

RECENT ADVANCES AND CHALLENGES ON BIG DATA ANALYSIS IN NEUROIMAGING

EDITED BY: Jian Kang, Brian Caffo and Han Liu
PUBLISHED IN: Frontiers in Neuroscience





frontiers

Frontiers Copyright Statement

© Copyright 2007-2017 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-128-9

DOI 10.3389/978-2-88945-128-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

RECENT ADVANCES AND CHALLENGES ON BIG DATA ANALYSIS IN NEUROIMAGING

Topic Editors:

Jian Kang, University of Michigan, USA

Brian Caffo, Johns Hopkins University, USA

Han Liu, Princeton University, USA

Big data is revolutionizing our ability to measure and study the human brain. New technology increases the resolution of images that are being study as well as enables researchers to study the brain as it functions. These technological advances are combined with efforts to collect neuroimaging data on large numbers of subjects, in some cases longitudinally. This combination of advances in measurement and scope of studies requires novel development in the statistical analysis. Fast, scalable, robust and accurate models and approaches need to be developed to make headway on these problems. This volume represents a unique collection of researchers providing deep insights on the statistical analysis of big neuroimaging data.

Citation: Kang, J., Caffo, B., Liu, H., eds. (2017). Recent Advances and Challenges on Big Data Analysis in Neuroimaging. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-128-9

Table of Contents

05	<i>Editorial: Recent Advances and Challenges on Big Data Analysis in Neuroimaging</i>
	Jian Kang, Brian Caffo and Han Liu
08	<i>Detecting the subtle shape differences in hemodynamic responses at the group level</i>
	Gang Chen, Ziad S. Saad, Nancy E. Adleman, Ellen Leibenluft and Robert W. Cox
26	<i>An empirical Bayes normalization method for connectivity metrics in resting state fMRI</i>
	Shuo Chen, Jian Kang and Guoqing Wang
37	<i>A Parcellation Based Nonparametric Algorithm for Independent Component Analysis with Application to fMRI Data</i>
	Shanshan Li, Shaojie Chen, Chen Yue and Brian Caffo
49	<i>Identification of Voxels Confounded by Venous Signals Using Resting-State fMRI Functional Connectivity Graph Community Identification</i>
	Klaudius Kalcher, Roland N. Boubela, Wolfgang Huf, Christian Našel and Ewald Moser
58	<i>The Voxel-Wise Functional Connectome Can Be Efficiently Derived from Co-activations in a Sparse Spatio-Temporal Point-Process</i>
	Enzo Tagliazucchi, Michael Siniatchkin, Helmut Laufs and Dante R. Chialvo
71	<i>Mixed Effects Models for Resampled Network Statistics Improves Statistical Power to Find Differences in Multi-Subject Functional Connectivity</i>
	Manjari Narayan and Genevera I. Allen
90	<i>An Efficient and Reliable Statistical Method for Estimating Functional Connectivity in Large Scale Brain Networks Using Partial Correlation</i>
	Yikai Wang, Jian Kang, Phebe B. Kemmer and Ying Guo
107	<i>Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fMRI data</i>
	Yuting Xu and Martin A. Lindquist
126	<i>Big Data Approaches for the Analysis of Large-Scale fMRI Data Using Apache Spark and GPU Processing: A Demonstration on Resting-State fMRI Data from the Human Connectome Project</i>
	Roland N. Boubela, Klaudius Kalcher, Wolfgang Huf, Christian Našel and Ewald Moser
134	<i>Statistical image analysis of longitudinal RAVENS images</i>
	Seonjoo Lee, Vadim Zipunnikov, Daniel S. Reich and Dzung L. Pham

- 148** *Reconstruction of human brain spontaneous activity based on frequency-pattern analysis of magnetoencephalography data*
Rodolfo R. Llinás, Mikhail N. Ustinin, Stanislav D. Rykunov, Anna I. Boyko, Vyacheslav V. Sychev, Kerry D. Walton, Guilherme M. Rabello and John Garcia
- 156** *An exploratory data analysis of electroencephalograms using the functional boxplots approach*
Duy Ngo, Ying Sun, Marc G. Genton, Jennifer Wu, Ramesh Srinivasan, Steven C. Cramer and Hernando Ombao
- 174** *Boosting brain connectome classification accuracy in Alzheimer's disease using higher-order singular value decomposition*
Liang Zhan, Yashu Liu, Yalin Wang, Jiayu Zhou, Neda Jahanshad, Jieping Ye and Paul M. Thompson
- 185** *Multimodal Imaging Signatures of Parkinson's Disease*
F. DuBois Bowman, Daniel F. Drake and Daniel E. Huddleston



Editorial: Recent Advances and Challenges on Big Data Analysis in Neuroimaging

Jian Kang^{1*}, Brian Caffo² and Han Liu³

¹ Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA, ² Department of Biostatistics, Johns Hopkins University, Baltimore, MD, USA, ³ Department of Operations Research and Financial Engineering, Princeton University, New Jersey, NJ, USA

Keywords: neuroimaging, big data analytics

The Editorial on the Research Topic

Recent Advances and Challenges on Big Data Analysis in Neuroimaging

“... the most powerful computer in the world isn’t nearly as intuitive as the one we’re born with. So there is this enormous mystery waiting to be unlocked.”

—President Obama Announcing the BRAIN Initiative

In its Big Data to Knowledge initiative, the US National Institutes of Health notes the wealth of biomedical and behavioral information will greatly advance our understanding of human health, disease and treatment—only if new analytic tools are developed and the understanding of these new tools is broadly disseminated (<https://datascience.nih.gov/bd2k>). Big Data encompasses the study of data formats from long, in the sense of multitudes of subjects, and wide, in the sense of complex measurements across relatively few subjects. Brain imaging tends to be of latter category. However, it is essential for our field to prepare for the inevitability of both long and wide neuroimaging data.

The stakes couldn’t be higher, as the promise of Big Data in neuroscience seems limitless. Recent advances in neuroimaging technology offers great hope for significant progress in furthering the understanding the human brain, with the potential to facilitate research in medicine, neuroscience, psychology, and many other disciplines. This technology enables the creation of massive amounts of high-resolution images, which capture the structure, function and composition of human brains. Parallels to brain imaging are often made with the scope, scale, scientific goals and importance of mapping and analyzing the human genome, and other “biomes” (proteome, transcriptome, microbiome). In fact, intra-brain structural and functional connections have their own portmanteau, the so called “connectome” (genome and connection). The implication of myriad of these new disciplines, including brain imaging, is the central idea of the measurement of the intrinsic, unique, fundamental, and personal measurements that will make true precision medicine a reality.

However, such breakthroughs in the development of effective personalized treatments of neurological and psychiatric disease require a massive effort in the: Measurement, informatics and analytic capacity to handle the large databases of subjects, increasingly fine temporal and spatial measures, and multiple technologies. To elaborate, the 100 billion neurons in the human brain, their trillions of structural and functional connections, glial structure, lesions and the electro-chemical function of the brain are captured through lenses of varying measurement types. The resulting images generate massive amounts of data so that even storage and representation of these data raise significant challenges. Furthermore, since the measurements capture the brain at multiple spatial and temporal scales, with different functional, structural, and compositional targets, the ability to synthesize this information is of fundamental importance for progress in understanding the brain and its pathologies. The term “big data” in this area encompasses this

OPEN ACCESS

Edited and reviewed by:

Jean-Baptiste Poline,
University of California, Berkeley, USA

*Correspondence:

Jian Kang
jiangkang@umich.edu

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 23 September 2016

Accepted: 24 October 2016

Published: 15 November 2016

Citation:

Kang J, Caffo B and Liu H (2016)
Editorial: Recent Advances and
Challenges on Big Data Analysis in
Neuroimaging.
Front. Neurosci. 10:505.
doi: 10.3389/fnins.2016.00505

intersection of data size, complexity and modalities. Thus, efficient analysis and process of big data and the development of high-performance computing tools is critical for modern neuroscientific studies.

Despite many existing successful efforts in the analysis of large neuroimaging datasets, there remains ample room for new methods to meet these challenges. In this Frontiers research topic, we selected 14 excellent research articles that present statistical challenges and/or proposed new approaches for dealing with neuroimaging big data.

The issue boasts of a total of 60 contributors, having a wealth of experience in the area and diverse backgrounds, including: Statisticians, neuroscientists, psychologists, and computer scientists. Their insights brought statistical and computational innovations to make significant progress on the most important questions in neuroimaging. Below we provide a brief overview of all the articles in this research topic.

Functional connectomics being a fundamental area for studying neural communications represents a focus of the issue, with a wide range of topics for studying the functional connectome using resting state fMRI (R-fMRI) data. In particular, Boubela et al. have developed parallel computing algorithms and efficient implementations using apache spark and graphical processing unit (GPU) techniques for analyzing big R-fMRI data. These computational tools are quite useful for scalable analysis of very large neuroimaging datasets. Chen et al.; (Bowman et al.) have proposed a novel empirical Bayes method to normalize functional brain connectivity metrics on a posterior probability scale. This method can facilitate appropriate quantifications of existing connectivity metrics and produce reproducible scientific findings. Kalcher et al. concentrated on an interesting and important problem: Identifying venous voxels in R-fMRI data in order to increase the specificity of fMRI analyses to microvasculature in the vicinity of the neural processes triggering the blood oxygenation level dependent (BOLD) response. They solved this challenging problem by applying a graph based clustering algorithm on thresholded correlation graphs. Wang et al. studied the difference between correlation-based graphs and partial correlation based graphs in terms of estimating functional connectivity using R-fMRI data. They have developed an efficient and reliable statistical procedure based on the constrained L1-minimization Approach (CLIME) in large-scale brain networks for single subject fMRI data analysis. They also have proposed a new *Dens*-based selection method that provides a more informative and a flexible tool to allow the users to select the tuning parameter based on the desired sparsity level. For the analysis of multiple subject fMRI data, Narayan and Allen defined functional connectivity using Gaussian graphical models. They proposed a mixed-effects model that treats both subject level networks and population level covariate effects as unknown parameters. They adopted resampling based methods to improve the power for detecting the differences in multi-subject functional connectivity. Adopting an alternative modeling approach for the brain network, Li et al. have proposed to use a non-parametric independent component analysis (ICA) to separate the latent source signals from the R-fMRI data. Their novel ICA algorithm is based on density estimation

and maximum likelihood, where the densities of the signals are estimated via p-spline based histogram smoothing and the mixing matrix is simultaneously estimated using an optimization algorithm. The proposed approach is very straightforward to implement and shows good performance for recovering the established brain networks. The dynamic nature of the functional connectivity was studied by Xu and Lindquist. They introduced a new data-driven algorithm to detect temporal change points in the functional connectivity and estimate a graph between region of interests (ROIs) by adopting a sparse matrix estimation approach and a hypothesis testing procedure to determine change points. This is referred as the Dynamic Connectivity Detection (DCD) algorithm which improves the recently developed Dynamic Connectivity Regression (DCR) algorithm in terms of computational efficiency and scalability for the large-scale data analysis.

