 0.9$ and $H_n(z)$ with $\alpha = 20$. (b) Zoomed portion.

The frequency responses of several notch filters $H_n(z)$ and bandpass filters $H_b(z)$ with various α values are shown in Figure 2(b) and 2(c), respectively. For ease of comparison, we have included the original simple notch filter $C_n(z)$ with $\rho = 0.9$ in Figure 2(b).

From Figure 2(b) and (c), it is clear that the passband of the notch filter is nearly ideal, which implies that the stopband of the bandpass filter is almost zero. Moreover, the BW of both is extremely narrow. That is, the bandpass filter will tend to pass an extremely narrow band signal with frequencies in the neighborhood of ω^* while the notch filter stops the signal with frequencies in the neighborhood of ω^* . Compared to [7], which introduced that a narrow bandpass filter could be obtained by containing a notch filter in a feedback path, our proposed bandpass filter not only has an extremely narrow BW but also preserves the near unity gain as α enlarges.

To see how $H_N(z)$ works, the proposed feedback structure is applied to a conventional notch filter example. Consider the notch filter $C_n(z)$ given in (1) with $\rho = 0.9$ and then substitute it into $C(z)$ in (2) and (3). The proposed notch filter and bandpass filter are given as

$$H_n(z) = \frac{(1 + \alpha)C_n(z)}{1 + \alpha C_n(z)} \quad (4)$$

and

$$H_b(z) = 1 - H_n(z) = \frac{1 - C_n(z)}{1 + \alpha C_n(z)} \quad (5)$$

Influence of the feedback structure

The usage of the feedback structure is to make the notch filter have a more rapid tendency at the notch point. Here, we investigate the slope of the proposed notch filter. Consider a stable notch filter $C(\omega)$, where the parameter z is replaced by ω using the relationship $z = e^{j\omega}$. By differentiating (2), the differential of the proposed notch filter is given as

$$H'_N(\omega) = \frac{(1 + \alpha)}{[1 + \alpha C(\omega)]^2} C'(\omega), \quad (6)$$

where $C'(\omega)$ represents the first-order differential of $C(\omega)$. Let ω_0 be the notch frequency. When $\omega = \omega_0$, both $C(\omega)$ and $H_N(\omega)$ are equal to zero because they are at the notch point. Substituting $\omega = \omega_0$ into (6), we obtain

$$H'_N(\omega_0) = (1 + \alpha) C'(\omega_0), \quad (7)$$

which indicates the relationship between the slope of the proposed filter and the slope of the original filter. By careful observation of (7), when $\alpha > 0$, their relationship can be expressed as

$$|H'_N(\omega_0)| > |C'(\omega_0)|. \quad (8)$$

It is clear from (8) that the proposed notch filter outperforms the original one.

We can also investigate the -3 dB points to prove the effectiveness of the feedback structure. When a filter is at a -3 dB point, the magnitude of its transfer function is equal to $1/\sqrt{2}$. There are two -3 dB points in one filter. To avoid obscurity, let $\omega_{-3\text{dB}}$ be the larger -3 dB frequency of the proposed notch filter, and let $\omega_{-3\text{dB}}^*$ be the larger -3 dB frequency of the original notch filter. It can be organized as follows

$$\begin{cases} |H_N(\omega_{-3\text{dB}})| = \frac{1}{\sqrt{2}} \\ |C(\omega_{-3\text{dB}}^*)| = \frac{1}{\sqrt{2}} \end{cases} \quad (9)$$

Note that $\omega_{-3\text{dB}}$ is unnecessarily equal to $\omega_{-3\text{dB}}^*$. To compare the -3 dB points between these two filters, we substitute $\omega = \omega_{-3\text{dB}}$ into (2), and the equation becomes

$$\begin{aligned} |H_N(\omega_{-3\text{dB}})| &= \frac{|(1 + \alpha)C(\omega_{-3\text{dB}})|}{|1 + \alpha C(\omega_{-3\text{dB}})|} \\ &= \frac{1}{\sqrt{2}}. \end{aligned} \quad (10)$$

If $\alpha > 0$, (10) can be rearranged as

$$|C(\omega_{-3\text{dB}})| \leq \frac{1}{\sqrt{2} + (\sqrt{2} - 1)\alpha}, \quad (11)$$

which shows that

$$|C(\omega_{-3\text{dB}})| < \frac{1}{\sqrt{2}} = |C(\omega_{-3\text{dB}}^*)|. \quad (12)$$

We obtain the relationship of -3 dB points, $\omega_{-3\text{dB}} < \omega_{-3\text{dB}}^*$, from (12) because $\omega_{-3\text{dB}}$ and $\omega_{-3\text{dB}}^*$ are on the right side of the notch point. Since $H_N(\omega)$ and $C(\omega)$ are symmetric with respect to the notch point $\omega = \omega_0$, we can define the notch BW as

$$\begin{cases} BW_{H_N} = 2(\omega_{-3\text{dB}} - \omega_0) \\ BW_C = 2(\omega_{-3\text{dB}}^* - \omega_0) \end{cases}, \quad (13)$$

where BW_{H_N} and BW_C are the BW of $H_N(\omega)$ and $C(\omega)$, respectively. With the relationship of -3 dB points between $H_N(\omega)$ and $C(\omega)$, the relationship between BW_{H_N} and BW_C is

$$BW_{H_N} < BW_C. \quad (14)$$

By investigating the slope and the -3 dB points of the proposed notch filter, we clearly see how the feedback structure improves the performance. Both (7) and (11) give mathematical representations of the influence of the feedback structure. According to (8) and (14), if we choose a coefficient α larger than zero, the

The task of a notch filter is to eliminate interference while preserving the target broadband signal.

proposed notch filter can perform better than the original one. However, the larger the coefficient α is, the longer the transient response is [8]. There is always a tradeoff between the transient response and the performance of a notch filter.

The pole of $H_n(z)$

The proposed feedback structure is suitable not only for the notch filter mentioned in (1) but also for any type of existing notch-like filter such as multiple notch filter [9] and comb filter [10]. For simplicity, take $C_n(z)$ and $H_n(z)$ as an example and illustrate why the method applied to any notch-like filter can improve the performance by deriving the pole of $H_n(z)$. Substitute $C_n(z)$ given in (1) into (4) and rearrange it as shown in (15) in the box at the bottom of the page. To determine the pole radius, let the denominator part equal zero and then use the quadratic formula. See the box at the bottom of the page for (16). By observation and doing some calculations for the square root term, it can be proven that the square root term is negative. Therefore, the radius of the poles is simplified as

$$|z| = \sqrt{\frac{\rho^2 + \alpha}{1 + \alpha}}. \quad (17)$$

Moreover, the inequality $\rho < |z| < 1$ would be automatically satisfied due to $0 \leq \rho < 1$ for stability. The larger α is, the closer the pole will be to the unit circle. This inequality also

implies that the pole of $H_n(z)$ is always closer to the unit circle than the one of $C_n(z)$ but never touches the unit circle, which always

remains stable. It is illustrated in Figure 3, where we use different marks to distinguish the pole from $C_n(z)$ and

$$H_n(z) = \frac{(1 + \alpha)(1 - 2\cos(\omega^*)z^{-1} + z^{-2})}{(1 + \alpha) - 2(\alpha + \rho)\cos(\omega^*)z^{-1} + (\rho^2 + \alpha)z^{-2}} \quad (15)$$

$$z = \frac{2(\rho + \alpha)\cos(\omega^*) \pm \sqrt{4(\rho + \alpha)^2\cos^2(\omega^*) - 4(1 + \alpha)(\rho^2 + \alpha)}}{2(1 + \alpha)} \quad (16)$$

the pole from $H_n(z)$. It is clear that the transfer function with feedback structure always has a better performance than the one without it.

Conclusions

The simple notch filter suffers from two limitations. Its BW is not narrow enough, and the gain in the passband is not near one. We present a trick to improve the performance of a notch filter by using the feedback structure and show the proof and the pole-zero plot to illustrate it.

Authors

Soo-Chang Pei (peisc@ntu.edu.tw) is a professor in the Department of Electrical Engineering at National Taiwan University, ROC.

Bo-Yi Guo (byguo1021@gmail.com) is a master's student at the Graduate Institute of Communication Engineering, National Taiwan University, ROC.

Wen-Yang Lu (plb85ucu@gmail.com) is an undergraduate student in the Department of Electrical and Computer Engineering at National Chiao Tung University, Taiwan, ROC.

References

- [1] S.-C. Pei and C.-C. Tseng, "Elimination of AC interference in electrocardiogram using IIR notch filter with transient suppression," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 11, pp. 1128–1132, Nov. 1995.
- [2] I. Claesson and A. Rossholm, "Notch filtering of humming GSM mobile telephone noise," in *Proc. Int. Conf. Information, Communication and Signal Processing*, 2005, pp. 1320–1323.
- [3] S.-C. Pei and C.-C. Tseng, "Two dimensional IIR digital notch filter design," *IEEE Trans. Circuits Syst. II*, vol. 41, no. 3, pp. 227–231, Mar. 1994.

[4] S.-C. Pei, W.-S. Lu, and C.-C. Tseng, "Two-dimensional FIR notch filter design using singular value decomposition," *IEEE Trans. Circuits Syst. I*, vol. 45, no. 3, pp. 290–294, Mar. 1998.

[5] L. B. Milstein, "Interference rejection techniques in spread spectrum communications," *Proc. IEEE*, vol. 76, no. 6, pp. 657–671, June 1988.

[6] C.-C. Tseng and S.-C. Pei, "Stable IIR notch filter design with optimal pole placement," *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2673–2681, Nov. 2001.

[7] S. Engelberg, "Precise variable-Q filter design," *IEEE Signal Processing Mag.*, vol. 25, no. 5, pp. 113–119, Sept. 2008.

[8] T. Li, J. Jiang, and L. Wang, "Pole-radius-varying IIR notch filter with transient suppression," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 6, pp. 1684–1691, June 2012.

[9] A. Thamrongmas and C. Charoenlarnppanart, "Modified pole re-position technique for optimal IIR multiple notch filter design," *ECTI Trans. Electr. Eng. Electron. Commun.*, vol. 9, no. 1, pp. 7–15, 2011.

[10] G. J. Dolecek and M. Laddomada, "A novel two-stage nonrecursive architecture for the design of generalized comb filters," *Digital Signal Process.*, vol. 22, pp. 859–868, Sept. 2012.

SP

Article Series on Signal Processing in Engineering Design Projects

Call for Information on Student Projects

Signal and information processing plays an important role in many engineering projects. A new article series being developed by a guest editor team for IEEE Signal Processing Magazine (SPM) seeks to provide a platform to share experiences and best practices on undergraduate design project trainings that incorporate a strong element of signal and information processing. SPM invites inputs on recent engineering design projects related to signal and information processing.

What? Input can be made by faculty or industry professionals who have served as mentors for the student projects; input from students is welcome, but endorsement by a faculty/mentor should be included with the submission. Due to space limit, project information will not be published in full in the SPM; instead, the guest editors will review the submissions to select representative projects based on quality, originality, topic coverage, and diversity. An article will be compiled to include summary or highlights of the projects, and published in SPM with acknowledgement of project authors and mentors.

How? Interested project authors and mentors should post their project information as a PDF file on the online repository of the IEEE Signal Processing Society, SigPort.org, under the "Event/Theme" category of "SPM Student Design Project Series" using coupon code "spm1227" and their IEEE web account. The submission should have a cover page that includes the project name, institution, name and email contact of project author(s) and mentor(s), project date/duration, and URL if available. A summary up to 3 pages should describe succinctly the project motivation and main elements as well as outcome and notably highlights. At least one image providing a visual summary or highlight of the project should be included; more relevant images/diagram/photos are welcome; an informative thumbnail image should be included on the SigPort platform. A full project report can be appended as one PDF file, or a URL of the report given.

Guest Editor Team: Hana Godrich, Arye Nehorai, Ali Tajer, Maria Sabrina Greco, and Changshui Zhang

When? Submissions are evaluated on a rolling basis before **31 July 2016**. For best considerations, authors are strongly encouraged to submit as soon as possible.



© GRAPHIC STOCK

ADVERTISING & SALES

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.

CVR 2

IEEE MDL/Marketing
www.ieee.org/go/freemonth

page 11

IEEE USA
www.ieeeusa.org/policy/govfel
+1 202 530 8347

CVR 4

Mathworks
www.mathworks.com/ltc
+1 508 647 7040

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

Mark David

Sr. Manager Advertising & Business Development
Phone: +1 732 465 6473
Fax: +1 732 981 1855
m.david@ieee.org

Mindy Belfer

Advertising Sales Coordinator
Phone: +1 732 562 3937
Fax: +1 732 981 1855
m.belfer@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

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/ 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**
Michael O'Kane
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
mo.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 Giovino
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

Michael O'Kane
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
mo.ieeemedia@ieee.org
Europe, Africa, Middle East



DATES AHEAD

Please send calendar submissions to:
Dates Ahead, Attention: Jessica Barragué
E-mail: j.barrague@ieee.org

2016

JUNE

Eighth International Conference on Quality of Multimedia Experience (QoMEX)

6–8 June, Lisbon, Portugal.
General Chair: Fernando Pereira
URL: <http://qomex2016.lx.it.pt/>

IEEE Second Conference on Network Software (NetSoft 2016)

6–10 June, Seoul, South Korea.
General Chair: James Won-Ki Hong
URL: <http://sites.ieee.org/netsoft/>

International Workshop on Content-Based Multimedia Indexing (CBMI)

15–17 June, Bucharest, Romania.
General Chairs: Bogdan Ionescu and Henning Müller
URL: <http://cbmi2016.upb.ro/>

IEEE Workshop on Statistical Signal Processing (SSP)

26–29 June, Palma de Mallorca, Spain.
General Chairs: Antonio Artés-Rodríguez and Joaquín Míguez
URL: <http://ssp2016.tsc.uc3m.es/>

JULY

IEEE Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)

10–13 July, Rio de Janeiro, Brazil.
General Chairs: Rodrigo C. de Lamare and Martin Haardt
URL: <http://delamare.cetuc.puc-rio.br/sam2016/index.html>

IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)

11–12 July, Bordeaux, France.
General Chair: Yannick Berthoumieu
URL: <http://ivmsp2016.org/>

IEEE International Conference on Multimedia and Expo (ICME)

11–15 July, Seattle, Washington, USA.
General Chairs: Tsuhan Chen, Ming-Ting Sun, and Cha Zhang
URL: <http://www.icme2016.org/>

Digital Object Identifier 10.1109/MSP.2016.2531038
Date of publication: 27 April 2016



IMAGE LICENSED BY INGRAM PUBLISHING

ICME 2016 will be held at the Westin Seattle, Seattle, Washington, United States, 11–15 July 2016.

AUGUST

IEEE 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)

23–26 August, Colorado Springs, Colorado, USA.
General Cochairs: Terry Boult and Ram Nevatia
URL: <http://avss2016.org/>

IEEE 24th European Signal Processing Conference (EUSIPCO)

29 August–2 September, Budapest, Hungary.
General Chair: Lajos Hanzo
URL: <http://www.eusipco2016.org/>

SEPTEMBER

IEEE International Workshop on Machine Learning for Signal Processing (MLSP)

13–16 September, Salerno, Italy.