In addition to the R-fMRI data analysis, the research topic also includes a new statistical approach to detecting subtle shape differences in the hemodynamic response at the group level in the fMRI studies (Chen et al.). This method estimates the shape features of hemodynamic response function using multiple basis functions and new dimension reduction methods. It is useful for improving the statistical power in detecting the brain activity signals at both the individual level and the group level.

In addition to the problems in the functional magnetic resonance imaging (fMRI) (Boubela et al.; Bowman et al.; Chen et al.; Chen et al.; Kalcher et al.; Li et al.; Narayan and Allen; Tagliazucchi et al.; Wang et al.; Xu and Lindquist), our research topic also covers a variety of other imaging modalities, such as structural magnetic resonance imaging (sMRI) (Lee et al.; Zhan et al.), diffusion tensor imaging (DTI) (Bowman et al.), magnetoencephalography (MEG) (Llinás et al.) and electroencephalograms (EEG) (Ngo et al.). Among those, Bowman et al. presented a statistical framework for analyzing neuroimaging data from multiple modalities to identify important biomarkers for Parkinson's disease (PD) risks. Their approach builds on the elastic net, performing regularization and variable selection with introducing additional criteria for parsimony and reproducibility.

Focusing on another progressive brain disease, the Alzheimer's disease (AD), Zhan et al. developed new methods to model brain structural networks from diffusion MRI and proposed a novel feature extraction and classification framework based on higher order singular value decomposition and the sparse logistic regression approach.

For the study of brain morphometry, Lee et al. developed new statistical approaches for the longitudinal regional analysis of volumes examined in normalized space (RAVENS). The method is a variant of the longitudinal functional principal component analysis (LFPCA) for high-dimensional images, which can separate registration errors from other longitudinal changes and baseline patterns, and thus address the limitations of the existing methods. Many statistical methods and computational algorithms have been developed for fMRI and MRI data analysis, limited statistical methods have been proposed to address the MEG analysis. Along this direction, we have included one article that focuses on frequency-pattern analysis of MEG data

to reconstruct the brain spontaneous activities (Llinás et al.). The proposed method is among the very first to successfully characterize brain electrical activities and localize the sources in anatomical brain space in combination with MRI data. In addition to the systematic statistical approaches for analysis of big neuroimaging, we also include an exploratory data analysis approach to EEG data: The functional boxplots approach. It analyzes log periodograms of EEG time series data in the spectral domain. It identifies a functional median, summarizes variability, and detects potential outliers.

In summary, our research topic has collected a series of new statistical approaches to addressing important questions in neuroimaging big data analyses from statistically efficient, computationally scalable and scientifically meaningful perspectives. It covers a broad range of imaging modalities, including fMRI, sMRI, dMRI, DTI, EEG, and MEG. It studies a variety of mental health diseases, including Parkinson's, autism spectrum disease, Alzheimer's and multiple sclerosis.

We hope that this issue will spur discussion and open a forum for statisticians, computer scientists, neuroscientists

and psychologists to further contribute the innovations in this important topic.

AUTHOR CONTRIBUTIONS

All authors contributed equally to this content.

FUNDING

BC was supported by NIH grants RO1 EB012547 and P41 EB015909 and JK was supported by NIH R01MH105561.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Kang, Caffo and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Detecting the subtle shape differences in hemodynamic responses at the group level

Gang Chen^{1*}, Ziad S. Saad¹, Nancy E. Adleman², Ellen Leibenluft³ and Robert W. Cox¹

¹ Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA, ² Department of Psychology, The Catholic University of America, Washington, DC, USA, ³ Section on Bipolar Spectrum Disorders, Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA

OPEN ACCESS

Edited by:

Brian Caffo,
Johns Hopkins University, USA

Reviewed by:

Xi-Nian Zuo,
Chinese Academy of Sciences, China
Théodore Papadopoulos,
French Institute for Research in
Computer Science and Automation,
France

*Correspondence:

Gang Chen
gangchen@mail.nih.gov

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 22 May 2015

Accepted: 28 September 2015

Published: 26 October 2015

Citation:

Chen G, Saad ZS, Adleman NE,
Leibenluft E and Cox RW (2015)
Detecting the subtle shape differences
in hemodynamic responses at the
group level. *Front. Neurosci.* 9:375.
doi: 10.3389/fnins.2015.00375

The nature of the hemodynamic response (HDR) is still not fully understood due to the multifaceted processes involved. Aside from the overall amplitude, the response may vary across cognitive states, tasks, brain regions, and subjects with respect to characteristics such as rise and fall speed, peak duration, undershoot shape, and overall duration. Here we demonstrate that the fixed-shape (FSM) or adjusted-shape (ASM) methods may fail to detect some shape subtleties (e.g., speed of rise or recovery, or undershoot). In contrast, the estimated-shape method (ESM) through multiple basis functions can provide the opportunity to identify some subtle shape differences and achieve higher statistical power at both individual and group levels. Previously, some dimension reduction approaches focused on the peak magnitude, or made inferences based on the area under the curve (AUC) or interaction, which can lead to potential misidentifications. By adopting a generic framework of multivariate modeling (MVM), we showcase a hybrid approach that is validated by simulations and real data. With the whole HDR shape integrity maintained as input at the group level, the approach allows the investigator to substantiate these more nuanced effects through the unique HDR shape features. Unlike the few analyses that were limited to main effect, two- or three-way interactions, we extend the modeling approach to an inclusive platform that is more adaptable than the conventional GLM. With multiple effect estimates from ESM for each condition, linear mixed-effects (LME) modeling should be used at the group level when there is only one group of subjects without any other explanatory variables. Under other situations, an approximate approach through dimension reduction within the MVM framework can be adopted to achieve a practical equipoise among representation, false positive control, statistical power, and modeling flexibility. The associated program 3dMVM is publicly available as part of the AFNI suite.

Keywords: hemodynamic response, basis function, multivariate general linear model, linear mixed-effects model, FMRI group analysis, AFNI

INTRODUCTION

When a region in the brain is activated, oxygen and glucose demands lead to blood vessel dilation, followed by increased blood to the tissue (neurons and astrocytes) under stress. The onset of a neuronal activity triggers a sequence of physiological events in the blood vessels of the surrounding area, typically characterized by the changes in cerebral blood flow as well as concentration fluctuations of deoxyhemoglobin and oxyhemoglobin. The blood oxygenation level dependent (BOLD) signal from the fMRI scanning mainly captures the concentration changes of deoxyhemoglobin; that is, the BOLD signal is a surrogate and signature of neuronal activations plus various sources of noise (e.g., physiological and random fluctuations). As an indirect measure of neuronal activity, the shape of the BOLD response may hold some crucial features about brain function. However, the cascade of events from neural activation to measurable MRI signal is complex and nonlinear under certain regimes (Friston et al., 1998b; Birn et al., 2001; Logothetis and Wandell, 2004; Logothetis, 2008; Magri et al., 2012): Even though the BOLD response is simply interpreted as changes in neuronal processing, the same neuronal activity may evoke different hemodynamic response (HDR) shape across trials, regions, conditions/tasks, subjects, and groups. For example, neurophysiological confounds such as neurovascular coupling or energy consumption changes could lead to different BOLD response features, potentially explaining the HDR variability in magnitude and shape across brain regions, cognitive conditions and populations (e.g., children with autism vs. controls, Reynell and Harris, 2013). Nevertheless, meaningful interpretation as well as detection power in fMRI data analysis may depend on the accurate modeling of the BOLD response both at the individual subject and group levels (e.g., Buxton et al., 2004; Handwerker et al., 2004; Stephen et al., 2007; Barbé et al., 2012; Badillo et al., 2013).

Under an experimentally-manipulated situation, the subject typically performs some tasks or is put under certain conditions in an event-related design, with each trial lasting for 2 s or less, and the HDR to each trial can be mathematically characterized by an impulse response function (IRF) that corresponds to a stimulus with a theoretically instantaneous duration and unit intensity. The voxel-wise EPI signal is then modeled through time series regression with explanatory variables (or regressors) of interest, each of which is constructed through the convolution between the stimulus timing and the IRF. In a block design, each task or condition has a duration of more than two seconds. As each block can be approximately considered as a sequence of events with an interval of scanning repetition time (TR), the theoretical HDR is usually hypothesized as the integral or linear summation of the consecutive IRFs, or the convolution of IRF over the stimulus duration.

We typically adopt some formative mathematical functions (usually called HDR functions or HRFs) to approximate the HDR based on the experimental data with the assumption of linearity and time-invariance (or stationarity) (Marrelec et al., 2003), and consider three common approaches to modeling the average HDR across trials. The first one presumes a fixed

shape IRF (e.g., gamma variate or wave form in AFNI, Cohen, 1997; canonical IRF in SPM, FSL, and NIPY, Friston et al., 1998a). With this model-based or fixed-shape method (FSM), the regression coefficient or β associated with each condition in the individual subject analysis reflects the major HDR magnitude (e.g., percent signal change). The second approach makes no assumption about the IRF's shape and estimates it with a set of basis functions. The number of basis functions varies depending on the kernel set and the duration over which the response is being modeled. A common approach to this estimated-shape method (ESM) consists of using a set of equally-spaced TENT (piecewise linear) functions or linear splines, and each of the resulting regression coefficient represents an estimate of the response amplitude at some time after stimulus onset. Regardless of the kernel set, however, ESM generates the same number of regressors as the number of basis functions (e.g., m) per condition or task, resulting in m regression coefficients which need to be considered simultaneously at the group level. In addition to the aforementioned TENT basis set, options for ESM at the voxel level include cubic splines, Legendre polynomials, sines, or user-defined functions in AFNI, and finite impulse function (FIR) in SPM, FSL, and NIPY, inverse logit (Lindquist et al., 2009), and high-order B-splines (Degras and Lindquist, 2014). In addition, the python package PyHRF offers an ESM at the parcel level through the joint detection-estimation framework (Vincent et al., 2014). It is of note that one significant advantage of adopting basis functions such as TENT or cubic splines is the flexibility of creating regressors through piecewise interpolation when the stimulus onset times are not aligned with the TR grids (e.g., the acquisition time is shorter than TR if one wants to present "silent trials" as a control condition to speech or other auditory stimulus). The third approach lies between the two extremes of FSM and ESM, and uses a set of two or three basis functions (Friston et al., 1998b). In this adjusted-shape method (ASM), the first basis (canonical IRF) captures the major HDR shape, and the second basis, the time derivative of the canonical IRF, provides some flexibility in modeling the delay or time-to-peak, while the third basis, dispersion curve (derivative relative to the dispersion parameter in the canonical IRF), allows the peak duration to vary.

With one parameter per condition, FSM is the most efficient¹ and statistically powerful among the three, if the presumed shape is reasonably close to the ground truth, and the group analysis strategies have been developed to reasonable maturity: The β values at the individual level are typically brought to the group level using the Student's t -test, permutation tests (Nichols and Holmes, 2002; Dehaene-Lambertz et al., 2006; Mériaux et al., 2006; Winkler et al., 2014), AN(C)OVA, general linear model (GLM) (Poline and Brett, 2012), multivariate modeling (MVM) (Chen et al., 2014), linear mixed-effects (LME) method (Bernal-Rusiel et al., 2013; Chen et al., 2013), or mixed-effect multilevel analysis (Worsley et al., 2002; Woolrich et al., 2004; Chen et al., 2012), with the assumption that each effect estimate is equally reliable across all subjects. However, deviations of the HDR from the presumed shape would result in biased estimates of the

¹The efficiency in the statistics context measures the optimality of a testing method. A more efficient test requires a smaller sample size to attain a fixed power level.

amplitude, in addition to failing to capture differences in shape such as during the undershoot or recovery phase. ESM is the most flexible among the three methods in terms of providing a more accurate characterization of the BOLD response and can achieve higher activation detection power in individuals. In addition, the estimated HDR curve with a unique signature shape offers much stronger support for the existence of activation than a single scaling factor or β value with FSM or ASM. Compared with FSM, ASM also results in a less biased response amplitude for the principal kernel, and can account for more variance compared to FSM; however, the common practice of using only the principal kernel's coefficient at the group level will not allow the detection of shape changes between conditions and/or groups when those exist.

Difficulties with using ESM (and to a lesser degree ASM) include the need for a larger number of kernel coefficients that need to be estimated. They require m times more regressors than FSM in the individual subject analysis, which translates to more data points and scanning time to reach similar statistical power in individuals. Secondly, the risk of over-fitting exists when some confounding effects such as head motion and physiological noise are stimulus-locked and not fully accounted for. Lastly, the most challenging step lies at the group level: How to simultaneously handle those m effect estimates? And how to summarize and interpret the results? To avoid the complexity involved in the multiple effect estimates from ESM or ASM, the popular approach at the group level is dimensional reduction, condensing the shape information over the multiple values into one number. For ESM, one method is to sum over all or a subset of effect estimates (e.g., ignoring a few time points at the beginning and the end) to obtain the area under the curve (AUC) (e.g., Beauchamp et al., 2003; Greene et al., 2007; McGregor et al., 2013). As the BOLD response curve can be characterized by parameters such as amplitude (or height), delay (or time-to-peak), duration (or HRF), another dimensional reduction proposal is to perform the group analysis on such a derived parameter from the estimated HDR (Lindquist et al., 2009; Degras and Lindquist, 2014). With two or three effect estimates per condition from ASM at the group level, the popular approach focuses on the β value of the canonical HDR while ignoring the parameters for the shape adjustments (i.e., the function of these other parameters is to absorb minor shape fluctuations that would otherwise be modeled as "noise"). One alternative is to estimate the HDR height using the Euclidean or L^2 -norm distance (L^2D) of the two or three effect estimates (Calhoun et al., 2004; Lindquist et al., 2009; Steffener et al., 2010). Essentially, these dimensional reduction methods transform the effect estimates in an k -dimensional space \mathbb{R}^k to one-dimensional \mathbb{R}^1 . As information loss is unavoidable in the process, statistical power in activation identification would suffer. This raises the question of whether a more preferable approach to significance testing might better exploit the information in the HDR shape at the group level.

A Motivational Example

To demonstrate and compare various modeling approaches at the group level, we adopt the same experimental data used in

our previous paper (Chen et al., 2014), with a typical group design that accounts for a confounding effect: varying age across subjects. Briefly, the experiment involved one between-subjects factor, group (two levels: 21 children and 29 adults) and one within-subject factor (two levels: congruent and incongruent conditions). Stimuli were large letters (either "H" or "S") composed of smaller letters ("H" or "S"). For half of the stimuli, the large letter and the component letters were congruent (e.g., "H" composed of "H"s) and for half they were incongruent (e.g., "H" composed of "S"s). Parameters for the whole brain BOLD data on a 3.0 T scanner were: voxel size of $3.75 \times 3.75 \times 5.0$ mm³, 24 contiguously interleaved axial slices, and TR of 1250 ms (TE = 25 ms, FOV = 240 mm, flip angle = 35°). Six runs of EPI data were acquired from each subject, and each run lasted for 380 s with 304 data points. The task followed an event-related design with 96 trials in each run, with three runs of congruent stimuli interleaved with three runs of incongruent stimuli (order counterbalanced across subjects). Subjects used a two button box to identify the large letter during global runs and the component letter during local runs. Each trial lasted 2500 ms: the stimulus was presented for 200 ms, followed by a fixation point for 2300 ms. Inter-trial intervals were jittered with a varying number of TRs, allowing for a trial-by-trial analysis of how the subject's BOLD response varied with changes in reaction time (RT). The experiment protocol was approved by the Combined Neuroscience Institutional Review Board at the NIMH, and the National Clinical Trials Identifier is NCT00006177.

The EPI time series went through the following preprocessing steps: slice timing and head motion corrections, spatial alignment to a Talairach template (TT_N27) at a voxel size of $3.5 \times 3.5 \times 3.5$ mm³, smoothing with an isotropic FWHM of 6 mm, and scaling each voxel time series by its mean value. The scaling step during preprocessing enables one to interpret each regression coefficient of interest as an approximate estimate of percent signal change relative to the temporal mean. The six runs of data were concatenated for the individual regression analysis with the discontinuities across runs properly handled (Chen et al., 2012). To capture the subtle HDR shape under a condition, two modeling approaches were adopted, ESM and ASM, for model comparison. With ESM, each trial was modeled with 10 tent basis functions, each of which spanned one TR (or 1.25 s). The subject's RT at each trial was incorporated as a per-trial modulation variable. In other words, two effects per condition were estimated in the time series regression at the individual level: one revealed the response curve associated with the average RT while the other showed the marginal effect of RT (response amplitude change when RT increases by 1 s) at each time point subsequent to the stimulus. In addition, the following confounding effects were included in the model for each subject, for each run: third-order Legendre polynomials accounting for slow drifts, incorrect trials (misses), censored time points with extreme head motion, and the six head motion parameters. The modeling strategy remained the same with ASM except that the three SPM basis functions (canonical IRF plus time and dispersion derivatives) were employed to model the BOLD responses instead of the 10 tents.

At the group level, it is the BOLD effects associated with the average RT that are of interest here. In addition

to the estimated HDR profiles, three other explanatory variables considered are: a) between-subjects factor, Group (two levels: children and adults), b) within-subject factors, Condition (two levels: congruent and incongruent), and c) quantitative covariate, age. The focus is on the interaction of HDR between Group and Condition: Do the two groups differ in the HDR profile contrast between the two conditions?

Preview

This paper is a sequel to our previous exploration (Chen et al., 2014) of the multivariate modeling (MVM) approach for fMRI group analysis. The layout is as follows. First, we explore and review various hypothesis testing strategies at the group level when the HDR is estimated through multiple basis functions. Second, simulation data were generated to reveal how each methodology performs in terms of controllability for false positives and false negatives, and the performance of these methods was assessed when they were applied to the experimental dataset at both individual and group levels. Finally, we compare all the modeling methodologies for ASM and ESM as well as with and without dimension reduction. The modeling strategies and testing methods discussed here are all performed at the voxel level. Multiple testing correction can be applied in the conventional fashion by controlling the false positive rate (Benjamini and Hochberg, 1995) or the family-wise error through Monte Carlo simulations (3dClustSim in AFNI, Forman et al., 1995) or random field theory (Worsley et al., 1992).

Our major contribution here is to demonstrate the importance of accounting for shape differences and to offer testing approaches at the group level within an MVM platform with the modeling flexibility that would not be available under the conventional GLM. Through our demonstration we propose that ESM should be adopted whenever appropriate or possible to identify the nuanced differences in HDR shape that would be difficult or unlikely to be revealed through FSM or ASM. Furthermore, we recommend that the investigator report the effect estimates such as the HDR curves to substantiate the results in addition to the statistical significance. The modeling framework and functionality are available in the program 3dMVM for public use in the AFNI suite (Cox, 1996).

Throughout this article, regular italic letters (e.g., α) stand for scalars, boldfaced italic letters in lower (\mathbf{a}) and upper (\mathbf{X}) cases for column vectors and matrices respectively. The word *multivariate* is used here in the sense of treating the effect estimates from the same subject or from the levels of a within-subject factor as the instantiations of simultaneous response (or outcome) variables (e.g., the effect estimates for the HDR). This usage differs from the popular connotation in the fMRI field when the spatial structure (multiple voxels) is modeled as the simultaneous response variables, including such methods as multivariate pattern analysis (Haxby, 2012), independent component analysis, and machine learning methods such as support vector machines. Major acronyms used in the paper are listed in Appendix A.