IEEE International 18th International Workshop on Multimedia Signal Processing (MMSP)

20–23 September, Montreal, Quebec, Canada.

Sensor Signal Processing for Defence (SSPD)

22–23 September, Edinburgh, Great Britain.
URL: <http://www.sspd.eng.ed.ac.uk>

IEEE International Conference on Image Processing (ICIP)

25–28 September, Phoenix, Arizona, USA.
General Chair: Lina Karam
URL: <http://www.ieeeicip2016.org>

NOVEMBER

50th Annual Asilomar Conference on Signals, Systems, and Computers (ASILOMAR)

6–9 November, Pacific Grove, California, USA.
General Chair: Philip Schniter
URL: <http://www.asilomarssconf.org/>

DECEMBER

Picture Coding Symposium (PCS)

4–7 December, Nuremberg, Germany.

Eighth IEEE International Workshop on Information Forensics and Security (WIFS)

5–7 December, Abu Dhabi, UAE.
General Chairs: Ernesto Damiani and Nasir Memon
URL: <http://wifs2016.mdabaie.com/>

IEEE Global Conference on Signal and Information Processing (GlobalSIP)

7–9 December, Greater Washington, D.C., USA.
General Chairs: Zhi Tian and Brian Sadler
URL: <http://2016.ieeeeglobalsip.org>

17th IEEE International Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP)

10–13 December, Curacao, Dutch Antilles.
General Chairs: André L.F. de Almeida and Martin Haardt
URL: <http://www.cs.huji.ac.il/conferences/CAMSAP17/>

IEEE Spoken Language Technology Workshop (SLT)

13–16 December, San Juan, Puerto Rico.

Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

13–16 December, Jeju, South Korea.

2017

MARCH

2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

5–9 March, New Orleans, Louisiana, USA.

SP

General Chairs

Zhi Tian
George Mason Univ.

Brian M. Sadler
Army Research Lab.

Technical Program Chairs

Philip Regalia
Catholic Univ. of America

Trac D. Tran
Johns Hopkins Univ.

Brian Mark
George Mason Univ.

Finance Chair

Jill Nelson
George Mason Univ.

Local Arrangement Chair

Nathalia Peixoto
George Mason Univ.

Publications Chair

Kathleen Wage
George Mason Univ.

Publicity Chairs

Piya Pal
Univ. MD, College Park

Seung-Jun Kim
Univ. MD, Baltimore Cty

Technical Workshop

Liaison Chair
Min Wu
Univ. MD, College Park

Government Panel Chair

Joel Goodman
Naval Research Lab

Industrial Liaison Chairs

Kristine Bell
Metron Inc.

Hang Liu
Catholic Univ.

International Liaison Chairs

Chengyang Yang
BUAA, China

Mounir Ghogho
Univ. of Leeds, UK

Advisory Committee

Monson Hayes
George Mason Univ.

IEEE GlobalSIP**Fourth IEEE Global Conference on Signal and Information Processing**

December 7-9, 2016, Greater Washington D.C., USA

IEEE Signal Processing Society

<http://2016.ieeeglobalsip.org/>

Call for Papers

The fourth IEEE Global Conference on Signal and Information Processing (GlobalSIP) will be held in Washington, DC, USA on December 7-9, 2016. GlobalSIP has rapidly assumed flagship status within the IEEE Signal Processing Society. It focuses broadly on signal and information processing with an emphasis on up-and-coming signal processing themes. The conference features world-class plenary speeches and overview talks, tutorials, exhibits, oral and poster sessions, and government panel discussions on emerging topics and funding opportunities in Signal and Information Processing. GlobalSIP2016 is comprised of co-located symposia selected based on responses to the Call for Symposium Proposals. Featured symposia include:

- General symposium
- Compressed sensing and deep learning
- Signal processing of big data
- Signal and information processing over networks
- Distributed optimization and resource management over networks
- Signal processing and transceiver design for 5G networks
- Secure communication, authentication and privacy
- Cognitive communications and Radar
- Big data analytics in medical imaging
- Signal processing for understanding crowd dynamics
- Signal and information processing for smart grid infrastructure
- Non-commutative theory and applications
- Sparse signal processing for communications
- Autonomous systems
- ESPA: (industrial) emerging signal processing applications

Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possibly references, and with one additional optional 5th page containing only references. Manuscripts should be original (not submitted/published elsewhere) and written in accordance with the standard IEEE double-column paper template.

Conference Highlights

- 14 technical symposia with plenary talks and keynotes overviewing emerging topics in SIP
- Government panel discussions on funding opportunities, trends and targeted topics
- New industrial symposium on emerging SP applications with demos and exhibitions
- Great venue with vibrant cultural, educational, and scientific identity, housing museums (many are free), monuments, art centers, universities, and federal agencies
- Opportunity to attend both GlobalSIP and Globecom(Dec 4-6, 2016) in one trip

Important Dates:

- **June 5, 2016** : Paper Submission Due
- August 5, 2016 : Final Acceptance decisions notifications sent to all authors
- September 5, 2016 : Camera-ready papers due

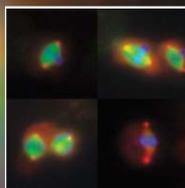




あなたは MATLAB を話しますか?

Over one million people around the world speak MATLAB. Engineers and scientists in every field from aerospace and semiconductors to biotech, financial services, and earth and ocean sciences use it to express their ideas.

Do you speak MATLAB?



*Cells in mitosis:
high-throughput microscopy
for image-based screens.
Provided by Roy Wollman,
Univ. California, Davis.*

Article available at
mathworks.com/ltc

MATLAB®

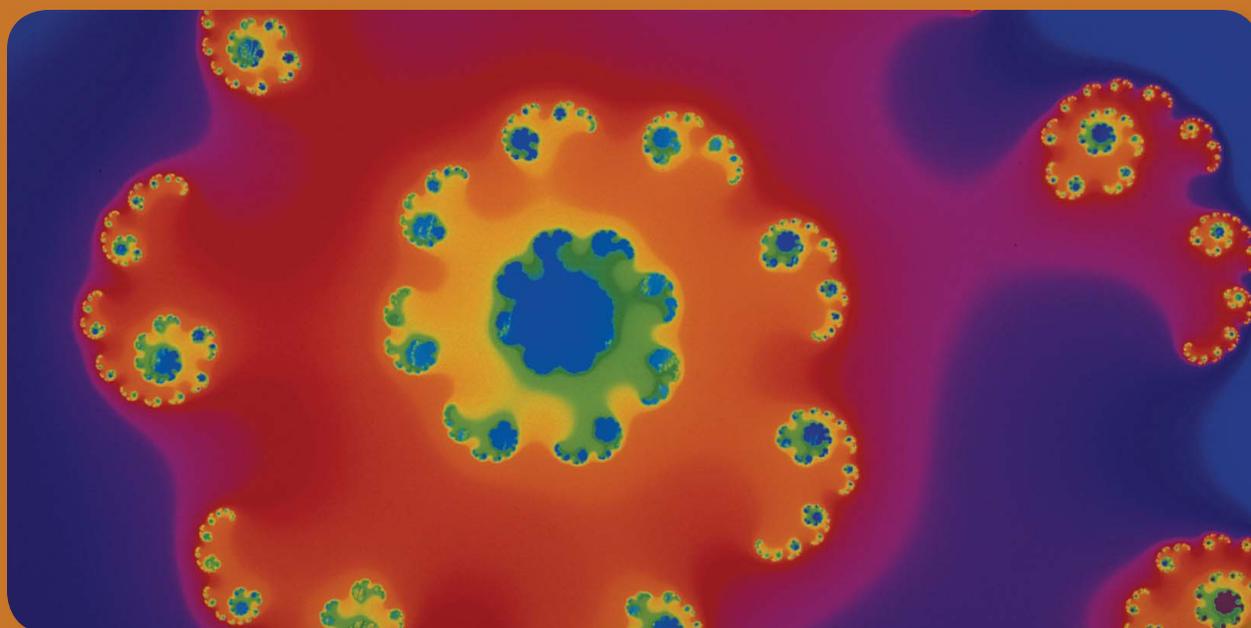
The language of technical computing

IEEE SIGNAL PROCESSING SOCIETY

Content Gazette

MAY 2016

ISSN 2167-5023

**IEEE Journal of Selected Topics in Signal Processing**<http://www.signalprocessingsociety.org/publications/periodicals/jstsp/>**IEEE Signal Processing Letters**<http://www.signalprocessingsociety.org/publications/periodicals/letters/>**IEEE Signal Processing Magazine**<http://www.signalprocessingsociety.org/publications/periodicals/spm/>**IEEE/ACM Transactions on Audio, Speech, and Language Processing**<http://www.signalprocessingsociety.org/publications/periodicals/taslp/>**IEEE Transactions on Image Processing**<http://www.signalprocessingsociety.org/publications/periodicals/image-processing/>**IEEE Transactions on Information Forensics and Security**<http://www.signalprocessingsociety.org/publications/periodicals/forensics/>**IEEE Transactions on Signal Processing**<http://www.signalprocessingsociety.org/publications/periodicals/tsp/>**IEEE Transactions on Multimedia**<http://www.signalprocessingsociety.org/tmm/>

Best AE Practices – ICASSP in Shanghai, China

Over the past several years, the Publications Board of the Signal Processing Society (SPS) has been organizing Panel discussions on Best Practices of Associated Editors (AEs) at its flagship conferences. The first panel discussion was held at ICASSP in Taipei in 2009. The purpose of the panels is to discuss important issues and challenges that Associate Editors face on a regular basis. The intent is also to come up with strategies for coordinating the review process, avoiding common problems, managing difficult situations, and handling plagiarism and author misconducts. In Shanghai, there were 36 participants in the discussion.

The panelists for the ICASSP event were selected by the Editors-in-Chief (EiCs) of the Transactions on Signal Processing (TSP), Transactions on Audio, Speech and Language Processing (TASLP), and Transactions on Information Forensics and Security (TIFS). The panelists were assigned topics that they would introduce for a few minutes with a discussion to follow. The panelists and their topics were:

1. Aleksandar Dogandžić – (TSP): *Controlling the length of the paper; addressing reviewer comments*
2. Amir Asif – (TSP): *Significance of results versus conceptual novelty; marginal contributions*
3. Yuantao Gu – (TSP): *Overriding recommendations of reviewers*
4. Patrick Bas – (TIFS): *Clear communication of decisions to authors; dealing with conflicting reviews*
5. Mads Græsbøll Christensen – (TASLP): *Making informed decisions based on the substance of the reviews*
6. Sharon Gannot – (TASLP): *Quality vs. timeliness of the review*

The remaining panelists were the Chair of the SPS Pub Board (Thrasos Pappas) and the EiCs of the Transactions on Image Processing (Scott Acton), the Transactions on Computational Imaging (Clem Carl), TASLP (Haizhou Li), and the Transactions on Signal and Information Processing over Networks (Petar Djurić).

The discussion provided interesting exchanges of opinions. In general, the AEs agreed that the reviews they receive from their reviewers simply represent recommendations and that they act on them as they see fit. They can choose to follow these recommendations or, in some unusual cases, to deviate from them. In any case, the AEs must clearly justify and communicate their decisions to the authors, and provide specific instructions for addressing the issues raised by the reviewers.

All agreed that the ultimate objective of the whole reviewing process is to recognize manuscripts with good quality and to help the authors improve them before final publication. To that end, the roles of the reviewers and the AEs are critical. The most important question in this process is to understand the novelty and significance of the manuscript. It was emphasized that there may be papers that are technically correct but the significance of the results may be minor; such papers should not be published. As for controlling the length of a paper, it was argued that the authors and the AE are primarily responsible for it.

The discussion about the ever-present dilemma Quality vs. Timeliness suggested that the AEs prefer to wait longer for certain reviews if they believe that they will be of high quality. Practice does show that some reviews are returned quickly but are of limited quality. In those cases, it is acceptable to sacrifice timeliness for quality, but the AEs should keep the authors informed why there is a delay in their decision.

During the discussion, the AEs were reminded of some policies. For instance, the AEs cannot assign themselves as reviewers but can provide reviewer-type feedback (however, not anonymously). They can also request assistance from other AEs.

In summary, the Best Practices Panel was a successful event. It provided the participating AEs with a number of answers to quandaries they may have had. Similar events will be held at future conferences of the society.

Petar M. Djurić
Moderator of the panel

IEEE TRANSACTIONS ON SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



APRIL 15, 2016

VOLUME 64

NUMBER 8

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

- Löwner-Based Blind Signal Separation of Rational Functions With Applications <http://dx.doi.org/10.1109/TSP.2015.2500179>
 *O. Debals, M. Van Barel, and L. De Lathauwer* 1909
- On the Achievable Rate of OFDM With Index Modulation <http://dx.doi.org/10.1109/TSP.2015.2500880>
 *M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor* 1919
- The Role of Principal Angles in Subspace Classification <http://dx.doi.org/10.1109/TSP.2015.2500889>
 *J. Huang, Q. Qiu, and R. Calderbank* 1933
- A Fast Hyperplane-Based Minimum-Volume Enclosing Simplex Algorithm for Blind Hyperspectral Unmixing
<http://dx.doi.org/10.1109/TSP.2015.2508778> *C.-H. Lin, C.-Y. Chi, Y.-H. Wang, and T.-H. Chan* 1946
- Distributed Bayesian Estimation of Linear Models With Unknown Observation Covariances <http://dx.doi.org/10.1109/TSP.2015.2488581> ..
 *Y. Wang and P. M. Djurić* 1962
- Consensus Algorithms With State-Dependent Weights <http://dx.doi.org/10.1109/TSP.2016.2515074> *O. Slučiak and M. Rupp* 1972
- Low-Complexity Algorithms for Low Rank Clutter Parameters Estimation in Radar Systems <http://dx.doi.org/10.1109/TSP.2015.2512535> ..
 *Y. Sun, A. Breloy, P. Babu, D. P. Palomar, F. Pascal, and G. Ginolhac* 1986
-



| | |
|--|------|
| Phase Noise Estimation in OFDM: Utilizing Its Associated Spectral Geometry http://dx.doi.org/10.1109/TSP.2015.2512532 | |
| <i>P. Mathecken, T. Riihonen, S. Werner, and R. Wichman</i> | 1999 |
| Joint Community and Anomaly Tracking in Dynamic Networks http://dx.doi.org/10.1109/TSP.2015.2510971 | |
| <i>B. Baingana and G. B. Giannakis</i> | 2013 |
| Recursive Sparse Point Process Regression With Application to Spectrotemporal Receptive Field Plasticity Analysis http://dx.doi.org/10.1109/TSP.2015.2512560 | |
| <i>A. Sheikhattar, J. B. Fritz, S. A. Shamma, and B. Babadi</i> | 2026 |
| Statistical Analysis of Interference for Nanoscale Electromechanical Wireless Communication at VHF-Band http://dx.doi.org/10.1109/TSP.2015.2512526 | |
| <i>J. J. Lehtomäki, A. O. Bicen, and I. F. Akyildiz</i> | 2040 |
| Sequence Design to Minimize the Weighted Integrated and Peak Sidelobe Levels http://dx.doi.org/10.1109/TSP.2015.2510982 | |
| <i>J. Song, P. Babu, and D. P. Palomar</i> | 2051 |
| Improving Radio Energy Harvesting in Robots Using Mobility Diversity http://dx.doi.org/10.1109/TSP.2016.2518999 | |
| <i>D. Bonilla Licea, S. A. Raza Zaidi, D. McLernon, and M. Ghogho</i> | 2065 |
| Generalized Cramér–Rao Bound for Joint Estimation of Target Position and Velocity for Active and Passive Radar Networks http://dx.doi.org/10.1109/TSP.2015.2510978 | |
| <i>Q. He, J. Hu, R. S. Blum, and Y. Wu</i> | 2078 |
| Unitary Beamformer Designs for MIMO Interference Broadcast Channels http://dx.doi.org/10.1109/TSP.2015.2508782 | |
| <i>S. M. Razavi</i> | 2090 |
| Algebraic Phase Unwrapping Based on Two-Dimensional Spline Smoothing Over Triangles http://dx.doi.org/10.1109/TSP.2015.2510986 .. | |
| <i>D. Kitahara and I. Yamada</i> | 2103 |
| A Multiscale Pyramid Transform for Graph Signals http://dx.doi.org/10.1109/TSP.2015.2512529 | |
| <i>D. I Shuman, M. J. Faraji, and P. Vandergheynst</i> | 2119 |
| Average SINR Calculation of a Persymmetric Sample Matrix Inversion Beamformer http://dx.doi.org/10.1109/TSP.2015.2512527 | |
| <i>J. Liu, W. Liu, H. Liu, B. Chen, X.-G. Xia, and F. Dai</i> | 2135 |
| Closed-Loop Compressive CSIT Estimation in FDD Massive MIMO Systems With 1 Bit Feedback http://dx.doi.org/10.1109/TSP.2016.2515070 | |
| <i>V. K. N. Lau, S. Cai, and A. Liu</i> | 2146 |
| Adaptive Radar Detection of a Subspace Signal Embedded in Subspace Structured Plus Gaussian Interference Via Invariance http://dx.doi.org/10.1109/TSP.2015.2507544 | |
| <i>A. De Maio and D. Orlando</i> | 2156 |
| Dual-Function Radar-Communications: Information Embedding Using Sidelobe Control and Waveform Diversity http://dx.doi.org/10.1109/TSP.2015.2505667 | |
| <i>A. Hassanien, M. G. Amin, Y. D. Zhang, and F. Ahmad</i> | 2168 |
| <hr/> | |
| EDICS—Editors' Information Classification Scheme http://dx.doi.org/10.1109/TSP.2016.2543942 | 2182 |
| Information for Authors http://dx.doi.org/10.1109/TSP.2016.2543943 | 2183 |



IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS



Now accepting paper submissions

The new *IEEE Transactions on Signal and Information Processing over Networks* publishes high-quality papers that extend the classical notions of processing of signals defined over vector spaces (e.g. time and space) to processing of signals and information (data) defined over networks, potentially dynamically varying. In signal processing over networks, the topology of the network may define structural relationships in the data, or may constrain processing of the data. Topics of interest include, but are not limited to the following:

Adaptation, Detection, Estimation, and Learning

- Distributed detection and estimation
- Distributed adaptation over networks
- Distributed learning over networks
- Distributed target tracking
- Bayesian learning; Bayesian signal processing
- Sequential learning over networks
- Decision making over networks
- Distributed dictionary learning
- Distributed game theoretic strategies
- Distributed information processing
- Graphical and kernel methods
- Consensus over network systems
- Optimization over network systems

Communications, Networking, and Sensing

- Distributed monitoring and sensing
- Signal processing for distributed communications and networking
- Signal processing for cooperative networking
- Signal processing for network security
- Optimal network signal processing and resource allocation

Modeling and Analysis

- Performance and bounds of methods
- Robustness and vulnerability
- Network modeling and identification

Modeling and Analysis (cont.)

- Simulations of networked information processing systems
- Social learning
- Bio-inspired network signal processing
- Epidemics and diffusion in populations

Imaging and Media Applications

- Image and video processing over networks
- Media cloud computing and communication
- Multimedia streaming and transport
- Social media computing and networking
- Signal processing for cyber-physical systems
- Wireless/mobile multimedia

Data Analysis

- Processing, analysis, and visualization of big data
- Signal and information processing for crowd computing
- Signal and information processing for the Internet of Things
- Emergence of behavior

Emerging topics and applications

- Emerging topics
- Applications in life sciences, ecology, energy, social networks, economic networks, finance, social sciences, smart grids, wireless health, robotics, transportation, and other areas of science and engineering

Editor-in-Chief: Petar M. Djurić, Stony Brook University (USA)

To submit a paper, go to: <https://mc.manuscriptcentral.com/tsipn-ieee>



Call for Papers
IEEE Signal Processing Society
IEEE Transactions on Signal and Information Processing over Networks

Special Issue on Distributed Information Processing in Social Networks

Over the past few decades, online social networks such as *Facebook* and *Twitter* have significantly changed the way people communicate and share information with each other. The opinion and behavior of each individual are heavily influenced through interacting with others. These local interactions lead to many interesting collective phenomena such as herding, consensus, and rumor spreading. At the same time, there is always the danger of mob mentality of following crowds, celebrities, or gurus who might provide misleading or even malicious information. Many efforts have been devoted to investigating the collective behavior in the context of various network topologies and the robustness of social networks in the presence of malicious threats. On the other hand, activities in social networks (clicks, searches, transactions, posts, and tweets) generate a massive amount of decentralized data, which is not only big in size but also complex in terms of its structure. Processing these data requires significant advances in accurate mathematical modeling and computationally efficient algorithm design.

Many modern technological systems such as wireless sensor and robot networks are virtually the same as social networks in the sense that the nodes in both networks carry disparate information and communicate with constraints. Thus, investigating social networks will bring insightful principles on the system and algorithmic designs of many engineering networks. An example of such is the implementation of consensus algorithms for coordination and control in robot networks. Additionally, more and more research projects nowadays are data-driven. Social networks are natural sources of massive and diverse big data, which present unique opportunities and challenges to further develop theoretical data processing toolsets and investigate novel applications. This special issue aims to focus on addressing distributed information (signal, data, etc.) processing problems in social networks and also invites submissions from all other related disciplines to present comprehensive and diverse perspectives.

Topics of interest include, but are not limited to:

- Dynamic social networks: time varying network topology, edge weights, etc.
- Social learning, distributed decision-making, estimation, and filtering
- Consensus and coordination in multi-agent networks
- Modeling and inference for information diffusion and rumor spreading
- Multi-layered social networks where social interactions take place at different scales or modalities
- Resource allocation, optimization, and control in multi-agent networks
- Modeling and strategic considerations for malicious behavior in networks
- Social media computing and networking
- Data mining, machine learning, and statistical inference frameworks and algorithms for handling big data from social networks
- Data-driven applications: attribution models for marketing and advertising, trend prediction, recommendation systems, crowdsourcing, etc.
- Other topics associated with social networks: graphical modeling, trust, privacy, engineering applications, etc.

Important Dates:

- Manuscript submission due: September 15, 2016
- First review completed: November 1, 2016
- Revised manuscript due: December 15, 2016
- Second review completed: February 1, 2017
- Final manuscript due: March 15, 2017
- Publication: June 1, 2017

Guest Editors:

- Zhenliang Zhang, Qualcomm Corporate R&D (zhenlian@qti.qualcomm.com)
- Wee Peng Tay, Nanyang Technological University (wptay@ntu.edu.sg)
- Moez Draief, Imperial College London (m.draief@imperial.ac.uk)
- Xiaodong Wang, Columbia University (xw2008@columbia.edu)
- Edwin K. P. Chong, Colorado State University (edwin.chong@colostate.edu)
- Alfred O. Hero III, University of Michigan (hero@eecs.umich.edu)

IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



APRIL 2016

VOLUME 24

NUMBER 04

ITASFA

(ISSN 2329-9290)

REGULAR PAPERS

| | | |
|--|---|-----|
| Joint Argument Inference in Chinese Event Extraction with Argument Consistency and Event Relevance http://dx.doi.org/10.1109/TASLP.2015.2497148 | <i>P. Li and G. Zhou</i> | 612 |
| Proportionate Adaptive Filtering for Block-Sparse System Identification http://dx.doi.org/10.1109/TASLP.2015.2499602 | <i>J. Liu and S. L. Grant</i> | 623 |
| Noise Reduction with Optimal Variable Span Linear Filters http://dx.doi.org/10.1109/TASLP.2015.2505416 | <i>J. R. Jensen, J. Benesty, and M. G. Christensen</i> | 631 |
| Enhancement and Noise Statistics Estimation for Non-Stationary Voiced Speech http://dx.doi.org/10.1109/TASLP.2016.2514492 | <i>S. M. Nørholm, J. R. Jensen, and M. G. Christensen</i> | 645 |
| Relationships Between Vocal Function Measures Derived from an Acoustic Microphone and a Subglottal Neck-Surface Accelerometer http://dx.doi.org/10.1109/TASLP.2016.2516647 | <i>D. D. Mehta, J. H. Van Stan, and R. E. Hillman</i> | 659 |
| Unsupervised Word Segmentation and Lexicon Discovery Using Acoustic Word Embeddings http://dx.doi.org/10.1109/TASLP.2016.2517567 | <i>H. Kamper, A. Jansen, and S. Goldwater</i> | 669 |



| | |
|--|-----|
| Joint Dereverberation and Noise Reduction Based on Acoustic Multi-Channel Equalization http://dx.doi.org/10.1109/TASLP.2016.2518804 .. | 680 |
| <i>I. Kodrasi and S. Doclo</i> | |
| Deep Sentence Embedding Using Long Short-Term Memory Networks: Analysis and Application to Information Retrieval http://dx.doi.org/10.1109/TASLP.2016.2520371 | 694 |
| <i>H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, J. Chen, X. Song, and R. Ward</i> | |
| Theory and Perceptual Evaluation of the Binaural Reproduction and Beamforming Tradeoff in the Generalized Spherical Array Beamformer http://dx.doi.org/10.1109/TASLP.2016.2522649 | 708 |
| <i>M. Jeffet, N. R. Shabtai, and B. Rafaely</i> | |
| A Single-Channel Non-Intrusive C50 Estimator Correlated With Speech Recognition Performance http://dx.doi.org/10.1109/TASLP.2016.2521486 | |
| <i>P. Peso Parada, D. Sharma, J. Lainez, D. Barreda, T. van Waterschoot, and P. A. Naylor</i> | 719 |
| Exploiting Turn-Taking Temporal Evolution for Personality Trait Perception in Dyadic Conversations http://dx.doi.org/10.1109/TASLP.2016.2531286 | 733 |
| <i>M.-H. Su, C.-H. Wu, and Y.-T. Zheng</i> | |
| Empirical Use of Information Retrieval to Build Synthetic Data for SMT Domain Adaptation http://dx.doi.org/10.1109/TASLP.2016.2517318 | 745 |
| <i>S. Abdul-Rauf, H. Schwenk, P. Lambert, and M. Nawaz</i> | |
| Postfilters to Modify the Modulation Spectrum for Statistical Parametric Speech Synthesis http://dx.doi.org/10.1109/TASLP.2016.2522655 .. | |
| <i>S. Takamichi, T. Toda, A. W. Black, G. Neubig, S. Sakti, and S. Nakamura</i> | 755 |
| Anti-Spoofing for Text-Independent Speaker Verification: An Initial Database, Comparison of Countermeasures, and Human Performance http://dx.doi.org/10.1109/TASLP.2016.2526653 | |
| <i>Z. Wu, P. L. De Leon, C. Demiroglu, A. Khodabakhsh, S. King, Z.-H. Ling, D. Saito, B. Stewart, T. Toda, M. Wester, and J. Yamagishi</i> | 768 |
| Adaptive Time-Frequency Analysis for Noise Reduction in an Audio Filter Bank With Low Delay http://dx.doi.org/10.1109/TASLP.2016.2526779 | 784 |
| <i>K. T. Andersen and M. Moonen</i> | |
| A Joint Training Framework for Robust Automatic Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2528171 | |
| <i>Z.-Q. Wang and D. L. Wang</i> | 796 |
| Learning Representations for Nonspeech Audio Events Through Their Similarities to Speech Patterns http://dx.doi.org/10.1109/TASLP.2016.2530401 | 807 |
| <i>H. Phan, L. Hertel, M. Maass, R. Mazur, and A. Mertins</i> | |
| EDICS—Editor’s Information Classification Scheme http://dx.doi.org/10.1109/TASLP.2016.2542598 | 823 |
| Information for Authors http://dx.doi.org/10.1109/TASLP.2016.2542618 | 825 |

ANNOUNCEMENT

| | |
|---|-----|
| Call for Papers—Special issue on Sound Scene and Event Analysis http://dx.doi.org/10.1109/TASLP.2016.2542599 | 827 |
|---|-----|

CALL FOR PAPERS

IEEE/ACM Transactions on Audio, Speech and Language Processing
Special issue on **Sound Scene and Event Analysis**

It is still difficult for a machine listening system to demonstrate the same capabilities as human listeners in the analysis of realistic acoustic scenes. Besides speech and music, the analysis of other types of sounds, generally referred to as environmental sounds, is the subject of growing interest from the community and is targeting an ever increasing set of audio categories. In realistic environments, multiple sources are often present simultaneously, and in reverberant conditions, which makes the computational scene analysis challenging.

Typical tasks on audio scene analysis are audio-based scene classification and audio event detection and recognition targeting categories such as “door knocks”, “gunshots”, “crowds”, “car engine noise”, as well as marine mammal and bird species, etc. The wide heterogeneity of possible sounds means that novel types of signal processing and machine learning methods should be developed including novel concepts for audio source segmentation and separation. Beyond recognizing sound scenes and sources of interest, a key task of complex audio scene analysis is sound-source localization.

Further, most of the methods developed until now are probably not tractable on big data so there is also a need for new approaches that are, by design, efficient on large scale problems. Acquiring large scale labelled databases is still problematic and such datasets are most likely collected on heterogeneous sets of acoustic conditions (mobile phone recordings, urban/domestic audio,...) most of which are usually offering a degraded version of the signal of interest with potential variable annotation strategies. Therefore methods to tackle large scale problems also have to be robust against signal degradation, acoustic variability, and annotation variability.