METHODS

As shown in Chen et al. (2014), we formulate the group analysis under a multivariate GLM or MVM platform that is expressed from a subject-wise perspective, $\beta_i^T = \mathbf{x}_i^T \mathbf{A} + \delta_i^T$, or through the variable-wise pivot, $\mathbf{b}_j = \mathbf{X} \mathbf{a}_j + \mathbf{d}_j$, or in the following concise form,

$$\mathbf{B}_{n \times m} = \mathbf{X}_{n \times q} \mathbf{A}_{q \times m} + \mathbf{D}_{n \times m}. \quad (1)$$

The n rows of the response matrix $\mathbf{B} = (\beta_{ij})_{n \times m} = (\beta_1^T, \beta_2^T, \dots, \beta_n^T)^T = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)$ represent the data from the n subjects while the m columns correspond to the levels of within-subject factor(s). For example, the effect estimates from the multiple basis functions under ESM or ASM can be considered the response values associated with the levels of a within-subject or repeated-measures factor (termed Component hereafter). When multiple within-subject factors occur, all their level combinations for each subject are *flattened* from a multi-dimensional space onto a one-dimensional row of \mathbf{B} . It is noteworthy that the within-subject factors are expressed as columns in \mathbf{B} on the left-hand side of the model (1), and only between-subjects variables such as subjects-grouping factors (e.g., sex, genotypes), subject-specific measures (e.g., age, IQ) and their interactions are treated as q explanatory variables on the right-hand side. The same linear system is assumed for all the m response variables, which share the same design matrix $\mathbf{X} = (\mathbf{x}_{ih}) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$. Without loss of generality, \mathbf{X} is assumed to have full column-rank q . Each column of the regression coefficient matrix $\mathbf{A} = (\alpha_{hj})$ corresponds to a response variable, and each row is associated with an explanatory variable. Lastly, the error matrix $\mathbf{D} = (\delta_{ij})_{n \times m} = (\delta_1, \delta_2, \dots, \delta_n)^T = (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m)$ is assumed nm -dimensional Gaussian: $\text{vec}(\mathbf{D}) \sim N(\mathbf{0}, \mathbf{I}_n \otimes \mathbf{\Sigma})$, where vec and \otimes are column stacking and direct (or Kronecker) product operators respectively. As in univariate modeling (UVM), the assumptions for model (1) are linearity, Gaussianity and homogeneity of variance-covariance structure (same $\mathbf{\Sigma}$ across all the between-subjects effects). When only one group of subjects is involved ($q = 1$), the parameter matrix \mathbf{A} becomes a row vector $(\alpha_1, \alpha_2, \dots, \alpha_m)$ that is associated with the m levels of a within-subject factor.

As demonstrated in Chen et al. (2014), MVM has a few advantages over its univariate counterpart. When the data are essentially multidimensional like the multiple effect estimates from ESM or ASM, MVM has a crucial role in formulating hypothesis testing. In addition, it characterizes and quantifies the intercorrelations among the variables based on the data rather than a presumed variance-covariance structure as in UVM. Furthermore, MVM in general provides a better control for false positives than UVM. Lastly, the conventional univariate testing (UVT) under GLM can be easily performed under the MVM framework with a few extra advantages. Here we discuss one aspect by which the group analysis of neuroimaging data will benefit from the MVM facility when the HDR profile is estimated from multiple basis functions instead of being presumed to have a fixed shape. Then in the section Simulations and Real Experiment Results, we elaborate and compare a few testing alternatives in

terms of power and false positives, using simulations and in terms of performance with real data.

Different Testing Strategies

Here we exemplify two simple and prototypical cases with the HDR profile modeled by m basis functions at the individual subject level: a) one group of subjects with the associated effects at the group level expressed as $\alpha_1, \alpha_2, \dots, \alpha_m$ under (1), and b) either two groups or two conditions and the two sets of effect estimates for HDR are α_{1j} and α_{2j} respectively, $j = 1, 2, \dots, m$. To simplify geometric representations, we assume equal number of subjects across groups in the case of group comparison, but the assumption is not required from the modeling perspective. The various modeling strategies discussed below for these two cases can be easily extended to situations with more explanatory variables, including factors and quantitative covariates.

Multivariate Testing (MVT)

As the analogs of one- and two-sample or paired t -tests under UVT, the two prototypes can be expressed with the following null hypotheses,

$$H_{01}^{MVT} : \alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_m = 0, \quad (2a)$$

$$H_{02}^{MVT} : \alpha_{11} = \alpha_{21}, \alpha_{12} = \alpha_{22}, \dots, \alpha_{1m} = \alpha_{2m}. \quad (2b)$$

In other words, the m regression coefficients associated with the m basis functions from each subject are brought to the group level and treated as the instantiated values of m simultaneous variables. When the effect estimates associated with the basis functions of ESM or ASM are treated as the values of m simultaneous response variables, the hypothesis (2a) or (2b) can be analyzed through MVT under the model (1). Geometrically, the data for H_{01}^{MVT} represent the group centroid $(\alpha_1, \alpha_2, \dots, \alpha_m)$ in the m -dimensional real coordinate space \mathbb{R}^m (Table 1), and the associated one-sample Hotelling T^2 -test is performed to reveal whether the group centroid lies in the rejection region (outside of an m -dimensional ellipse centering around the origin in the case of H_{01}^{MVT}). Similarly, the data for H_{02}^{MVT} are expressed as two group centroids, $(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1m})$ and $(\alpha_{21}, \alpha_{22}, \dots, \alpha_{2m})$, and the corresponding two-sample Hotelling T^2 -test is conducted to see if the hypothesis (2b) about the two centroids can be rejected. The hypothesis (2b) can be easily generalized to the situation with more than two groups of subjects (e.g., three genotypes) as well as more than one subject-grouping variable (e.g., sex, genotypes, and handedness) through the formulation of general linear testing (Chen et al., 2014). One noteworthy feature of MVT is that it allows those simultaneous effects to have different scales or units, unlike the traditional AN(C)OVA or univariate GLM in which all the levels of a factor are usually of the same dimension.

Linear Mixed-effects Modeling (LME)

As demonstrated in Chen et al. (2013), linear mixed-effects modeling (LME) can be adopted for group analysis when the HDR is estimated through multiple basis functions. Specifically, the m regression coefficients from each subject associated with the m basis functions are modeled as values corresponding to m levels of a within-subject factor under the LME framework. When

no other explanatory variables are present in the model, the LME methodology can be formulated by (2a) with an intercept of 0. That is, the m effects are coded by m indicator variables instead of any conventional contrast coding. Suppose that the m effect estimates associated with the m basis functions from the i th subject are $\beta_{i1}, \beta_{i2}, \dots, \beta_{im}$, the LME model can be specified as,

$$\beta_{ij} = \alpha_j x_{ij} + \delta_i + \epsilon_{ij}, i = 1, 2, \dots, n, j = 1, 2, \dots, m.$$

where the random effect δ_i characterizes the deviation or shift of the i th subject's HDR from the overall group HDR, the residual term ϵ_{ij} indicates the deviation of each effect estimate β_{ij} from the i th subject's HDR, and the indicator variables x_{ij} take the cell mean coding,

$$x_{ij} = \begin{cases} 1, & \text{if } i\text{th subject is at } j\text{th level,} \\ 0, & \text{otherwise.} \end{cases}$$

so that the parameters $\alpha_j, j = 1, 2, \dots, m$ capture the overall group HDR. The significance of the overall HDR at the group level can be tested through LME on the same hypothesis as (2a),

$$H_0^{LME} : \alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_m = 0. \quad (3)$$

It is of note that the LME approach does not work when other explanatory variables (multiple groups, conditions, or quantitative covariates) are involved because (2a) or (2b) cannot be formulated due to the parameterization constraint through dummy coding. For instance, when there are two groups involved, the typical contrast coding for the two groups renders one dummy variable (e.g., the contrast of one group vs. the other when effect coding is adopted); however, such a coding strategy relies on the existence of an intercept in the model. If the two groups are coded by two indicator variables, the model matrix would become overparameterized.

Area-under-the-Curve (AUC)

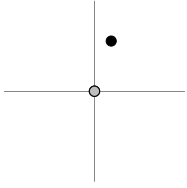
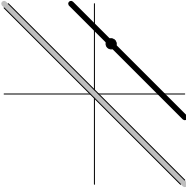
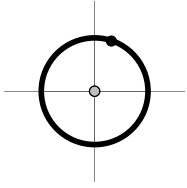
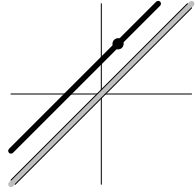

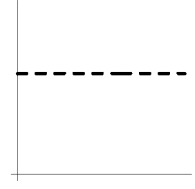
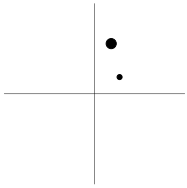
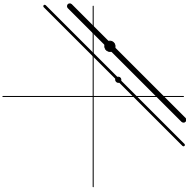
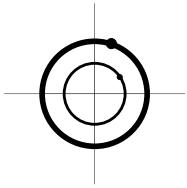
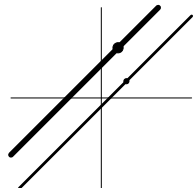
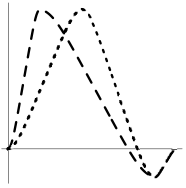

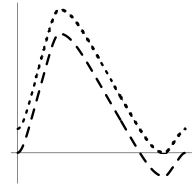
The multiple estimates associated with the multiple basis functions can be reduced to a single value, which is the area under the curve of the estimated response function. The AUC hypotheses for the two prototypes (2a) and (2b) become

$$H_{01}^{AUC} : \sum_{j=1}^m \alpha_j = 0, \quad (4a)$$

$$H_{02}^{AUC} : \sum_{j=1}^m \alpha_{1j} = \sum_{j=1}^m \alpha_{2j}. \quad (4b)$$

That is, the sum of the m coefficients (or area under the HDR curve) is used to summarize the overall response amplitude per subject in one- or two-sample t -test at the group level. The AUC hypotheses (4a) and (4b) are essentially a zero-way interaction (or intercept) and a one-way interaction (or the main effect of Group or Condition) respectively and can be performed under the AN(C)OVA, GLM, or MVM framework. Their geometrical interpretations are as follows (cf. Table 1). The data for H_{01}^{AUC} lie on an \mathbb{R}^{m-1} isosurface (or hyperplane) $\alpha_1 + \dots + \alpha_m = c$, and the

TABLE 1 | Schematic comparisons among various testing methods.