We invite papers on various topics on Sound Scene and Event Analysis, including but not limited to :

- * Audio scene classification;
- * Sound event detection and classification
- * Large-scale environmental audio data sets;
- * Acoustic features for environmental sound analysis;
- * Source localization methods for environmental audio scene analysis
- * Source separation for environmental audio scene analysis
- * Big data in environmental audio;
- * Environmental sound recognition;
- * Computational auditory scene analysis;

The authors are required to follow the Author's Guide for manuscript submission to the IEEE /ACM Transactions on Audio, Speech, and Language Processing at <http://www.signalprocessingsociety.org/publications/periodicals/taslp/>

Important Dates:

Manuscript submission due: July 1st, 2016
First review completed: Sept. 30th 2016
Revised manuscript due: October 20th, 2016
Second review completed: Dec. 1st, 2016
Final manuscript due: Dec. 31st, 2016
Publication date: February 2017

Guest Editors:

Gaël Richard, Télécom ParisTech, France (lead guest editor)
Tuomas Virtanen, Tampere University of Technology, Finland
Juan Pablo Bello, New York University, USA
Nobutaka Ono, National Institute of Informatics, Japan

IEEE TRANSACTIONS ON IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine



APRIL 2016

VOLUME 25

NUMBER 4

IIPRE4

(ISSN 1057-7149)

PAPERS

| | | |
|--|--|------|
| Framelet-Based Sparse Unmixing of Hyperspectral Images http://dx.doi.org/10.1109/TIP.2016.2523345 | <i>G. Zhang, Y. Xu, and F. Fang</i> | 1516 |
| Fourier Spectral Filter Array for Optimal Multispectral Imaging http://dx.doi.org/10.1109/TIP.2016.2523683 | <i>J. Jia, K. J. Barnard, and K. Hirakawa</i> | 1530 |
| Blind Super Resolution of Real-Life Video Sequences http://dx.doi.org/10.1109/TIP.2016.2523344 | <i>E. Faramarzi, D. Rajan, F. C. A. Fernandes, and M. P. Christensen</i> | 1544 |
| The Semi-Variogram and Spectral Distortion Measures for Image Texture Retrieval http://dx.doi.org/10.1109/TIP.2016.2526902 | <i>T. D. Pham</i> | 1556 |
| Learning a Combined Model of Visual Saliency for Fixation Prediction http://dx.doi.org/10.1109/TIP.2016.2522380 | <i>J. Wang, A. Borji, C.-C. J. Kuo, and L. Itti</i> | 1566 |
| Regularization Strategies for Discontinuity-Preserving Optical Flow Methods http://dx.doi.org/10.1109/TIP.2016.2526903 | <i>N. Monzón, A. Salgado, and J. Sánchez</i> | 1580 |
| Dense and Sparse Reconstruction Error Based Saliency Descriptor http://dx.doi.org/10.1109/TIP.2016.2524198 | <i>H. Lu, X. Li, L. Zhang, X. Ruan, and M.-H. Yang</i> | 1592 |
| Texture Classification Using Dense Micro-Block Difference http://dx.doi.org/10.1109/TIP.2016.2526898 | <i>R. Mehta and K. Egiazarian</i> | 1604 |
| Effective Five Directional Partial Derivatives-Based Image Smoothing and a Parallel Structure Design http://dx.doi.org/10.1109/TIP.2016.2526785 | <i>C. Cho and S. Lee</i> | 1617 |
| LBP-Based Segmentation of Defocus Blur http://dx.doi.org/10.1109/TIP.2016.2528042 | <i>X. Yi and M. Eramian</i> | 1626 |
| Image Segmentation Using Parametric Contours With Free Endpoints http://dx.doi.org/10.1109/TIP.2016.2529180 | <i>H. Benninghoff and H. Garcke</i> | 1639 |
| Data-Driven Soft Decoding of Compressed Images in Dual Transform-Pixel Domain http://dx.doi.org/10.1109/TIP.2016.2526910 | <i>X. Liu, X. Wu, J. Zhou, and D. Zhao</i> | 1649 |
| Non-Parametric Blur Map Regression for Depth of Field Extension http://dx.doi.org/10.1109/TIP.2016.2526907 | <i>L. D'Andrès, J. Salvador, A. Kochale, and S. Süsstrunk</i> | 1660 |



| | | |
|---|--|------|
| A Diffusion and Clustering-Based Approach for Finding Coherent Motions and Understanding Crowd Scenes http://dx.doi.org/10.1109/TIP.2016.2531281 | <i>W. Lin, Y. Mi, W. Wang, J. Wu, J. Wang, and T. Mei</i> | 1674 |
| Learn Sparse Dictionaries for Edit Propagation http://dx.doi.org/10.1109/TIP.2016.2523429 | <i>X. Chen, J. Li, D. Zou, and Q. Zhao</i> | 1688 |
| Statistical Performance Analysis of a Fast Super-Resolution Technique Using Noisy Translations http://dx.doi.org/10.1109/TIP.2016.2526901 | <i>P. Chainais and A. Leray</i> | 1699 |
| Weakly Supervised Fine-Grained Categorization With Part-Based Image Representation http://dx.doi.org/10.1109/TIP.2016.2531289 | <i>Y. Zhang, X.-S. Wei, J. Wu, J. Cai, J. Lu, V.-A. Nguyen, and M. N. Do</i> | 1713 |
| Robust Fringe Projection Profilometry via Sparse Representation http://dx.doi.org/10.1109/TIP.2016.2530313 | <i>Budianto and D. P. K. Lun</i> | 1726 |
| Joint Cross-Range Scaling and 3D Geometry Reconstruction of ISAR Targets Based on Factorization Method http://dx.doi.org/10.1109/TIP.2016.2526905 | <i>L. Liu, F. Zhou, X.-R. Bai, M.-L. Tao, and Z.-J. Zhang</i> | 1740 |
| Learning Iteration-wise Generalized Shrinkage-Thresholding Operators for Blind Deconvolution http://dx.doi.org/10.1109/TIP.2016.2531905 | <i>W. Zuo, D. Ren, D. Zhang, S. Gu, and L. Zhang</i> | 1751 |
| Graph-Based Compression of Dynamic 3D Point Cloud Sequences http://dx.doi.org/10.1109/TIP.2016.2529506 | <i>D. Thanou, P. A. Chou, and P. Frossard</i> | 1765 |
| Robust Visual Tracking via Convolutional Networks Without Training http://dx.doi.org/10.1109/TIP.2016.2531283 | <i>K. Zhang, Q. Liu, Y. Wu, and M.-H. Yang</i> | 1779 |
| Automatic Design of Color Filter Arrays in the Frequency Domain http://dx.doi.org/10.1109/TIP.2016.2531287 | <i>C. Bai, J. Li, Z. Lin, and J. Yu</i> | 1793 |
| Reference View Selection in DIBR-Based Multiview Coding http://dx.doi.org/10.1109/TIP.2016.2530303 | <i>T. Maugey, G. Petrazzuoli, P. Frossard, M. Cagnazzo, and B. Pesquet-Popescu</i> | 1808 |
| Analysis of Decorrelation Transform Gain for Uncoded Wireless Image and Video Communication http://dx.doi.org/10.1109/TIP.2016.2535288 | <i>R. Xiong, F. Wu, J. Xu, X. Fan, C. Luo, and W. Gao</i> | 1820 |
| DeepTrack: Learning Discriminative Feature Representations Online for Robust Visual Tracking http://dx.doi.org/10.1109/TIP.2015.2510583 | <i>H. Li, Y. Li, and F. Porikli</i> | 1834 |
| Multiscale Edge Detection Using a Finite Element Framework for Hexagonal Pixel-Based Images http://dx.doi.org/10.1109/TIP.2016.2529720 | <i>B. Gardiner, S. A. Coleman, and B. W. Scotney</i> | 1849 |
| A Feature Learning and Object Recognition Framework for Underwater Fish Images http://dx.doi.org/10.1109/TIP.2016.2535342 | <i>M.-C. Chuang, J.-N. Hwang, and K. Williams</i> | 1862 |
| The Lattice-Based Screen Set: A Square N-Color All-Orders Moiré-Free Screen Set http://dx.doi.org/10.1109/TIP.2016.2523428 | <i>Y.-Y. Chen, T. Kashti, M. Fischer, D. Shaked, R. Ulichney, and J. P. Allebach</i> | 1873 |
| Robust All-in-Focus Super-Resolution for Focal Stack Photography http://dx.doi.org/10.1109/TIP.2016.2523419 | <i>M. Lee and Y.-W. Tai</i> | 1887 |
| Unsupervised Co-Segmentation for Indefinite Number of Common Foreground Objects http://dx.doi.org/10.1109/TIP.2016.2526900 | <i>K. Li, J. Zhang, and W. Tao</i> | 1898 |
| Blind Image Blur Estimation via Deep Learning http://dx.doi.org/10.1109/TIP.2016.2535273 | <i>R. Yan and L. Shao</i> | 1910 |

| | | |
|---|--|------|
| EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIP.2016.2540859 | | 1922 |
| Information for Authors http://dx.doi.org/10.1109/TIP.2016.2540860 | | 1923 |

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING

NEW!



Editor-in-Chief

W. Clem Karl
Boston University

Technical Committee

Charles Bouman
Eric Miller
Peter Corcoran
Jong Chul Ye
Dave Brady
William Freeman

The IEEE Transactions on Computational Imaging publishes research results where computation plays an integral role in the image formation process. All areas of computational imaging are appropriate, ranging from the principles and theory of computational imaging, to modeling paradigms for computational imaging, to image formation methods, to the latest innovative computational imaging system designs. Topics of interest include, but are not limited to the following:

**30% FASTER
TURN AROUND
THAN TIP!**

Computational Imaging Methods and Models

- Coded image sensing
- Compressed sensing
- Sparse and low-rank models
- Learning-based models, dictionary methods
- Graphical image models
- Perceptual models

Computational Image Formation

- Sparsity-based reconstruction
- Statistically-based inversion methods
- Multi-image and sensor fusion
- Optimization-based methods; proximal iterative methods, ADMM

Computational Photography

- Non-classical image capture
- Generalized illumination
- Time-of-flight imaging
- High dynamic range imaging
- Plenoptic imaging

Computational Consumer Imaging

- Mobile imaging, cell phone imaging
- Camera-array systems
- Depth cameras, multi-focus imaging
- Pervasive imaging, camera networks

Computational Acoustic Imaging

- Multi-static ultrasound imaging
- Photo-acoustic imaging
- Acoustic tomography

Computational Microscopy

- Holographic microscopy
- Quantitative phase imaging
- Multi-illumination microscopy
- Lensless microscopy
- Light field microscopy

Imaging Hardware and Software

- Embedded computing systems
- Big data computational imaging
- Integrated hardware/digital design

Tomographic Imaging

- X-ray CT
- PET
- SPECT

Magnetic Resonance Imaging

- Diffusion tensor imaging
- Fast acquisition

Radar Imaging

- Synthetic aperture imaging
- Inverse synthetic aperture imaging

Geophysical Imaging

- Multi-spectral imaging
- Ground penetrating radar
- Seismic tomography

Multi-spectral Imaging

- Multi-spectral imaging
- Hyper-spectral imaging
- Spectroscopic imaging

**NOW WITH
NO PAGE
CHARGES!**

For more information on the IEEE Transactions on Computational Imaging see <http://www.signalprocessingsociety.org/publications/periodicals/tci/>



General Chairs

André L. F. de Almeida
andre@qtel.ufc.br
Federal University of Ceará,
Brazil

Martin Haardt
martin.haardt@tu-ilmenau.de
Ilmenau University of Technology,
Germany

Technical Program Chairs

Xiaoli Ma
xiaoli@gatech.edu
Georgia Institute of Technology,
USA

Shahram ShahbazPanahi
shahram.shahbazpanahi@uoit.ca
University of Ontario Institute of
Technology,
Canada

Finance Chair

Cihan Tepedelenioglu
cihan@asu.edu
Arizona State University,
USA

Publicity and Publications Chair

Ami Wiesel
ami.wiesel@huji.ac.il
Hebrew University of Jerusalem,
Israel

Local Arrangements Chair

Geert Leus
g.j.t.leus@tudelft.nl
Delft University of Technology,
The Netherlands

Robert W. Heath Jr.
rheath@utexas.edu
University of Texas at Austin,
USA



Join Twitter Conversation
#CAMSAP2017



CAMSAP 2017

Call for Papers

The Seventh IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing

Curaçao, Dutch Antilles

December 10-13, 2017

<http://www.cs.huji.ac.il/conferences/CAMSAP17>



Following the success of the first six editions of the IEEE workshop on Computational Advances in Multi-Sensor Adaptive Processing, we are pleased to announce the seventh workshop in this series. IEEE CAMSAP 2017 will be held in Curaçao, Dutch Antilles, and will feature a number of plenary talks from the world's leading researchers in the area, special focus sessions, and contributed papers. All papers will undergo peer review in order to provide feedback to the authors and ensure a high-quality program.

Topics and applications of interest for the workshop include, but are not limited to, the following.

TOPICS OF INTEREST

- Array processing, waveform diversity, space-time adaptive processing
- Convex optimization and relaxation
- Computational linear & multi-linear algebra
- Computer-intensive methods in signal processing (bootstrap, MCMC, EM, particle filtering, etc.)
- Signal and information processing over networks
- Sparse signal processing

APPLICATIONS

- Big data
- Biomedical signal processing
- Communication systems
- Computational imaging
- Radar
- Sensor networks
- Smart grids
- Sonar

Submission of Papers: Prospective authors are invited to submit original full-length papers, with up to four pages for technical content including figures and references, using the formatting guidelines on the website for reviewing purposes. All accepted papers must be presented at the workshop to appear in the proceedings. Best student paper awards, selected by a CAMSAP committee, will also be presented at the workshop.

Special Session Proposals: In addition to contributed sessions, the workshop will also have a number of special sessions. Prospective organizers of special sessions are invited to submit a proposal form, available on the workshop website, by e-mail to the Special Sessions Chair.

IMPORTANT DEADLINES

| | |
|--|------------------------|
| Submission of proposals for special sessions | March, 2017 |
| Notification of special session acceptance | March 15, 2017 |
| Submission of papers | July, 2017 |
| Notification of paper acceptance | September, 2017 |
| Final paper submission | |



IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

MAY 2016

VOLUME 11

NUMBER 5

ITIFA6

(ISSN 1556-6013)

REGULAR PAPERS

| | | |
|--|--|------|
| HMOG: New Behavioral Biometric Features for Continuous Authentication of Smartphone Users http://dx.doi.org/10.1109/TIFS.2015.2506542 | Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani | 877 |
| An Efficient Data-Driven Clustering Technique to Detect Attacks in SCADA Systems http://dx.doi.org/10.1109/TIFS.2015.2512522 | A. Almalawi, A. Fahad, Z. Tari, A. Alamri, R. AlGhamdi, and A. Y. Zomaya | 893 |
| Authentication and Authorization Scheme for Various User Roles and Devices in Smart Grid http://dx.doi.org/10.1109/TIFS.2015.2512525 .. | N. Saxena, B. J. Choi, and R. Lu | 907 |
| Exploring the Usefulness of Light Field Cameras for Biometrics: An Empirical Study on Face and Iris Recognition http://dx.doi.org/10.1109/TIFS.2015.2512559 | R. Raghavendra, K. B. Raja, and C. Busch | 922 |
| Complementary Cohort Strategy for Multimodal Face Pair Matching http://dx.doi.org/10.1109/TIFS.2015.2512561 | Y. Sun, K. Nasrollahi, Z. Sun, and T. Tan | 937 |
| Perceptual Visual Security Index Based on Edge and Texture Similarities http://dx.doi.org/10.1109/TIFS.2016.2515503 | T. Xiang, S. Guo, and X. Li | 951 |
| Boosting 3D LBP-Based Face Recognition by Fusing Shape and Texture Descriptors on the Mesh http://dx.doi.org/10.1109/TIFS.2016.2515505 | N. Werghi, C. Tortorici, S. Berretti, and A. Del Bimbo | 964 |
| Privacy-Preserving Data Aggregation in Mobile Phone Sensing http://dx.doi.org/10.1109/TIFS.2016.2515513 | Y. Zhang, Q. Chen, and S. Zhong | 980 |
| Security Analysis and Improvements on Two Homomorphic Authentication Schemes for Network Coding http://dx.doi.org/10.1109/TIFS.2016.2515517 | C. Cheng, J. Lee, T. Jiang, and T. Takagi | 993 |
| Audio Authentication by Exploring the Absolute-Error-Map of ENF Signals http://dx.doi.org/10.1109/TIFS.2016.2516824 | G. Hua, Y. Zhang, J. Goh, and V. L. L. Thing | 1003 |
| Secure Transmission Against Pilot Spoofing Attack: A Two-Way Training-Based Scheme http://dx.doi.org/10.1109/TIFS.2016.2516825 | Q. Xiong, Y.-C. Liang, K. H. Li, Y. Gong, and S. Han | 1017 |
| PHY-Layer Authentication Using Duobinary Signaling for Spectrum Enforcement http://dx.doi.org/10.1109/TIFS.2016.2516904 | V. Kumar, J.-M. Park, and K. Bian | 1027 |
| A Framework for the Analysis and Evaluation of Algebraic Fault Attacks on Lightweight Block Ciphers http://dx.doi.org/10.1109/TIFS.2016.2516905 | F. Zhang, S. Guo, X. Zhao, T. Wang, J. Yang, F.-X. Standaert, and D. Gu | 1039 |
| Jammer-Assisted Resource Allocation in Secure OFDMA With Untrusted Users http://dx.doi.org/10.1109/TIFS.2016.2516912 | R. Saini, A. Jindal, and S. De | 1055 |
| Network Diversity: A Security Metric for Evaluating the Resilience of Networks Against Zero-Day Attacks http://dx.doi.org/10.1109/TIFS.2016.2516916 | M. Zhang, L. Wang, S. Jajodia, A. Singhal, and M. Albanese | 1071 |

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/>

APRIL 2016

VOLUME 18

NUMBER 4

ITMUF8

(ISSN 1520-9210)

PAPERS

Audio/Speech/Language Analysis and Synthesis

Human Visual System-Based Saliency Detection for High Dynamic Range Content <http://dx.doi.org/10.1109/TMM.2016.2522639> Y. Dong, M. T. Pourazad, and P. Nasiopoulos 549

Compression and Coding

Complexity Control Based on a Fast Coding Unit Decision Method in the HEVC Video Coding Standard <http://dx.doi.org/10.1109/TMM.2016.2524995> A. Jiménez-Moreno, E. Martínez-Enríquez, and F. Díaz-de-María 563

Multiview and 3D Video Compression Using Neighboring Block Based Disparity Vectors <http://dx.doi.org/10.1109/TMM.2016.2525010> ... Y. Chen, X. Zhao, L. Zhang, and J.-W. Kang 576

Free-Energy Principle Inspired Video Quality Metric and Its Use in Video Coding <http://dx.doi.org/10.1109/TMM.2016.2525004> L. Xu, W. Lin, L. Ma, Y. Zhang, Y. Fang, K. N. Ngan, S. Li, and Y. Yan 590

A Low-Power Video Recording System With Multiple Operation Modes for H.264 and Light-Weight Compression <http://dx.doi.org/10.1109/TMM.2016.2525861> H. Kim, C. E. Rhee, and H.-J. Lee 603

Image/Video/Graphics Analysis and Synthesis

Factorization Algorithms for Temporal Psychovisual Modulation Display <http://dx.doi.org/10.1109/TMM.2016.2523425> Z. Gao, G. Zhai, and J. Zhou 614

Display Technology for Multimedia

Link Adaptation for High-Quality Uncompressed Video Streaming in 60-GHz Wireless Networks <http://dx.doi.org/10.1109/TMM.2016.2525012> M. Choi, G. Lee, S. Jin, J. Koo, B. Kim, and S. Choi 627

Multimodal Human-Human and Human-Computer Dialog

Predicting the Performance in Decision-Making Tasks: From Individual Cues to Group Interaction <http://dx.doi.org/10.1109/TMM.2016.2521348> U. Avci and O. Aran 643



Multimodal Human-Machine Interfaces and Interaction

Multimodal Personality Recognition in Collaborative Goal-Oriented Tasks <http://dx.doi.org/10.1109/TMM.2016.2522763> L. Batrinca, N. Mana, B. Lepri, N. Sebe, and F. Pianesi 659

Comparison and Evaluation of Sonification Strategies for Guidance Tasks <http://dx.doi.org/10.1109/TMM.2016.2531978> G. Parsehian, C. Gondre, M. Aramaki, S. Ystad, and R. Kronland-Martinet 674

Content Description and Annotation

Social Diffusion Analysis With Common-Interest Model for Image Annotation <http://dx.doi.org/10.1109/TMM.2015.2477277> C. Lei, D. Liu, and W. Li 687

Knowledge and Semantics Modeling for Multimedia Databases

Folksonomy-Based Visual Ontology Construction and Its Applications <http://dx.doi.org/10.1109/TMM.2016.2527602> Q. Fang, C. Xu, J. Sang, M. S. Hossain, and A. Ghoneim 702

Multimedia Search and Retrieval

Holons Visual Representation for Image Retrieval <http://dx.doi.org/10.1109/TMM.2016.2530399> L. Dong, Y. Liang, G. Kong, Q. Zhang, X. Cao, and E. Izquierdo 714

Query-Adaptive Small Object Search Using Object Proposals and Shape-Aware Descriptors <http://dx.doi.org/10.1109/TMM.2016.2532601> .. S. D. Bhattacharjee, J. Yuan, Y.-P. Tan, and L.-Y. Duan 726

Multimedia Streaming and Transport

mDASH: A Markov Decision-Based Rate Adaptation Approach for Dynamic HTTP Streaming <http://dx.doi.org/10.1109/TMM.2016.2522650> C. Zhou, C.-W. Lin, and Z. Guo 738

Bandwidth-Efficient Packet Scheduling for Live Streaming With Network Coding <http://dx.doi.org/10.1109/TMM.2016.2530411> S. Huang, E. Izquierdo, and P. Hao 752

Wireless and Mobile Multimedia

A Decision-Tree-Based Perceptual Video Quality Prediction Model and Its Application in FEC for Wireless Multimedia Communications <http://dx.doi.org/10.1109/TMM.2016.2525862> A. Hameed, R. Dai, and B. Balas 764

Multimedia Algorithms, Systems, and Interfaces

Learning Personalized Models for Facial Expression Analysis and Gesture Recognition <http://dx.doi.org/10.1109/TMM.2016.2523421> G. Zen, L. Porzi, E. Sangineto, E. Ricci, and N. Sebe 775

Multimedia Networking and Processing In The Cloud and Data Centers

Core Failure Mitigation in Integer Sum-of-Product Computations on Cloud Computing Systems <http://dx.doi.org/10.1109/TMM.2016.2532603> I. Anarado and Y. Andreopoulos 789

CALL FOR PAPERS

IEEE TRANSACTIONS ON MULTIMEDIA Special Issue on Visualization and Visual Analytics for Multimedia <http://dx.doi.org/10.1109/TMM.2016.2539838> 802

Information for Authors <http://dx.doi.org/10.1109/TMM.2016.2538678> 803

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING



www.ieee.org/sp/index.html

MARCH 2016

VOLUME 10

NUMBER 2

IJSTGY

(ISSN 1932-4553)

ISSUE ON STOCHASTIC SIMULATION AND OPTIMIZATIONS IN SIGNAL PROCESSING

EDITORIAL

Introduction to the Issue on Stochastic Simulation and Optimizations in Signal Processing <http://dx.doi.org/10.1109/IJSTSP.2016.2524963> ..
..... *S. McLaughlin, M. Pereyra, A. O. Hero, J.-Y. Tournet, J.-C. Pesquet* 221

OVERVIEW ARTICLE

A Survey of Stochastic Simulation and Optimization Methods in Signal Processing <http://dx.doi.org/10.1109/IJSTSP.2015.2496908> ..
..... *M. Pereyra, P. Schniter, E. Chouzenoux, J.-C. Pesquet, J.-Y. Tournet, A. O. Hero, and S. McLaughlin* 224

PAPERS

Mini-Batch Semi-Stochastic Gradient Descent in the Proximal Setting <http://dx.doi.org/10.1109/IJSTSP.2015.2505682> ..
..... *J. Konečný, J. Liu, P. Richtárik, and M. Takáč* 242

Online Optimization Under Adversarial Perturbations <http://dx.doi.org/10.1109/IJSTSP.2015.2496911> ..
..... *M. A. Donmez, M. Raginsky, and A. C. Singer* 256

Robust Censoring Using Metropolis-Hastings Sampling <http://dx.doi.org/10.1109/IJSTSP.2015.2506142> ..
..... *G. Kail, S. P. Chepuri, and G. Leus* 270

A Randomized Block Sampling Approach to Canonical Polyadic Decomposition of Large-Scale Tensors
<http://dx.doi.org/10.1109/IJSTSP.2015.2503260> ..
..... *N. Vervliet and L. De Lathauwer* 284



| | |
|---|-----|
| Stochastic Spectral Descent for Discrete Graphical Models http://dx.doi.org/10.1109/JSTSP.2015.2505684 | 296 |
| <i>D. Carlson, Y.-P. Hsieh, E. Collins, L. Carin, and V. Cevher</i> | |
| Langevin and Hamiltonian Based Sequential MCMC for Efficient Bayesian Filtering in High-Dimensional Spaces http://dx.doi.org/10.1109/JSTSP.2015.2497211 | 312 |
| <i>F. Septier and G. W. Peters</i> | |
| Blocked Particle Gibbs Schemes for High Dimensional Interacting Systems http://dx.doi.org/10.1109/JSTSP.2015.2509940 | 328 |
| <i>J. Murphy and S. J. Godsill</i> | |
| Gradient Scan Gibbs Sampler: An Efficient Algorithm for High-Dimensional Gaussian Distributions http://dx.doi.org/10.1109/JSTSP.2015.2510961 | 343 |
| <i>O. Féron, F. Orieux, and J.-F. Giovannelli</i> | |
| Rao-Blackwellized Particle Smoothers for Conditionally Linear Gaussian Models http://dx.doi.org/10.1109/JSTSP.2015.2506543 | 353 |
| <i>F. Lindsten, P. Bunch, S. Särkkä, T. B. Schön, and S. J. Godsill</i> | |
| A Shrinkage-Thresholding Metropolis Adjusted Langevin Algorithm for Bayesian Variable Selection http://dx.doi.org/10.1109/JSTSP.2015.2496546 | 366 |
| <i>A. Schreck, G. Fort, S. Le Corff, and E. Moulines</i> | |
| Unified Importance Sampling Schemes for Efficient Simulation of Outage Capacity Over Generalized Fading Channels http://dx.doi.org/10.1109/JSTSP.2015.2500201 | 376 |
| <i>N. Ben Rached, A. Kammoun, M.-S. Alouini, and R. Tempone</i> | |
| Compressive Hyperspectral Imaging via Approximate Message Passing http://dx.doi.org/10.1109/JSTSP.2015.2500190 | 389 |
| <i>J. Tan, Y. Ma, H. Rueda, D. Baron, and G. R. Arce</i> | |
| Cooling-Aware Energy and Workload Management in Data Centers via Stochastic Optimization http://dx.doi.org/10.1109/JSTSP.2015.2500189 | 402 |
| <i>T. Chen, X. Wang, and G. B. Giannakis</i> | |
| A Differential Evolution-Based Approach for Fitting a Nonlinear Biophysical Model to fMRI BOLD Data http://dx.doi.org/10.1109/JSTSP.2015.2502553 | 416 |
| <i>P. Mesejo, S. SAILLET, O. David, C. G. Bénar, J. M. Warnking, and F. Forbes</i> | |
| <hr/> | |
| Information for Authors http://dx.doi.org/10.1109/JSTSP.2016.2528880 | 428 |

CALL FOR PAPERS**IEEE Journal of Selected Topics in Signal Processing****Special Issue on Cooperative Signal Processing for Heterogeneous and Multi-Task
Wireless Sensor Networks****Aims and Scope**

Portable devices usually operate on their own to solve a single or multiple signal processing tasks in a non-cooperative fashion. To achieve superior performance without a dedicated and power-hungry central device, distributed and cooperative processing techniques over wireless sensor networks (WSNs) have received a lot of attention. However, traditional WSNs typically assume an homogeneous setting, in which the cooperation is limited to homogeneous devices that observe the same phenomenon and that are interested in solving a single network-wide signal processing task.

Due to the heterogeneity of the devices, which may belong to the so-called Internet-of-Things in today's digital age, there is a growing interest in system configurations that overcome the limitations of traditional WSNs. Toward this goal, these systems, referred to as heterogeneous and multi-task WSNs, generalize the homogeneous WSNs to settings, in which randomly located heterogeneous devices cooperate with each other although they observe different but overlapping phenomena, and are interested in solving different but related signal processing tasks. Since the cooperation takes place among devices that generate a large volume of data and that may have different observation models and targeted tasks, the algorithm design needs to address different new challenges of big data analytics to allow for a superior performance as compared to the case where the devices would operate on their own, or where they would exchange raw sensor data in an uncontrolled fashion. In particular, besides keeping minimal communication bandwidth and transmit power when solving the multiple tasks simultaneously, the new challenges require a novel theoretical framework for distributed detection, classification, estimation, coding and topology inference when the devices can have different goals (e.g., the estimation of signals with partially overlapping latent subspaces or detection of different but overlapping sets of events) and modes of operation (e.g., different sampling rates, image resolution). Moreover, as the devices may be selfish and/or malicious, new operating principles need to be designed in order to encourage cooperation and avoid selfish/malicious behavior.

This special issue covers the signal processing theory, modeling, algorithms, and implementation aspects related to heterogeneous multi-tasks WSNs and their use in various applications.

Topics of interest include (but are not limited to):

- Ad-hoc heterogeneous and multi-task wireless sensor networks (and applications thereof)
- Adaptive distributed learning over multi-task networks
- Multi-task decentralized and distributed optimization
- Big data analytics over distributed multi-task networks
- Multi-task distributed detection and labeling
- Distributed multi-task parameter and signal estimation
- Audio/speech or image processing over distributed multi-task networks
- Distributed localization and tracking over multi-task networks
- Distributed topology inference and control
- Cooperative communications over heterogeneous networks
- Game-theoretical tools for trust management and cooperation stimulation over multi-task networks

Important Dates:

- | | |
|--|--|
| - Manuscript submission due: June 1, 2016 | - Second review completed: December 1, 2016 |
| - First review completed: August 15, 2016 | - Final manuscript due: January 1, 2017 |
| - Revised manuscript due: October 1, 2016 | - Publication date: April 2017 |

Prospective authors should visit <http://www.signalprocessingsociety.org/publications/periodicals/jstsp/> for information on paper submission. Manuscripts should be submitted using the Manuscript Central system at <http://mc.manuscriptcentral.com/jstsp-ieee>.