One-sample				
Method ^a	MVT/LME	AUC	L2D	EXC (XUV and XMV)
H_0	$\alpha_1 = \dots = \alpha_m = 0$	$\alpha_1 + \dots + \alpha_m = 0$	$(\alpha_1^2 + \dots + \alpha_m^2)^{1/2} = 0$	$\alpha_1 = \dots = \alpha_m$
Dimensions in \mathbb{R}^m	0	$m - 1$	$m - 1$	1
DFs for F -statistic ^b	$m, n - m - q + 1$	$1, n - q$	$1, n - q$	$m - 1, (m - 1)(n - q)$
Geometric representation ^c of H_0 and H_1 ($m = 2$)				
Geometric representation ^d of HDR when detection failure occurs due to improper H_0 formulation	no		no	
Two-sample or paired				
Method	MVT	AUC	L2D	EXC (XUV and XMV)
H_0	$\alpha_{11} = \alpha_{21}, \dots, \alpha_{1m} = \alpha_{2m}$	$\sum_{j=1}^m \alpha_{1j} = \sum_{j=1}^m \alpha_{2j}$	$(\sum_{j=1}^m \alpha_{1j}^2)^{1/2} = (\sum_{j=1}^m \alpha_{2j}^2)^{1/2}$	$\alpha_{11} - \alpha_{21} = \dots = \alpha_{1m} - \alpha_{2m}$
Dimensions in \mathbb{R}^m	0	$m - 1$	$m - 1$	1
DFs for F -statistic	$m, n - m - q + 1$	$1, n - q$	$1, n - q$	$m - 1, (m - 1)(n - q)$
Geometric representation ^e of H_0 and H_1				
Geometric representation ^f of HDR when detection failure occurs due to improper H_0 formulation	no			

^a The table is meant to show the dimensions of each null hypothesis and an instantiation in the rejection domain while the whole rejection domain is not represented here. For example, the reject region of one-sample Hotelling T^2 -test for MVT (2a) is outside of an m -dimensional ellipse.

^b An interesting fact is that the numerator degrees of freedom for the F -statistic under MVT and UVT are the dimensions of the complementary space to the associated null hypothesis H_0 , or the dimensions of the alternative hypothesis H_1 .

^c The two axes represent the two weights associated with the two basis functions. The whole rejection regions are not shown here, and the shaded (gray) and solid (black) areas correspond respectively to the null hypothesis H_0 space and an instantiation (and its dimension) in the alternative hypothesis H_1 space. Detection failure occurs when the group centroid falls on the diagonal line other than the origin under AUC and EXC.

^d The horizontal and vertical axes represent time and the amplitude of HDR curve (dashed line).

^e The two axes represent the two weights associated with the two basis functions. The whole rejection regions are not shown here, and the shaded and solid areas correspond respectively to the null hypothesis H_0 space and an instantiation (and its dimension) in the alternative hypothesis H_1 space. The two types of line thickness (or dot size) differentiate the two groups (or conditions).

^f The horizontal and vertical axes represent time and the amplitude of HDR curves. The two line types, dashed and dotted, differentiate the two groups or conditions.

associated test for AUC (4a) is executed on the distance between the data isosurface and the null isosurface $\alpha_1 + \dots + \alpha_m = 0$. As the correct null hypothesis for MVT (2a) is only a subset of AUC (4a), the rejection domain of AUC (4a) is only a subset of the rejection domain for MVT (2a), leading to a misrepresentation in (4a) and a detection failure when a data point lies on $\alpha_1 + \dots + \alpha_m = 0$ but not at the origin (i.e., the HDR curve has roughly equal area below and above the x -axis, e.g., a large undershoot). Similarly for H_{02}^{AUC} .

Euclidean Distance (L2D)

As an alternate dimension reduction approach, the null hypotheses associated with the Euclidean or L^2 distance (L2D) for ESM can be formulated respectively as

$$H_{01}^{L2D} : \left(\sum_{j=1}^m \alpha_j^2 \right)^{1/2} = 0, \quad (5a)$$

$$H_{02}^{L2D} : \left(\sum_{j=1}^m \alpha_{1j}^2 \right)^{1/2} = \left(\sum_{j=1}^m \alpha_{2j}^2 \right)^{1/2}. \quad (5b)$$

In other words, one captures the overall magnitude for each subject using the L^2 -distance of the m regression coefficients from no response, and then performs one- or two-sample t -test on the distances.

For ASM, the null hypotheses with the focus on the canonical basis are

$$H_0^{CAN} : \alpha_1 = 0, \quad (6a)$$

$$H_0^{CAN} : \alpha_{11} = \alpha_{21}. \quad (6b)$$

And the null hypotheses for L2D (Calhoun et al., 2004; Steffener et al., 2010) are tested with the first two bases,

$$H_0^{L2D} : \text{sgn}(\alpha_1)(\alpha_1^2 + \alpha_2^2)^{1/2} = 0, \quad (7a)$$

$$H_0^{L2D} : \text{sgn}(\alpha_{11})(\alpha_{11}^2 + \alpha_{12}^2)^{1/2} = \text{sgn}(\alpha_{21})(\alpha_{21}^2 + \alpha_{22}^2)^{1/2} \quad (7b)$$

or with all the three bases,

$$H_0^{L2D} : \text{sgn}(\alpha_1)(\alpha_1^2 + \alpha_2^2 + \alpha_3^2)^{1/2} = 0, \quad (8a)$$

$$H_0^{L2D} : \text{sgn}(\alpha_{11})(\alpha_{11}^2 + \alpha_{12}^2 + \alpha_{13}^2)^{1/2} = \text{sgn}(\alpha_{21})(\alpha_{21}^2 + \alpha_{22}^2 + \alpha_{23}^2)^{1/2}, \quad (8b)$$

where sgn is the sign function. That is, the L2D for ASM is similar to the L2D for ESM, but using the two or three weights associated with the two or three basis functions in ASM and assigning the sign of the canonical response to the resultant L^2 -distance.

Their geometrical interpretations are as follows (Table 1). The data for H_{01}^{L2D} lie on an \mathbb{R}^{m-1} iso-sphere, and the associated test for (5a) is executed on the radius of the \mathbb{R}^{m-1} iso-sphere, leading to no geometrical distortion (but not necessarily true statistically). On the other hand, the data for H_{02}^{L2D} are on two \mathbb{R}^{m-1} iso-sphere surfaces, and the associated test for (5b) acts on the radius difference between the two \mathbb{R}^{m-1} iso-spheres, resulting a detection failure when the two HDR curves have roughly the same radius.

Effect-by-Component Interaction (EXC: XUV and XMV)

By treating the m effect estimates from ESM as m levels of a within-subject factor Component, one can test the hypothesis for the effect-by-component interaction (EXC); that is, the m regression coefficients associated the m basis functions are taken to the group level without any condensation:

$$H_{01}^{EXC} : \alpha_1 = \alpha_2 = \dots = \alpha_m, \quad (9a)$$

$$H_{02}^{EXC} : \alpha_{11} - \alpha_{21} = \alpha_{12} - \alpha_{22} = \dots = \alpha_{1m} - \alpha_{2m}. \quad (9b)$$

As discussed in Chen et al. (2014), EXC (9) can be tested through two methods, one univariate testing for the interaction (XUV), and one multivariate testing for the interaction (XMV). More specifically, with XUV one tests the equality among the m components in (9) by treating them as the m levels of a within-subject factor in an AN(C)OVA or univariate GLM platform. In contrast, the equality among the m components in (9) is tested in XMV as m simultaneous variables in a MAN(C)OVA or multivariate GLM (Appendix B).

The geometrical interpretations of the hypotheses are the following (Table 1). EXC (9a) tests the main effect (or first-way interaction) of Component, representing a straight line in \mathbb{R}^m . The associated test for (9a) is executed on the distance between the data line and the null line (a diagonal line through the origin). As the correct null hypothesis (2a) is only a subset of H_{01}^{EXC} , its rejection domain is only a subset of the rejection domain for MVT (2a), leading to a misrepresentation in (9a) and a detection failure when the group centroid lies on the null line but not at the origin (i.e., the HDR curve is roughly a flat line). Similarly, EXC (9b) as a two-way interaction between Group/Condition and Component is represented by two lines, and the corresponding test acts on the distance between the two lines: are the HDR profiles parallel with each other between the two groups or conditions? As the correct null hypothesis (2b) is only a subset of EXC (9b), the rejection domain of EXC (9b) is only a subset of MVT (2b), resulting in a misrepresentation in (9b) and a detection failure when the two HDR curves are roughly parallel with each other (Table 1).

SIMULATIONS AND REAL EXPERIMENT RESULTS

Among all the testing strategies, LME and MVT are the most precise (points in Table 1). Among all the dimensional reduction methods, the two EXC methods, XUV and XMV, are of the closest approximation to the null hypothesis (lines), while AUC and L2D are the least accurate (\mathbb{R}^{m-1} planes and sphere surfaces respectively). We need to address the question of whether the geometric accuracy order translates to statistical power through simulations and to performance when the methods are applied to real data.

Simulations of Group Analysis with Different Testing Methods

As the spatial extent of fMRI data analysis is independently controlled through false positive rate or family-wise error, the

simulations here were performed at a voxel to examine and compare the false positives and power performance among the testing methods. Simulated data were generated with the following parameters, imitating a typical fMRI group analysis with six scenarios (top row in **Figure 1**): a) one group of subjects with a small undershoot at the end of HDR curve; b) one group of subjects with a moderate undershoot at the end; c) two homoscedastic groups (same variance between groups) with equal number of subjects in each with a similar HDR profile but a factor of 2 difference in amplitude; d) two homoscedastic groups with equal number of subjects in each with HDR having the same amplitude but with a 2 s difference in peak location; e) two heteroscedastic groups (different variance between groups) with equal number of subjects in each with a similar HDR profile but a factor of 2 difference in amplitude; and f) two heteroscedastic groups with equal number of subjects in each with HDR having the same amplitude but with a 2 s difference in peak location. The HDRs are presumably estimated through 7 basis functions (e.g., TENT in AFNI) at the individual level, and the associated 7 effect components $\{\beta_i, i = 1, 2, \dots, 7\}$ at the TR grids are assumed to follow a multivariate Gaussian distribution with a first order autoregressive AR(1) structure for their variance-covariance matrix

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^6 \\ \rho & 1 & \rho & \dots & \rho^5 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^6 & \rho^5 & \rho^4 & \dots & 1 \end{bmatrix}.$$

The choice of a simple Σ structure here is to allow manageable number of simulations while in the same time providing a reasonable structure similar to the one adopted for the Gaussian prior in Marrelec et al. (2003) that guarantees the HDR smoothness. To explore the impact of sample size, the number of subjects in each group was simulated at $n = 9, 12, 15, 18, 21, 24, 27, 30$ with $\rho = 0.3$ for each of the six scenarios. The standard error σ varied (shown in **Figure 1**) across the scenarios to obtain comparable power for each n . 5000 datasets were simulated, each of which was analyzed through 3dMVM with two explanatory variables, Group (between-subjects factor with 2 levels) and Component (within-subject factor with 7 levels that are associated with the 7 basis functions). False positive rate (FPR) and power were assessed by counting the datasets with their respective F - or t -statistic surpassing the threshold corresponding to the nominal significance level of 0.05. Similarly, one- or two-sample t -test was performed on the AUC and L2D values respectively.