Lead Guest Editor:

Abdelhak M. Zoubir, Technische Universität Darmstadt, Germany, zoubir@spg.tu-darmstadt.de

Guest Editors:

Daniel P. Palomar, Hong Kong University of Science and Technology, Hong Kong, palomar@ust.hk

Jorge Plata-Chaves, KU Leuven, Belgium, jplata@esat.kuleuven.be

Alejandro Ribeiro, University of Pennsylvania, USA, aribeiro@seas.upenn.edu

Anna Scaglione, Arizona State University, USA, Anna.Scaglione@asu.edu

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

APRIL 2016

VOLUME 23

NUMBER 4

ISPLEM

(ISSN 1070-9908)

LETTERS

| | |
|--|-----|
| Matched-Field Performance Prediction with Model Mismatch http://dx.doi.org/10.1109/LSP.2016.2524645 | 409 |
| <i>Y. Le Gall, F.-X. Socheleau, and J. Bonnel</i> | |
| Bayesian Blind Identification of Nonlinear Distortion with Memory for Audio Applications http://dx.doi.org/10.1109/LSP.2016.2525005 .. | 414 |
| <i>F. R. Ávila, H. T. Carvalho, and L. W. P. Biscainho</i> | |
| Distance Based Leakage Alignment for Side Channel Attacks http://dx.doi.org/10.1109/LSP.2016.2521441 | 419 |
| <i>W. Yang, Y. Cao, Y. Zhou, H. Zhang, and Q. Zhang</i> | |
| Cumulative Impulse Strength for Epoch Extraction http://dx.doi.org/10.1109/LSP.2016.2519500 | 424 |
| <i>A. P. Prathosh, P. Sujith, A. G. Ramakrishnan, and P. Kumar Ghosh</i> | |
| Estimate the Primary-Link SNR Using Full-Duplex Relay for Underlay Spectrum Sharing http://dx.doi.org/10.1109/LSP.2016.2530091 | 429 |
| <i>G. Zhao, B. Huang, L. Li, and Z. Chen</i> | |
| Image Denoising Using Quadtree-Based Nonlocal Means With Locally Adaptive Principal Component Analysis http://dx.doi.org/10.1109/LSP.2016.2530406 | 434 |
| <i>C. Zuo, L. Jovanov, B. Goossens, H. Q. Luong, W. Philips, Y. Liu, and M. Zhang</i> | |
| Blind Deconvolution With Nonlocal Similarity and l0 Sparsity for Noisy Image http://dx.doi.org/10.1109/LSP.2016.2530855 | 439 |
| <i>W. Ren, J. Tian, and Y. Tang</i> | |
| Asymptotic Expansions for Heavy-Tailed Data http://dx.doi.org/10.1109/LSP.2016.2526625 | 444 |
| <i>G. Pastor, I. Mora-Jiménez, A. J. Caamaño, and R. Jäntti</i> | |
| Compressive Sensing Reconstruction of Correlated Images Using Joint Regularization http://dx.doi.org/10.1109/LSP.2016.2527680 | 449 |
| <i>K. Chang, P. L. Kevin Ding, and B. Li</i> | |
| A New Data Extrapolation Approach Based on Spectral Partitioning http://dx.doi.org/10.1109/LSP.2016.2533602 | 454 |
| <i>V. K. Nguyen, M. D. E. Turley, and G. A. Fabrizio</i> | |
| Region Based Exemplar References for Image Segmentation Evaluation http://dx.doi.org/10.1109/LSP.2016.2517101 | 459 |
| <i>B. Peng, X. Wang, and Y. Yang</i> | |
| Unscented von Mises–Fisher Filtering http://dx.doi.org/10.1109/LSP.2016.2529854 | 463 |
| <i>G. Kurz, I. Gilitschenski, and U. D. Hanebeck</i> | |
| A Robust Gaussian Approximate Fixed-Interval Smoother for Nonlinear Systems With Heavy-Tailed Process and Measurement Noises http://dx.doi.org/10.1109/LSP.2016.2533543 | 468 |
| <i>Y. Huang, Y. Zhang, N. Li, and J. Chambers</i> | |



| | | |
|--|---|-----|
| When Compressive Sensing Meets Data Hiding http://dx.doi.org/10.1109/LSP.2016.2536110 | <i>G. Hua, Y. Xiang, and G. Bi</i> | 473 |
| Transferred Deep Convolutional Neural Network Features for Extensive Facial Landmark Localization http://dx.doi.org/10.1109/LSP.2016.2533721 | <i>S. Zhang, H. Yang, and Z.-P. Yin</i> | 478 |
| Forcing Multiple Spectral Compatibility Constraints in Radar Waveforms http://dx.doi.org/10.1109/LSP.2016.2532739 | <i>A. Aubry, V. Carotenuto, and A. D. Maio</i> | 483 |
| Estimation Performance for the Bayesian Hierarchical Linear Model http://dx.doi.org/10.1109/LSP.2016.2528579 | <i>M. N. El Korso, R. Boyer, P. Larzabal, and B.-H. Fleury</i> | 488 |
| Enhanced Low-Rank Matrix Approximation http://dx.doi.org/10.1109/LSP.2016.2535227 | <i>A. Parekh and I. W. Selesnick</i> | 493 |
| Using Convolution to Estimate the Score Function for Intractable State-Transition Models http://dx.doi.org/10.1109/LSP.2016.2526667 | <i>L. Dai and T. B. Schön</i> | 498 |
| Discriminative Training of NMF Model Based on Class Probabilities for Speech Enhancement http://dx.doi.org/10.1109/LSP.2016.2532903 | <i>H. Chung, E. Plourde, and B. Champagne</i> | 502 |
| Long-Range Motion Trajectories Extraction of Articulated Human Using Mesh Evolution http://dx.doi.org/10.1109/LSP.2016.2536647 | <i>Y. Wu, X. He, B. Kang, H. Song, and T. Q. Nguyen</i> | 507 |
| Globally Optimal Base Station Clustering in Interference Alignment-Based Multicell Networks http://dx.doi.org/10.1109/LSP.2016.2536159 | <i>R. Brandt, R. Mochaourab, and M. Bengtsson</i> | 512 |
| Improving Saliency Detection Via Multiple Kernel Boosting and Adaptive Fusion http://dx.doi.org/10.1109/LSP.2016.2536743 | <i>X. Zhou, Z. Liu, G. Sun, L. Ye, and X. Wang</i> | 517 |
| Interlaced Double-Precision 2-D Angle Estimation Algorithm Using L-Shaped Nested Arrays http://dx.doi.org/10.1109/LSP.2016.2537826 .. | <i>C. Niu, Y. Zhang, and J. Guo</i> | 522 |
| A Large-Scale Open-Source Acoustic Simulator for Speaker Recognition http://dx.doi.org/10.1109/LSP.2016.2537844 | <i>M. Ferràs, S. Madikeri, P. Motlicek, S. Dey, and H. Bourlard</i> | 527 |
| A Joint Relay-and-Antenna Selection Scheme in Energy-Harvesting MIMO Relay Networks http://dx.doi.org/10.1109/LSP.2016.2538290 .. | <i>J. Men, J. Ge, and C. Zhang</i> | 532 |
| Message Passing Algorithms for Upper and Lower Bounding the Coded Modulation Capacity in a Large-Scale Linear System http://dx.doi.org/10.1109/LSP.2016.2531991 | <i>D. Zhang, M. Matthé, L. L. Mendes, and G. Fettweis</i> | 537 |
| No-Reference Quality Assessment for Multiply-Distorted Images in Gradient Domain http://dx.doi.org/10.1109/LSP.2016.2537321 | <i>Q. Li, W. Lin, and Y. Fang</i> | 541 |
| Multidimensional Scaling-Based TDOA Localization Scheme Using an Auxiliary Line http://dx.doi.org/10.1109/LSP.2016.2537371 | <i>W. Jiang, C. Xu, L. Pei, and W. Yu</i> | 546 |
| Distributed Spectrum Estimation Based on Alternating Mixed Discrete-Continuous Adaptation http://dx.doi.org/10.1109/LSP.2016.2539328 | <i>T. G. Miller, S. Xu, R. C. de Lamare, and H. V. Poor</i> | 551 |
| Regenerated Phase-Shifted Sinusoid-Assisted Empirical Mode Decomposition http://dx.doi.org/10.1109/LSP.2016.2537376 | <i>C. Wang, Q. Kemaio, and F. Da</i> | 556 |

IEEE Signal Processing

Volume 33 | Number 3 | May 2016

MAGAZINE

BRAIN ANALYTICS

Enhancing Understanding of Brain Function

The Meaning of the Signal Processing Life

Signal Processing Helps Put Robot Users in Control

Filter Design Tips & Tricks

IEEE Signal Processing Society

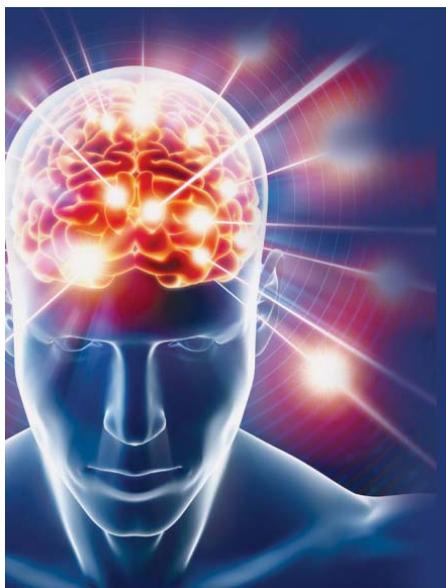
IEEE

Contents

Volume 33 | Number 3 | May 2016

FEATURES

- 12 FROM THE EDITORS**
Z. Jane Wang, Wade Trappe,
and Shuguang Cui
- 14 THE CONNECTED BRAIN**
Adeel Razi and Karl J. Friston
- 36 MAPPING BRAIN ANATOMICAL
CONNECTIVITY USING
DIFFUSION MAGNETIC
RESONANCE IMAGING**
Junning Li, Yonggang Shi,
and Arthur W. Toga
- 52 TIME-VARYING BRAIN
CONNECTIVITY IN fMRI DATA**
Vince D. Calhoun
and Tülay Adalı
- 67 THE PROMISE OF
HYBRID PET/MRI**
Habib Zaidi and
Minerva Becker
- 86 JOINT BLIND SOURCE
SEPARATION FOR
NEUROPHYSIOLOGICAL
DATA ANALYSIS**
Xun Chen, Z. Jane Wang,
and Martin J. McKeown



ON THE COVER

A complementary set of tutorial overview and survey articles demonstrating the importance of incorporating signal processing strategies into the advances in neuroimaging techniques, data analytics, and modeling for brain function is presented in this issue of *IEEE Signal Processing Magazine*. This cluster of feature articles showcases the inherently interdisciplinary nature of brain mapping research and the intriguing signal processing questions.

COVER IMAGE: ©ISTOCKPHOTO.COM/HENRIK5000

- 6 Reader's Choice**
Top Downloads in IEEE Xplore
- 8 Special Reports**
Signal Processing Helps Put Robot
Users in Control
John Edwards
- 108 Tips & Tricks**
Improving FIR Filters by Using
Cascade Techniques
*David Shiung, Ya-Yin Yang,
and Chu-Sing Yang*
- Narrowband Notch Filter Using
Feedback Structure
*Soo-Chang Pei, Bo-Yi Guo,
and Wen-Yang Lu*

PG. 8



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\$213.00 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 © 2016 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.2016.2523299

www.signalprocessingsociety.org [21] MAY 2016

CALL FOR PAPERS AND PROPOSALS
2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)
12-14 December 2016 — Reston, USA
<http://sites.ieee.org/wf-iot-2016/>

IoT: Smart Innovation for Vibrant Ecosystems

The 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT) seeks contributions on how to nurture and cultivate IoT technologies and applications for the benefit of society. Original papers are solicited in, but are not limited to, the following topics:

IoT Enabling Technologies

- 5G Networks and IoT
- Software Defined Network (SDN) and IoT
- Sensor and Actuator Networks
- Ultra-low power IoT Technologies and Embedded Systems Architectures
- Wearables, Body Sensor Networks, Smart Portable Devices
- Design Space Exploration Techniques for IoT Devices and Systems
- Heterogeneous Networks, Web of Things, Web of Everything
- IoT Protocols (IPv6, 6LoWPAN, RPL, 6TiSCH, W3C)
- Internet of Nano Things
- Sensors Data Management, IoT Mining and Analytics
- Adaptive Systems and Models at Runtime
- Distributed Storage, Data Fusion
- Routing and Control Protocols
- Resource Management, Access Control
- Mobility, Localization and Management Aspects
- Identity Management and Object Recognition
- Localization Technologies
- Edge Computing, Fog Computing and IoT
- Machine to Machine (M2M)/Devices-to-Devices communications and IoT
- Industrial IoT and Factory of Things and Internet of Things

IoT Application and Services

- Cyber-physical systems, Context Awareness, Situation Awareness, Ambient Intelligence
- Collaborative Applications and Systems
- Service Experiences and Analysis
- Smart Cities, Smart Public Places, Smart Home/Building Automation
- e-Health, e-Wellness, Automotive, Intelligent Transport
- Smart Grid, Energy Management
- Consumer Electronics, Assisted Living, Rural Services and Production
- Industrial IoT Service Creation and Management Aspects
- Crowd-sensing, human centric sensing
- Big data and IoT Data Analytics
- Internet Applications Naming and Identifiers
- Semantic Technologies, Collective Intelligence
- Cognitive and Reasoning about Things and Smart Objects
- Mobile Cloud Computing (MCC) and IoT
- IoT Multimedia

IoT Societal Impacts

- Human Role in the IoT, Social Aspects and Services
- Value Chain Analysis and Evolution Aspects
- New Human-Device Interactions for IoT, Do-It-Yourself
- Social Models and Networks
- Green IoT: Sustainable Design and Technologies
- Urban Dynamics and crowdsourcing services
- Metrics, Measurements, and Evaluation of IoT Sustainability and ROI

Organizing Committee

Co-Chairs

Najim Dehak
Johns Hopkins University

Pedro Torres-Carrasquillo
MIT Lincoln Laboratory

Technical Co-Chairs

Dilek Hakkani-Tur
Microsoft

Julia Hirschberg
Columbia University

Douglas Reynolds
MIT Lincoln Laboratory

Frank Seide
Microsoft

Zheng Hua Tan
Aalborg University

Dan Povey
Johns Hopkins University

Regional Publicity Chairs

Nancy Chen
I2R, A*STAR

Reinhold Häb-Umbach
Paderborn University

Joseph P. Campbell, Jr.
MIT Lincoln Laboratory

Finance Chair

Douglas O'Shaughnessy
INRS

Sponsorship Chairs

Geoff Zweig
Microsoft

Erik McDermott
Google

Jerome Bellegarda
Apple

Publication Chairs
Patrick Cardinal
ETS

John H. L. Hansen
UT Dallas

Tutorial, Panel and Invited Speakers

Roger K. Moore
Sheffield University

Tara Sainath
Google

Demo Chairs

Chuck Wooters
AST

Andreas Stocke
Microsoft

Advisory Board

James R. Glass
MIT

Jason Williams
Microsoft

2016 IEEE Workshop on Spoken Language Technology



13-16 December 2016 • San Juan, Puerto Rico

The Sixth IEEE Workshop on Spoken Language Technology (SLT) will be held from December 13-16, 2016 in San Juan, Puerto Rico. The theme for this year will be "machine learning from signal to concepts". The workshop is expected to provide researchers around the world the opportunity to interact and present their newest and most advanced research in the fields of speech and language processing. The program for SLT 2016 will include oral and posters sessions, keynotes, plus invited speakers in the field of spoken language as well as tutorials and multiple special sessions.

Topics

Submission of papers is desired on a large variety of areas of spoken language technology, with emphasis on the following topics on previous workshops:

| | |
|---|-------------------------------|
| Speech recognition and synthesis | Human/computer interaction |
| Spoken language understanding | Spoken dialog systems |
| Spoken document retrieval | Speech data mining |
| Question answering from speech | Spoken document summarization |
| Assistive technologies | Spoken language databases |
| Natural language processing | Speaker/language recognition |
| Educational and healthcare applications | Multimodal processing |

Venue

IEEE SLT 2016 will take place in San Juan, Puerto Rico at the InterContinental Hotel in the tourist area of Isla Verde. These areas feature beautiful beaches and a vibrant night life besides a large number of dining options. Additionally, the Old San Juan area is just a few miles away. Additional details about SLT 2016 can be found at: www.slt2016.org

Important Dates

| | |
|---------------------------------|----------------------|
| Special Session Proposals: | June 8, 2016 |
| Paper Submission: | July 22, 2016 |
| Notification of Review Results: | September 14, 2016 |
| Demo Submission: | September 16, 2016 |
| Early Registration Deadline: | October 14, 2016 |
| Workshop: | December 13-16, 2016 |

Submission Details

Authors are invited to prepare a full-length manuscript of 4-6 pages, including reference materials and figures, to the SLT 2016 website: www.slt2016.org



General Chairs

Zhi Tian
George Mason Univ.

Brian M. Sadler
Army Research Lab.

Technical Program Chairs

Philip Regalia
Catholic Univ. of America

Trac D. Tran
Johns Hopkins Univ.

Brian Mark
George Mason Univ.

Finance Chair

Jill Nelson
George Mason Univ.

Local Arrangement Chair

Nathalia Peixoto
George Mason Univ.

Publications Chair

Kathleen Wage
George Mason Univ.

Publicity Chairs

Piya Pal
Univ. MD, College Park

Seung-Jun Kim
Univ. MD, Baltimore Cty

Technical Workshop**Liaison Chair**

Min Wu
Univ. MD, College Park

Government Panel Chair

Joel Goodman
Naval Research Lab

Industrial Liaison Chairs

Kristine Bell
Metron Inc.

Hang Liu
Catholic Univ.

International Liaison Chairs

Chengyang Yang
BUAA, China

Mounir Ghogho
Univ. of Leeds, UK

Advisory Committee

Monson Hayes
George Mason Univ.

IEEE GlobalSIP



Fourth IEEE Global Conference on Signal and Information Processing

December 7-9, 2016, Greater Washington D.C., USA



<http://2016.ieeeglobalsip.org/>

Call for Papers

The fourth IEEE Global Conference on Signal and Information Processing (GlobalSIP) will be held in Washington, DC, USA on December 7-9, 2016. GlobalSIP has rapidly assumed flagship status within the IEEE Signal Processing Society. It focuses broadly on signal and information processing with an emphasis on up-and-coming signal processing themes. The conference features world-class plenary speeches and overview talks, tutorials, exhibits, oral and poster sessions, and government panel discussions on emerging topics and funding opportunities in Signal and Information Processing. GlobalSIP2016 is comprised of co-located symposia selected based on responses to the Call for Symposium Proposals. Featured symposia include:

- General symposium
- Compressed sensing and deep learning
- Signal processing of big data
- Signal and information processing over networks
- Distributed optimization and resource management over networks
- Signal processing and transceiver design for 5G networks
- Secure communication, authentication and privacy
- Cognitive communications and Radar
- Big data analytics in neuro-imaging
- Signal processing for understanding crowd dynamics
- Signal and information processing for smart grid infrastructure
- Non-commutative theory and applications
- Sparse signal processing for communications
- Autonomous systems
- Emerging signal processing applications

Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possibly references, and with one additional optional 5th page containing only references. Manuscripts should be original (not submitted/published elsewhere) and written in accordance with the standard IEEE double-column paper template.

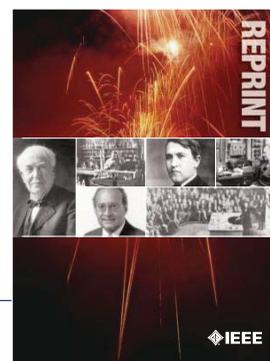
Conference Highlights

- 14 technical symposia with plenary talks and keynotes overviewing emerging topics in SIP
- Government panel discussions on funding opportunities, trends and targeted topics
- New industrial symposium on emerging SP applications with demos and exhibitions
- Great venue with vibrant cultural, educational, and scientific identity, housing museums (many are free), monuments, art centers, universities, and federal agencies
- Opportunity to attend both GlobalSIP and Globecom (Dec 4-6, 2016) in one trip

Important Dates:

- **June 5, 2016** : Paper Submission Due
- August 5, 2016 : Final Acceptance decisions notifications sent to all authors
- September 5, 2016 : Camera-ready papers due





IEEE ORDER FORM FOR REPRINTS

Purchasing IEEE Papers in Print is easy, cost-effective and quick.

Complete this form, send via our secure fax (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
 Piscataway, NJ 08855-1331

Email the Reprint Department at reprints@ieee.org for questions regarding this form

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.
- \$21.50 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: _____

2012 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 | \$129 | \$213 | \$245 | \$248 | \$288 | \$340 | \$371 | \$408 | \$440 | \$477 | \$510 | \$543 |
| 100 | \$245 | \$425 | \$479 | \$495 | \$573 | \$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: _____

2016 IEEE MEMBERSHIP APPLICATION

(students and graduate students must apply online)

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 | 9. Design/development engineering—digital |
| 2. Engineering management | 10. Hardware engineering |
| 3. Project engineering management | 11. Software design/development |
| 4. Research and development management | 12. Computer science |
| 5. Design engineering management—analogue | 13. Science/physics/mathematics |
| 6. Design engineering management—digital | 14. Engineering (not elsewhere specified) |
| 7. Research and development engineering | 15. Marketing/sales/purchasing |
| 8. Design/development engineering—analogue | 16. Consulting |
| | 17. Education/teaching |
| | 18. Retired |
| | 19. Other _____ |

C. Principal responsibility

- | | |
|--|-----------------------|
| 1. Engineering and scientific management | 6. Education/teaching |
| 2. Management other than engineering | 7. Consulting |
| 3. Engineering design | 8. Retired |
| 4. Engineering | 9. Other _____ |
| 5. Software: science/mngmnt/engineering | |

D. Title

- | | |
|--|--------------------------------|
| 1. Chairman of the Board/President/CEO | 10. Design Engineering Manager |
| 2. Owner/Partner | 11. Design Engineer |
| 3. General Manager | 12. Hardware Engineer |
| 4. VP Operations | 13. Software Engineer |
| 5. VP Engineering/Dir. Engineering | 14. Computer Scientist |
| 6. Chief Engineer/Chief Scientist | 15. Dean/Professor/Instructor |
| 7. Engineering Management | 16. Consultant |
| 8. Scientific Management | 17. Retired |
| 9. Member of Technical Staff | 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*

(continued on next page)



5 Add IEEE Society Memberships (Optional)

The 39 IEEE Societies support your technical and professional interests. Many society memberships include a personal subscription to the core journal, magazine, or newsletter of that society. **For a complete list of everything included with your IEEE Society membership, visit www.ieee.org/join.** All prices are quoted in US dollars.

Please check the appropriate box.

| | | BETWEEN 16 AUG 2015- 28 FEB 2016 PAY | BETWEEN 1 MAR 2016- 15 AUG 2016 PAY |
|---|---------|---|--|
| IEEE Aerospace and Electronic Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | AES010 | 25.00 <input type="checkbox"/> | 12.50 <input type="checkbox"/> |
| IEEE Antennas and Propagation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | AP003 | 15.00 <input type="checkbox"/> | 7.50 <input type="checkbox"/> |
| IEEE Broadcast Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | BT002 | 15.00 <input type="checkbox"/> | 7.50 <input type="checkbox"/> |
| IEEE Circuits and Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | CAS004 | 22.00 <input type="checkbox"/> | 11.00 <input type="checkbox"/> |
| IEEE Communications <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | COM019 | 30.00 <input type="checkbox"/> | 15.00 <input type="checkbox"/> |
| IEEE Components, Packaging, & Manu. Tech. <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | CPMT021 | 15.00 <input type="checkbox"/> | 7.50 <input type="checkbox"/> |
| IEEE Computational Intelligence <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | CIS011 | 29.00 <input type="checkbox"/> | 14.50 <input type="checkbox"/> |
| IEEE Computer <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | C016 | 56.00 <input type="checkbox"/> | 28.00 <input type="checkbox"/> |
| IEEE Consumer Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | CE008 | 20.00 <input type="checkbox"/> | 10.00 <input type="checkbox"/> |
| IEEE Control Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | CS023 | 25.00 <input type="checkbox"/> | 12.50 <input type="checkbox"/> |
| IEEE Dielectrics and Electrical Insulation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | DEI032 | 26.00 <input type="checkbox"/> | 13.00 <input type="checkbox"/> |
| IEEE Education <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | ED025 | 20.00 <input type="checkbox"/> | 10.00 <input type="checkbox"/> |
| IEEE Electromagnetic Compatibility <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | EMC027 | 31.00 <input type="checkbox"/> | 15.50 <input type="checkbox"/> |
| IEEE Electron Devices <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | ED015 | 18.00 <input type="checkbox"/> | 9.00 <input type="checkbox"/> |
| IEEE Engineering in Medicine and Biology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | EMB018 | 40.00 <input type="checkbox"/> | 20.00 <input type="checkbox"/> |
| IEEE Geoscience and Remote Sensing <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | GRS029 | 19.00 <input type="checkbox"/> | 9.50 <input type="checkbox"/> |
| IEEE Industrial Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | IE013 | 9.00 <input type="checkbox"/> | 4.50 <input type="checkbox"/> |
| IEEE Industry Applications <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | IA034 | 20.00 <input type="checkbox"/> | 10.00 <input type="checkbox"/> |
| IEEE Information Theory <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | IT012 | 30.00 <input type="checkbox"/> | 15.00 <input type="checkbox"/> |
| IEEE Instrumentation and Measurement <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | IM009 | 29.00 <input type="checkbox"/> | 14.50 <input type="checkbox"/> |
| IEEE Intelligent Transportation Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | ITSS038 | 35.00 <input type="checkbox"/> | 17.50 <input type="checkbox"/> |
| IEEE Magnetics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | MAG033 | 26.00 <input type="checkbox"/> | 13.00 <input type="checkbox"/> |
| IEEE Microwave Theory and Techniques <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | MTT017 | 17.00 <input type="checkbox"/> | 8.50 <input type="checkbox"/> |
| IEEE Nuclear and Plasma Sciences <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | NPS005 | 35.00 <input type="checkbox"/> | 17.50 <input type="checkbox"/> |
| IEEE Oceanic Engineering <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | OE022 | 19.00 <input type="checkbox"/> | 9.50 <input type="checkbox"/> |
| IEEE Photonics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | PHO036 | 34.00 <input type="checkbox"/> | 17.00 <input type="checkbox"/> |
| IEEE Power Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | PEL035 | 25.00 <input type="checkbox"/> | 12.50 <input type="checkbox"/> |
| IEEE Power & Energy <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | PE031 | 35.00 <input type="checkbox"/> | 17.50 <input type="checkbox"/> |
| IEEE Product Safety Engineering <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | PSE043 | 35.00 <input type="checkbox"/> | 17.50 <input type="checkbox"/> |
| IEEE Professional Communication <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | PC026 | 31.00 <input type="checkbox"/> | 15.50 <input type="checkbox"/> |
| IEEE Reliability <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | RL007 | 35.00 <input type="checkbox"/> | 17.50 <input type="checkbox"/> |
| IEEE Robotics and Automation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | RA024 | 9.00 <input type="checkbox"/> | 4.50 <input type="checkbox"/> |
| IEEE Signal Processing <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | SP001 | 22.00 <input type="checkbox"/> | 11.00 <input type="checkbox"/> |
| IEEE Social Implications of Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | SIT030 | 33.00 <input type="checkbox"/> | 16.50 <input type="checkbox"/> |
| IEEE Solid-State Circuits <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | SSC037 | 22.00 <input type="checkbox"/> | 11.00 <input type="checkbox"/> |
| IEEE Systems, Man, & Cybernetics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | SMC028 | 12.00 <input type="checkbox"/> | 6.00 <input type="checkbox"/> |
| IEEE Technology & Engineering Management <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | TEM014 | 35.00 <input type="checkbox"/> | 17.50 <input type="checkbox"/> |
| IEEE Ultrasonics, Ferroelectrics, & Frequency Control <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | UFFC020 | 20.00 <input type="checkbox"/> | 10.00 <input type="checkbox"/> |
| IEEE Vehicular Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> | VT006 | 18.00 <input type="checkbox"/> | 9.00 <input type="checkbox"/> |

Legend—Society membership includes:

- One or more Society publications
- Online access to publication
- Society newsletter
- CD-ROM of selected society publications

Complete both sides of this form, sign, and return to:
 IEEE MEMBERSHIP APPLICATION PROCESSING
 445 HOES LN, PISCATAWAY, NJ 08854-4141 USA
 or fax to +1 732 981 0225
 or join online at www.ieee.org/join

Please reprint your full name here

6 2016 IEEE Membership Rates (student rates available online)

IEEE member dues and regional assessments are based on where you live and when you apply. Membership is based on the calendar year from 1 January through 31 December. All prices are quoted in US dollars.

Please check the appropriate box.

| | BETWEEN 16 AUG 2015- 28 FEB 2016 PAY | BETWEEN 1 MAR 2016- 15 AUG 2016 PAY |
|----------------------------------|---|--|
| RESIDENCE | | |
| United States..... | \$197.00 <input type="checkbox"/> | \$98.50 <input type="checkbox"/> |
| Canada (GST)*..... | \$173.35 <input type="checkbox"/> | \$86.68 <input type="checkbox"/> |
| Canada (NB, NF and ON HST)*..... | \$185.11 <input type="checkbox"/> | \$92.56 <input type="checkbox"/> |
| Canada (Nova Scotia HST)*..... | \$188.05 <input type="checkbox"/> | \$94.03 <input type="checkbox"/> |
| Canada (PEI HST)*..... | \$186.58 <input type="checkbox"/> | \$93.29 <input type="checkbox"/> |
| Canada (GST and QST Quebec)..... | \$188.01 <input type="checkbox"/> | \$94.01 <input type="checkbox"/> |
| Africa, Europe, Middle East..... | \$160.00 <input type="checkbox"/> | \$80.00 <input type="checkbox"/> |
| Latin America..... | \$151.00 <input type="checkbox"/> | \$75.50 <input type="checkbox"/> |
| Asia, Pacific..... | \$152.00 <input type="checkbox"/> | \$76.00 <input type="checkbox"/> |

*IEEE Canada Business No. 125634188

Minimum Income or Unemployed Provision

Applicants who certify that their prior year income did not exceed US\$14,700 (or equivalent) or were not employed are granted 50% reduction in: full-year dues, regional assessment and fees for one IEEE Membership plus one Society Membership. If applicable, please check appropriate box and adjust payment accordingly. Student members are not eligible.