Among the six scenarios, all the testing methods showed proper control of FPR except for L2D with one group of subjects. L2D exhibits high power but at the cost of poor FPR control. This is in part due to the reduction of effect estimates to a positive value regardless the signs of the individual components in ESM. It is possible to reduce this problem in ASM when the sign of the principal kernel is assigned to the resulting L2D measure as shown in (7) and (8). Also, L2D achieved the lowest power with two groups of subjects. AUC simply sums over all the components, significantly misrepresenting the effects when

the undershoot becomes moderate. This is reflected in the results where reasonable power is achieved when the undershoot is small and lower power is obtained when the undershoot is moderate. With two groups, AUC performed well in power when the two groups had the same HDR shape, but behaved as poorly as L2D when the two groups had different HDR shapes. As expected, AUC is only sensitive to peak amplitude differences, but is insensitive to shape subtleties. Except for L2D and AUC, the other methods tend to converge in power when the sample size is large enough (e.g., 30 or more). With one group, LME outperformed all other candidates. XUV had a balanced performance on power among all the scenarios, constantly surpassing XMV. Lastly, MVT was slightly more powerful than XUV with two groups when their HDRs were of the same shape with a large number of subjects (e.g., 20 or more per group).

In summary, our simulations show that LME is preferred when there is only one group of subjects with no other explanatory variables present. Under other circumstances, XUV is the preferred choice, especially with the typical sample size of most studies, while MVT, AUC, and XMV may provide some auxiliary detection power.

Results with Experimental Data

How do the testing approaches perform when applied to real data? Would their performances be consistent with the simulations? To address these questions, we ran 3dMVM on the ESM data presented in the Introduction section with $n = 50$ (2 groups: 21 children and 29 adults), $m = 20$ (2 conditions with each having 10 component estimates at 10 TR grids) and design matrix X of $q = 4$ columns in the MVM (1): all ones (intercept associated with the average effect across groups), effect coding for the two groups, the average age effect between the two groups, and the interaction group:age (or group difference in age effect). The age values were centered within each group so that the group effect can be interpreted as the difference between the two groups at their respective average age. The effect of interest was on the interaction of group and condition: Did the two groups have the same HDR profile difference between the two conditions? Five F -statistics from MVT, XUV (with sphericity correction), AUC, L2D, and XMV, were obtained and then, due to different degrees of freedom, converted to Z -values for direct comparisons (**Figure 2A**). To take advantage of the geometrical representation in **Table 1** when interpreting the effect of interest, we reduce the within-subject factor Condition to the contrast between the two conditions, so that the interaction effect essentially becomes the group contrast in terms of the HDR profile difference between the two conditions (**Figure 2C**).

Consistent with the simulation results, XUV achieved the highest detection power in most regions (**Figure 2A** top) while L2D showed low power (and likely high FPR) due to no differentiation between the positive and negative effect estimates for ESM. All the other three methods, MVT, AUC, and XMV, were generally less powerful than XUV. The strong performance of XUV can be seen in the estimated HDR curves at Voxel 1 (**Figures 2B** left,C) extracted from a cluster (left postcentral gyrus). More specifically, the adults had roughly the same HDR profile between the two conditions except for a faster recovery

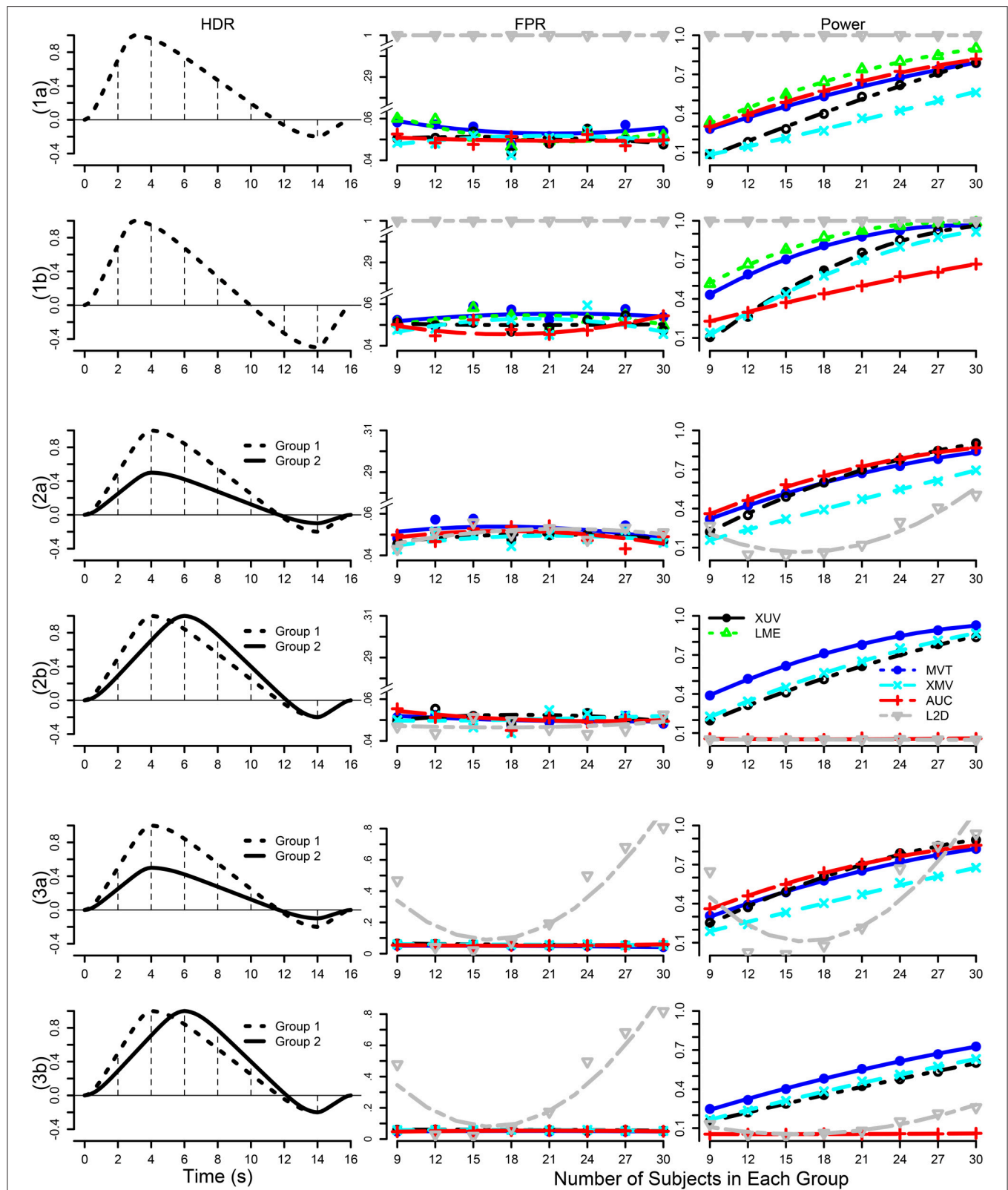


FIGURE 1 | Simulation parameters and results. The six rows correspond to the scenarios in which the presumed HDRs (first column) with a poststimulus undershoot were generated by the convolution program *waver* in AFNI, and sampled at TR = 2 s (shown with vertical dotted lines): (1) one group with a small undershoot (Continued)

FIGURE 1 | Continued

(1a, $\sigma = 1.8$) and a moderate (1b, $\sigma = 1.8$) undershoot, (2) two homoscedastic groups with the same HDR shape but different amplitudes (2a, $\sigma = 0.5$) and with same peak amplitude but a difference of two seconds in peak location (2b, $\sigma = 0.3$), (3) two heteroscedastic groups with the same HDR shape but different amplitudes (3a, $\sigma = 0.3$) and with same peak amplitude but a difference of two seconds in peak location (3b, $\sigma = 0.3$). FPR and power are shown in the second and third columns with a varying number of subjects in each group at a temporal correlation coefficient ρ of 0.3 under six testing approaches: XUV, LME, MVT, XMV, AUC, and L2D. The curves for FPR and power were fitted to the simulation results (plotting symbols) through LOESS smoothing with second order local polynomials.

phase under the Congruent condition than the Incongruent condition; in contrast, the upstroke and peak were more elevated under the Congruent condition in the children than the Incongruent condition except for the recovery phase during the last 3 TRs. Geometrically, the interaction effect between Group and Condition at Voxel 1 is represented by the fact that the HDR profiles of condition difference were intersecting between the two groups (**Figure 2C**). MVT and XMV achieved a moderate power while AUC and L2D failed to reach the significance level of 0.05 at Voxel 1 (**Figure 2B** left). On the other hand, the detection failure of XUV at Voxel 2 (left precuneus) was caused by the fact that the condition contrast was roughly parallel between the two groups (**Figure 2C**), as geometrically demonstrated in **Table 1**. MVT, AUC, and XMV showed their auxiliary role when XUV failed (**Figure 2B** left).

With the ASM analysis results, five tests were performed using 3dMVM. First, the popular approach of focusing on the effect estimate β_0 associated with the first basis (canonical) function through the hypothesis (6b) was adopted (**Figure 2A** bottom). Secondly, the L2D approach (7) was used on the first two basis functions (not shown here) as well as all three. Thirdly, MVT was performed using (2b) with the three coefficients. Lastly, the HDR curve at each condition was reassembled for each subject using the three coefficients, and the reconstructed effect estimates only at the first 10 TRs were analyzed with 3dMVM for two reasons: a) with the three SPM curves covering 32 s or 25 TRs, the model would contain too many parameters relative to the data size; b) the effect estimates after the first 10 TRs were mostly negligible. Two tests, XUV and AUC, were performed while MVT and XMV were impossible because the rank was 3 among the 10 effect estimates from the linearly reconstructed HDR per condition.

The detection power for both β_0 and L2D with ASM was very low (**Figure 2A** bottom), illustrating the fact that focusing on the peak or the combined effects associated the two or three basis functions would largely fail to detect subtle differences during the BOLD uprising and recovery phases. In contrast, MVT (with the coefficients from three basis functions of ASM), XUV and AUC (with the reconstructed HDRs from ASM) outperformed the conventional approaches of β_0 and L2D in SPM. Such failure of ASM is specifically demonstrated at Voxel 1 where the peak alone or the summarized values from the three coefficients were not as powerful as the reassembled HDR profiles (**Figure 2B** right). It is noteworthy that XUV with ASM was less powerful than its ESM counterpart, showcasing the coarser characterization with three parameters in ASM than the estimation at every time point in ESM. Furthermore, for both ESM and ASM, even though XUV was mostly more powerful than the alternatives, MVT and AUC (as well as XMV for ESM and β_0 for ASM) played

a supplementary role when XUV failed (Voxel 2 in **Figure 2B** right).

To recapitulate the performance of the five testing methods in situations when LME cannot be applied, ESM provided a more accurate estimation for the HDR curves than ASM, leading to a higher success in detection power. In addition, with the typical sample size in most studies, XUV as an approximate approach had the lowest power loss at the group level compared to other dimensional alternatives as well as the test with the most accurate hypothesis formulation, MVT. However, MVT plus the lesser accurate approximations such as AUC and XMV may play an auxiliary or even irreplaceable role in situations when XUV suffers from power loss (e.g., **Table 1** or Voxel 2 in **Figure 2**).

DISCUSSION

There are many characteristics that could describe the HDR shape: onset latency, onset-to-peak, peak location, peak duration, magnitude or shape of the undershoot after the onset or during the recovery phase, and habituation or saturation effect. Because of the multiple facets of HDR shape, a lot of effects may well have gone undetected at both individual and group levels in most neuroimaging data analyses, and the failures to capture the shape nuances might have partially contributed to the poor reliability and reproducibility in the field. With a few exceptions, most analyses adopt FSM or ASM mainly for the simplicity of group analysis, as each condition or task is associated with one effect estimate, while other coefficients (e.g., time and dispersion derivatives in ASM) are a priori ignored. That is, activation detection intuitively focuses on the estimated magnitude around the activation peak while statistical inference on the whole HDR shape is generally considered a daunting hurdle. FSM may work well for situations such as a contrast between a condition and fixation. However, it would fail to detect shape subtleties such as prolonged plateau at the peak, slower or faster rise or fall, bigger or longer undershoot, or overall duration. Therefore, FSM through a presumed HDR (gamma variate in AFNI, canonical function in FSL and SPM) is very crude even in an experiment with a block design (Saad et al., 2006; Shan et al., 2013). ASM is an improvement over FSM; however, its flexibility is still limited. For instance, when one is interested in contrasting two conditions (or groups) or in investigating higher-order interactions, the three ASM basis functions may still not be enough in capturing the undershoot subtleties. In addition, characterizing the whole HDR curve with its peak value from ASM for group analysis may suffer from significant power loss, as demonstrated in our real experimental data. Response shapes can vary considerably over space (e.g., Handwerker et al., 2004; Gonzalez-Castillo et al., 2012; Badiillo et al., 2013), and we believe it is important to model

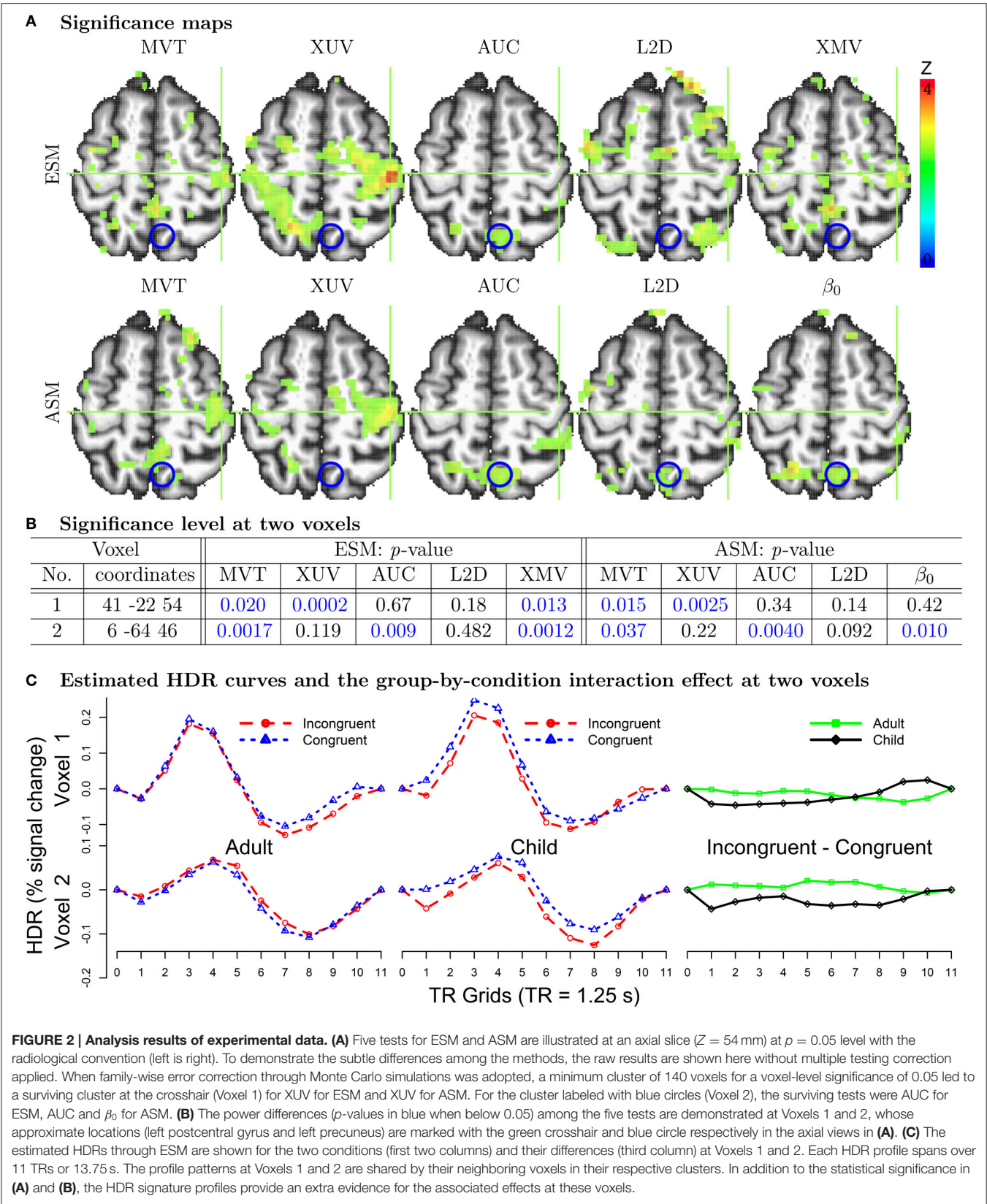


FIGURE 2 | Analysis results of experimental data. (A) Five tests for ESM and ASM are illustrated at an axial slice ($Z = 54$ mm) at $p = 0.05$ level with the radiological convention (left is right). To demonstrate the subtle differences among the methods, the raw results are shown here without multiple testing correction applied. When family-wise error correction through Monte Carlo simulations was adopted, a minimum cluster of 140 voxels for a voxel-level significance of 0.05 led to a surviving cluster at the crosshair (Voxel 1) for XUV for ESM and XUV for ASM. For the cluster labeled with blue circles (Voxel 2), the surviving tests were AUC for ESM, AUC and β_0 for ASM. **(B)** The power differences (p -values in blue when below 0.05) among the five tests are demonstrated at Voxels 1 and 2, whose approximate locations (left postcentral gyrus and left precuneus) are marked with the green crosshair and blue circle respectively in the axial views in **(A)**. **(C)** The estimated HDRs through ESM are shown for the two conditions (first two columns) and their differences (third column) at Voxels 1 and 2. Each HDR profile spans over 11 TRs or 13.75 s. The profile patterns at Voxels 1 and 2 are shared by their neighboring voxels in their respective clusters. In addition to the statistical significance in **(A)** and **(B)**, the HDR signature profiles provide an extra evidence for the associated effects at these voxels.

more accurately the HDRs at the individual level and test for shape rather just amplitude at the group level, particularly when detecting subtle differences between conditions or groups. The dominant adoption of FSM or ASM with a relatively rigid HDR shape reflects the daunting challenge in adopting ESM at the group level, and it is this challenge that motivated our exploration of various group analysis strategies with ESM.

Overview of the Testing Methodologies

Among all the testing strategies for ESM (Table 1), MVT and LME maintain an accurate characterization for the hypothesis. In contrast, the dimensional reduction methods AUC, L2D, and EXC (XUV and XMV) project the original space of the alternative hypothesis from \mathbb{R}^m to \mathbb{R}^1 , \mathbb{R}^1 , and \mathbb{R}^{m-1} , respectively. Any dimensional reduction usually translates to information loss or geometrical distortion. Based on the results from our simulations and real data applications, we believe that the major testing methods for ESM are LME, XUV, MVT, XMV, and AUC, which all have the proper controllability for FPR. If sample size is not an issue in fMRI studies, MVT (e.g., hypothesis 2a or 2b) would be the most accurate approach in terms of hypothesis characterization. However, in practice the number of subjects is usually not large enough for MVT due to resource limitations (e.g., financial cost, time, and manpower), leading to an underpowered performance of MVT as shown in our simulations and real data. Among all the workaround methods through dimensional reduction, XUV has the least hypothesis distortion and the lowest power loss. With one group of subjects and no other explanatory variables present, XUV surpasses MVT, XMV, and AUC in power. However, with an accurate representation of the hypothesis, LME is slightly more efficient than XUV, and should be considered as the first choice (e.g., Alvarez et al., 2008). For all other situations, LME modeling is not feasible due to the constraint of variable parameterization, and we opt for the workaround methods through dimensional reduction, among which AUC is insensitive to subtle shape differences while XMV mostly underperforms unless when the temporal correlation is relatively high (e.g., 0.65 or higher; Chen et al., 2014). XUV achieves the best balance between dimensional reduction and statistical power. However, as XUV tests for parallelism, not exactly the same as the accurate representation characterized in MVT, it may fail in detecting the situation where the HDR profiles are roughly parallel. To compensate for the occasions when XUV fails, other dimensional reduction methods (MVT, AUC, XMV) may offer some complementary detection power.

In light of the discussion here, we strongly encourage the adoption of the ESM approach to achieving two goals: detecting activations and estimating the hemodynamics by characterizing the HDR shape. In addition to the large power gain at both individual and group levels, ESM provides the estimated HDR shape information at the group level, providing an extra layer of validation about the effect veracity through the graphical display of the familiar HDR shape, and alleviating the misconceptions and malpractices prevalent in statistical analysis (e.g., P-hacking, graphical presentation of statistic values instead of effect estimates, overuse of statistical significance; Motulsky, 2014).

The HDR profile information from ESM offers a precious boost especially when a cluster fails to survive the typical stringent thresholding for multiple testing correction but still reaches the significance level of 0.05 at the voxel level. Such a reassuring support of ESM is unavailable from the alternatives of FSM and ASM, with which typically the investigator would be only able to report the peak HDR magnitude or statistic values at a region.