- I certify I earned less than US\$14,700 in 2015
- I certify that I was unemployed in 2015

7 More Recommended Options

- Proceedings of the IEEE..... print \$47.00 or online \$41.00
- Proceedings of the IEEE (print/online combination).....\$57.00
- IEEE Standards Association (IEEE-SA).....\$53.00
- IEEE Women in Engineering (WIE).....\$25.00

8 Payment Amount

Please total the Membership dues, Society dues, and other amounts from this page:

- IEEE Membership dues ⑥.....\$ _____
- IEEE Society dues (optional) ⑤.....\$ _____
- IEEE-SA/WIE dues (optional) ⑦.....\$ _____
- Proceedings of the IEEE (optional) ⑦.....\$ _____
- Canadian residents pay 5% GST or appropriate HST (BC—12%; NB, NF, ON—13%;NS—15%) on Society payments & publications only.....TAX \$ _____

AMOUNT PAID **TOTAL \$** _____

Payment Method

All prices are quoted in US dollars. You may pay for IEEE membership by credit card (see below), check, or money order payable to IEEE, drawn on a US bank.

- Check
-
-
-
-

Credit Card Number

MONTH YEAR CARDHOLDER'S 5-DIGIT ZIP/POD (BILLING STATEMENT ADDRESS) USA ONLY

Name as it appears on card _____

Signature _____

- Auto Renew my Memberships and Subscriptions (available when paying by credit card).
- I agree to the Terms and Conditions located at www.ieee.org/autorenew

9 Were You Referred to IEEE?

- Yes No If yes, provide the following:
- Member Recruiter Name _____
- IEEE Recruiter's Member Number (Required) _____

CAMPAIGN CODE _____ PROMO CODE _____

Information for Authors

(Updated/Effective January 2015)

For Transactions and Journals:

Authors are encouraged to submit manuscripts of Regular papers (papers which provide a complete disclosure of a technical premise), or Comment Correspondences (brief items that provide comment on a paper previously published in these TRANSACTIONS).

Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere.

Every manuscript must:

- i. provide a clear statement of the problem and what the contribution of the work is to the relevant research community;
- ii. state why this contribution is significant (what impact it will have);
- iii. provide citation of the published literature most closely related to the manuscript; and
- iv. state what is distinctive and new about the current manuscript relative to these previously published works.

By submission of your manuscript to these TRANSACTIONS, all listed authors have agreed to the authorship list and all the contents and confirm that the work is original and that figures, tables and other reported results accurately reflect the experimental work. In addition, the authors all acknowledge that they 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. The corresponding author is responsible for obtaining consent from all co-authors and, if needed, from sponsors before submission.

In order to be considered for review, a paper must be within the scope of the journal and represent a novel contribution. A paper is a candidate for an Immediate Rejection if it is of limited novelty, e.g. a straightforward combination of theories and algorithms that are well established and are repeated on a known scenario. Experimental contributions will be rejected without review if there is insufficient experimental data. These TRANSACTIONS are published in English. Papers that have a large number of typographical and/or grammatical errors will also be rejected without review.

In addition to presenting a novel contribution, acceptable manuscripts must describe and cite related work in the field to put the contribution in context. Do not give theoretical derivations or algorithm descriptions that are easily found in the literature; merely cite the reference.

New and revised manuscripts should be prepared following the "Manuscript Submission" guidelines below, and submitted to the online manuscript system, ScholarOne Manuscripts. 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 Manuscript system.

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-column, single-spaced format using a font size of 10 points or larger, having a margin of at least 1 inch on all sides. Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOne-Manuscripts site. Since many reviewers prefer a larger font, you are strongly encouraged to also submit a single-column, double-spaced version (11 point font or larger), which is easy to create with the templates provided **IEEE Author Digital Toolbox** (http://www.ieee.org/publications_standards/publications/authors/authors_journals.html). Page length restrictions will be determined by the double-column

version. 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 .PDF 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. *EDICS (Not applicable to Journal of Selected Topics in Signal Processing).* All submissions must be classified by the author with an EDICS (Editors' Information Classification Scheme) selected from the list of EDICS published online at the at the publication's EDICS webpage (*please see the list below). 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.
4. *Additional Documents for Review.* Please upload pdf versions of all items in the reference list that are not publicly available, such as unpublished (submitted) papers. Graphical abstracts and supplemental materials intended to appear with the final paper (see below) must also be uploaded for review at the time of the initial submission for consideration in the review process. 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. *Supplemental Materials.* IEEE Xplore can publish multimedia files (audio, images, video), datasets, and software (e.g. 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 IEEE Author Digital Toolbox under "Multimedia." To make your work reproducible by others, these TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper.
6. *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.
7. *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.

Comment Correspondence. Comment Correspondences provide brief comments on material previously published in these TRANSACTIONS. These items may not exceed 2 pages in double-column, single spaced format, using 9 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 "Manuscript Submission" above for instructions).

Authors may also submit manuscripts of overview articles, but note that these include an additional white paper approval process <http://www.signalprocessingsociety.org/publications/overview-articles/>. [This does not apply to the Journal of Selected Topics in Signal Processing. Please contact the Editor-in-Chief.]

Digital Object Identifier

Manuscript Length. For the initial submission of a regular paper, the manuscript may not exceed 13 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. Supplemental materials and graphical abstracts are not included in the page count. For regular papers, the revised manuscript may not exceed 16 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. For Overview Papers, the maximum length is double that for regular submissions at each stage (please reference <http://www.signalprocessingsociety.org/publications/overview-articles/> for more information).

Note that any paper in excess of 10 pages will be subject to mandatory overlength page charges. 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. Note: Papers submitted to the TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.

Exceptions to manuscript length requirements may, under extraordinary circumstances, be granted by the Editor-in-Chief. However, such exception does not obviate your requirement to pay any and all overlength or additional charges that attach to the manuscript.

Resubmission of Previously Rejected Manuscripts. Authors of manuscripts rejected from any journal are allowed to resubmit their manuscripts only once. At the time of submission, you will be asked whether your manuscript is a new submission or a resubmission of an earlier rejected manuscript. If it is a resubmission of a manuscript previously rejected by any journal, you are expected to submit supporting documents identifying the previous submission and detailing how your new version addresses all of the reviewers' comments. Papers that do not disclose connection to a previously rejected paper or that do not provide documentation as to changes made may be immediately rejected.

Author Misconduct. Author misconduct includes plagiarism, self-plagiarism, and research misconduct, including falsification or misrepresentation of results. All forms of misconduct are unacceptable and may result in sanctions and/or other corrective actions. 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 author's previously published work that also involves other authors. Self-plagiarism involves the verbatim copying or reuse of an authors own prior work without appropriate citation, including 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. For more information on the definitions, investigation process, and corrective actions related to author misconduct, see the Signal Processing Society Policies and Procedures Manual, Section 6.1. <http://www.signalprocessingsociety.org/about-sps/governance/policy-procedure/part-2>. Author misconduct may also be actionable by the IEEE under the rules of Member Conduct.

Extensions of the Author's Prior 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 their related prior work; the papers cannot be identical; and the journal publication must include substantively novel aspects such as new experimental results and analysis or added theoretical work. The journal publication should clearly specify how the journal paper offers novel contributions when citing the prior work. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper, and the prior work must be cited as the primary source.

Submission Format. Authors are required to prepare manuscripts employing the on-line style files developed by IEEE, which include guidelines for abbreviations, mathematics, and graphics. 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 the **IEEE Author Digital Toolbox** under "Template for all TRANSACTIONS." (LaTeX and MS Word). Please note the following requirements about the abstract:

- The abstract must be a concise yet comprehensive reflection of what is in your article.
- The abstract must be self-contained, without abbreviations, footnotes, displayed equations, or references.

- The abstract must be between 150-250 words.
- The abstract should include a few keywords or phrases, as this will help readers to find it. Avoid over-repetition of such phrases as this can result in a page being rejected by search engines.

In addition to written abstracts, papers may include a graphical abstract; see http://www.ieee.org/publications_standards/publications/authors/authors_journals.html for options and format requirements.

IEEE supports the publication of author names in the native language alongside the English versions of the names in the author list of an article. For more information, see "Author names in native languages" (http://www.ieee.org/publications_standards/publications/authors/auth_names_native_lang.pdf) on the IEEE Author Digital Toolbox page.

Open Access. The publication is a hybrid journal, allowing either Traditional manuscript submission or Open Access (author-pays OA) manuscript submission. Upon submission, if you choose to have your manuscript be an Open Access article, you commit to pay the discounted \$1,750 OA fee if your manuscript is accepted for publication in order to enable unrestricted public access. Any other application charges (such as overlength page charge and/or charge for the use of color in the print format) will be billed separately once the manuscript formatting is complete but prior to the publication. If you would like your manuscript to be a Traditional submission, your article will be available to qualified subscribers and purchasers via IEEE Xplore. No OA payment is required for Traditional submission.

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 (two 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. (**NOTE: Papers accepted to IEEE TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.) 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. No mandatory overlength page charges will be applied to overview articles in the Society's journals.

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 charges for print are the responsibility of the author. Details of the associated charges can be found on the IEEE Publications page.

Payment of fees on color reproduction is not negotiable or voluntary, and the author's agreement to publish the manuscript in these TRANSACTIONS is considered acceptance of this requirement.

*EDICS Webpages:

IEEE TRANSACTIONS ON SIGNAL PROCESSING:

<http://www.signalprocessingsociety.org/publications/periodicals/tsp/TSP-EDICS/>

IEEE TRANSACTIONS ON IMAGE PROCESSING:

<http://www.signalprocessingsociety.org/publications/periodicals/image-processing/tip-edics/>

IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE / ACM:

<http://www.signalprocessingsociety.org/publications/periodicals/taslp/taslp-edics/>

IEEE TRANSACTIONS ON INFORMATION, FORENSICS AND SECURITY:

<http://www.signalprocessingsociety.org/publications/periodicals/forensics/forensics-edics/>

IEEE TRANSACTIONS ON MULTIMEDIA:

<http://www.signalprocessingsociety.org/tmm/tmm-edics/>

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING:

<http://www.signalprocessingsociety.org/publications/periodicals/tci/tci-edics/>

IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS:

<http://www.signalprocessingsociety.org/publications/periodicals/tsipn/tsipn-edics/>

2016 IEEE SIGNAL PROCESSING SOCIETY MEMBERSHIP APPLICATION

Mail to: IEEE OPERATIONS CENTER, ATTN: Matthew Plotner, 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



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: _____

2016 SPS MEMBER RATES

16 Aug-28 Feb 1 Mar-15 Aug

Pay Full Year Pay Half Year

Signal Processing Society Membership Fee* \$ 22.00 \$ 11.00

Fee includes: IEEE Signal Processing Magazine (electronic and digital), Inside Signal Proc. eNewsletter (electronic) and IEEE Signal Processing Society Content Gazette (electronic).

Add \$17 to enhance SPS Membership and also receive: \$17.00 \$ 8.50

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 \$209.00 \$104.50

Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on: Print \$160.00 \$ 80.00

Image Processing, IEEE Transactions on: Print \$207.00 \$103.50

Information Forensics and Security, IEEE Trans. on: Print \$179.00 \$ 89.50

IEEE Journal of Selected Topics in Signal Processing: Print \$176.00 \$ 88.00

Affective Computing, IEEE Transactions on: Electronic \$ 36.00 \$ 18.00

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

IEEE Trans. on Cognitive Comm. & Networking Electronic \$ 27.00 \$ 13.50

IEEE Trans. on Computational Imaging Electronic \$ 28.00 \$ 14.00

IEEE Trans. on Big Data Electronic \$ 26.00 \$ 13.00

IEEE Trans. on Molecular, Biological, & Multi-scale Communications Electronic \$ 25.00 \$ 12.50

IEEE Internet of Things Journal Electronic \$ 26.00 \$ 13.00

IEEE Trans. on Cloud Computing Electronic \$ 43.00 \$ 21.50

IEEE Trans. on Computational Social Systems Electronic \$ 30.00 \$ 15.00

IEEE Trans. on Signal & Info Proc. Over Networks Electronic \$ 28.00 \$ 14.00

IEEE Biometrics Compendium: Online \$ 30.00 \$ 15.00

Computing in Science & Engrg. Mag.: Electronic and Digital \$ 39.00 \$ 19.50

Print \$ 69.00 \$ 34.50

Medical Imaging, IEEE Transactions on: Print \$ 74.00 \$ 37.00

Electronic \$ 53.00 \$ 26.50

Print & Electronic \$ 89.00 \$ 44.50

Mobile Computing, IEEE Transactions on: ELE/Print Abstract/CD-ROM \$ 41.00 \$ 20.50

Multimedia, IEEE Transactions on: Electronic \$ 43.00 \$ 21.50

IEEE MultiMedia Magazine: Electronic and Digital \$ 39.00 \$ 19.50

Print \$ 69.00 \$ 34.50

Network Science and Engrg., IEEE Trans. on: Electronic \$ 34.00 \$ 17.00

IEEE Reviews in Biomedical Engineering: Print \$ 25.00 \$ 12.50

Print & Electronic \$ 40.00 \$ 20.00

IEEE Security and Privacy Magazine: Electronic and Digital \$ 39.00 \$ 19.50

Print \$ 69.00 \$ 34.50

IEEE Sensors Journal: Electronic \$ 40.00 \$ 20.00

Smart Grid, IEEE Transactions on: Print \$100.00 \$ 50.00

Electronic \$ 40.00 \$ 20.00

Print & Electronic \$120.00 \$ 60.00

Wireless Communications, IEEE Transactions on: Print \$127.00 \$ 63.50

Electronic \$ 49.00 \$ 24.50

Print & Electronic \$127.00 \$ 63.50

IEEE Wireless Communications Letters: Electronic \$ 19.00 \$ 9.50

*IEEE membership required or requested
 Affiliate application to join SP Society only. Amount Paid \$ _____

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 _____ Male Female

Day _____ Month _____ Year _____

7. CONTACT INFORMATION

Office Phone/Office Fax _____ Home Phone/Home Fax _____

Office E-Mail _____ Home E-Mail _____

8. 2016 IEEE MEMBER RATES

| IEEE DUES | 16 Aug-14 Feb 15 | 1 Mar-15 Aug 15 |
|--------------------------------------|-----------------------------------|----------------------------------|
| Residence | Pay Full Year | Pay Half Year** |
| United States | \$197.00 <input type="checkbox"/> | \$98.50 <input type="checkbox"/> |
| Canada (incl. GST) | \$173.35 <input type="checkbox"/> | \$86.68 <input type="checkbox"/> |
| Canada (incl. HST for PEI) | \$186.58 <input type="checkbox"/> | \$93.29 <input type="checkbox"/> |
| Canada (incl. HST for Nova Scotia) | \$198.05 <input type="checkbox"/> | \$99.03 <input type="checkbox"/> |
| Canada (incl. HST for NB, NF and ON) | \$185.11 <input type="checkbox"/> | \$92.56 <input type="checkbox"/> |
| Canada (incl. GST and QST Quebec) | \$188.01 <input type="checkbox"/> | \$94.01 <input type="checkbox"/> |
| Africa, Europe, Middle East | \$160.00 <input type="checkbox"/> | \$80.00 <input type="checkbox"/> |
| Latin America | \$151.00 <input type="checkbox"/> | \$75.50 <input type="checkbox"/> |
| Asia, Pacific | \$152.00 <input type="checkbox"/> | \$76.00 <input type="checkbox"/> |

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 12%, 13% 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.



9. IEEE Membership Affiliate Fee (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 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): _____