Our recommendation of adopting ESM not only applies to event-related experiments, but also are adaptable to modeling the attenuation or habituation effect in block designs (Saad et al., 2006). In addition, this approximation modeling methodology of XUV assisted with MVT, AUC, and XMV has been applied to DTI data in which the simultaneous variables (white matter network groups such as corpus callosum, corona radiata, left and right hemispheric projection fibers, left and right hemispheric association fibers) were modeled by multiple explanatory variables (e.g., sex, age, behavioral measures) for each response variable such as fractional anisotropy, axial diffusivity, mean diffusivity, radial diffusivity, T1 relaxation time, proton density, and volume (Taylor et al., 2015).

The proposed modeling strategies have been implemented into the open-source program 3dMVM in AFNI, which offers the investigator all the testing results in the output including XUV and the auxiliary approaches (MVT, XMV, and AUC). MVT for the components from ESM presents a unique challenge when one or more within-subject factors are included in the model, and we offer a testing strategy that still fits in the MVM framework (Appendix B). As an alternative, these tests could be conducted in the traditional univariate GLM except for the two multivariate methods, MVT and XMV. In other words, some of the testing methods (MVT and XMV) are truly multivariate, while others (XUV, AUC, and L2D) are essentially univariate. However, as we demonstrated in Chen et al. (2014), these univariate tests are sometimes difficult to perform under the univariate framework, as shown by the implementation challenges faced by some of the neuroimaging packages. Instead, these univariate tests can be more conveniently formulated under the MVM platform by treating the levels of each within subject factor as simultaneous variables in (1) and then constructing the univariate testing statistics through a conversion process. For example, those univariate tests presented in Figure 2 cannot be performed under the univariate GLM framework due to the incorporation of a covariate (age) in the presence of two within subject factors (Condition and HDR effects). It is in this sense that we frame our discussion here under the MVM perspective.

Limitations of the ESM Approach

It is noteworthy that the reliability information from the individual subject analysis is not considered at the group level with the modeling methods discussed here, unlike the mixed-effect multilevel analysis (Worsley et al., 2002; Woolrich et al., 2004; Chen et al., 2012). In addition, the number of basis functions monotonically increases among FSM, ASM, and ESM, therefore it is expected that the goodness of fit at the individual subject analysis level improves across the three methods. On the other hand, as each condition is characterized through multiple

(e.g., ≥ 7) basis functions in ESM, a reliable estimation of the HDR curve at the individual level pays a price through the lower degrees of freedom and requires enough (e.g., 20 or more) trials per condition, and may encounter the risk of numerical instability due to high correlations or even multicollinearity among the regressors. These latter issues can be exacerbated by poor stimulus timing designs. In addition, the typical regression analysis at the individual level assumes the linearity of HDR across trials. Although available (e.g., `3dNLfim` in AFNI), a non-linear approach is usually difficult to handle and still requires some extent of prior information about the HDR shape. Furthermore, the ESM approach is generally considered to be susceptible to noise or effects unrelated to the effects of interest (e.g., head motion, physiological confounds). In other words, the confounding effects may leak into the HDR estimation through over-fitting. However, the false positives from the potential over-fitting at the individual level is less a concern at the group level for the following reasons: a) the likelihood is reduced unless most subjects systematically have similar or same confounding effects; b) cluster-based inferences may reduce the risk of false positives; and most importantly c) examination of the estimated HDR profiles offer an extra safeguard to filter out the potential false positives.

Comparisons with Other Modeling Approaches

Some (not all) of the dimensional reduction methods for ESM discussed here have been sporadically and individually applied to real data in the literature. For example, a popular practice with ASM is to solely focus on the coefficient of the principal basis function (e.g., canonical curve in SPM) with other coefficients (e.g., time and dispersion derivatives) being *a priori* abandoned. As our results with real data showed, the investigator may fail to detect most activations when the effect lies in the HDR shape nuances but not the peak. One suggestion for ASM was to extend the definition of amplitude in (6) to the L^2 -distance by including either the effect for the time derivative (7) or the effects for both time and dispersion derivatives (8) (Calhoun et al., 2004; Worsley and Taylor, 2006; Steffener et al., 2010). A similar approach was to express the effect estimates from the first two basis functions of ASM as a complex number (Wang et al., 2012). However, the potential issues with L2D or its analogs (e.g., Worsley and Taylor, 2006) are the following. a) The definition of amplitude extension in (7) and (8) is under the premise that all the three basis functions are orthogonal with each other (Calhoun et al., 2004). However, only the first two basis functions are orthogonal with each other, but not the third one. b) The second and third basis functions are not normalized; that is, they are not scaled to have a maximum value of 1, unlike the first basis function. In addition, the three effect estimates have different dimensions: the first is of percent signal change while the other two of percent signal change by the unit of time. Therefore, it is difficult to render a physically meaning interpretation with the L2D measures. c) All the effect estimates including negative values are folded into a positive L2D measure, which cannot be differentiated among those effect estimates on the same circle or sphere (see **Table 1**). In addition, it may lead to the violation of

the Gaussian distribution assumption, as illustrated in the poor controllability of FPR (**Figure 1**). d) Their power performance is not satisfactory (**Figures 1, 2**). As an alternative, MVT or LME through the hypothesis (2a) or (2b) on the two or three effect estimates from ASM, as shown in **Figure 2A**, provides a more accurate characterization because it allows for different units or dimensions across the effects.

Similarly for ESM, two dimensional reduction methods have separately been adopted in data analyses. For example, AUC was employed in Beauchamp et al. (2003), Greene et al. (2007), and McGregor et al. (2013). Although not explicitly stated, XUV was used in several real applications to identify the HDR effect under a condition through the main effect (or one-way interaction) of the ESM components in a one-way within-subject ANOVA (Weissman et al., 2006; Geier et al., 2007; Church et al., 2008), to detect the group or condition differences in the overall HDR shape through the group-by-component or condition-by-component interaction in a two-way ANOVA (e.g., Schlaggar et al., 2002; Church et al., 2008; Shuster et al., 2014), and to explore the three-way group-by-task-by-component interaction (Church et al., 2008). However, two limitations were not addressed in those analyses: the potential identification failure of XUV (**Table 1** and Voxel 2 in **Figure 2**), and the limited applicability of univariate GLM.

Some comparisons were performed in terms of amplitude, peak latency, and duration in the estimated HDR among various modeling methods (e.g., FSM, L2D, ESM, a nonlinear model, and inverse logit model; Lindquist et al., 2009). The inverse logit model was deemed the best among the candidates in both simulations and real data, and slightly more powerful than ESM. However, the comparisons were not optimal. First, the dimensional reduction from the HDR shape in \mathbb{R}^m to the three quantities (amplitude, delay, and duration) in \mathbb{R}^3 might be compromised in power when detecting the shape subtleties—this point can be highly dependent on the experiment. Secondly, the reliability for the estimation of the three characteristics was suboptimal. For example, the lackluster performance of ESM in Lindquist et al. (2009) might be caused by the inaccurate amplitude based on the first local peak because such an approach could be misleading especially when more than one local peak occurs. Lastly, the final group analyses were still focused on the amplitude with the Student's *t*-test, an effective dimensional reduction from \mathbb{R}^m to \mathbb{R}^1 .

A multivariate approach (Zhang et al., 2012) was previously proposed, analogous to our method except for the following differences. It was demonstrated among the voxels within only five structurally pre-defined regions; smoothing the estimated HDR from each subject by a Gaussian kernel and imposing regularization on the smoothed HDR were performed to improve the temporal continuities of the HDR; and group analysis was run through multivariate testing of one-sample or pair-wise comparisons among conditions, equivalent to MVT (2a or 2b) discussed here. Another approach (Zhang et al., 2013) assumed that the HDR under each condition would only vary in amplitude and latency across subjects; that is, the HDR shape was presumed same across all subjects. Specifically, the HDR curve for each condition was characterized at the group level by two parameters:

one was of interest (amplitude) and the other of no interest (delay). In addition, the HDR shape (fixed across subjects) was modeled by cubic splines plus their time derivatives. Once the amplitude was estimated for each subject in a one-tier model that incorporated both within- and across-subject variances, a second round of group analysis was performed only on the amplitudes (ignoring the delay) through typical one-sample or paired *t*-test to make inference about a condition or contrast. The approach was demonstrated among the voxels within only three structurally predefined regions.

Recently, a hierarchical approach was proposed for ESM through integrating both individual and group levels into one model (Degras and Lindquist, 2014) in which the HDR curves were captured through multiple higher-order B-spline functions. Even though only demonstrated on one slice of data, the approach is appealing because the variability at both levels is accounted for. However, the current implementation in Matlab is hindered by the following constraints or limitations. a) Spatial parcellation based on anatomical structure was required to determine the temporal correlation structure in the noise component. More applicable approaches would be based on a priori regions that are functionally parcellated through, for example, hierarchical clustering (Thirion et al., 2006; Ji, 2010), joint parcellation detection-estimation (Badillo et al., 2014), consensus clustering (Badillo et al., 2013), k-means clustering (Ji, 2010), etc. (b) The HDR shape may vary across different stimulus conditions under some scenarios (e.g., Ciuciu et al., 2003), and a presumption of the same shape HDR as in Degras and Lindquist (2014) may decrease the detection power when the shape subtleties are of interest. The same HDR assumption is reasonable under other circumstances and has proven sufficient for encoding or decoding the brain activity (Pedregosa et al., 2015). c) Final statistical inference in Degras and Lindquist (2014) through an asymptotic *t*-test was still based on the scaling factors of the same HDR curve shared by all conditions, a dimensional reduction approach from \mathbb{R}^m to \mathbb{R}^1 . An alternative approach is the incorporation of both individual and group levels in a mixed-effects model under the Bayesian framework (Chaari et al., 2013; Badillo et al., 2014). Applied at a priori regions that are functionally parcellated, this jointed detection and estimation method may render a robust procedure less sensitive to outliers than the conventional two-tier methods

under the assumption that all the voxels share the same HDR within a region or parcel.

CONCLUSION

Here we demonstrate with simulations and experimental data that the fixed-shape (FSM) or adjusted-shape (ASM) method may fail to detect most of the shape subtleties (e.g., the speed of rise or recovery, undershoot) in hemodynamic response (HDR). In contrast, the estimated-shape method (ESM) through multiple basis functions would more accurately characterize the cerebral blood flow regulation, and significantly improve the detection power at both individual and group levels. In addition, we propose an analysis scheme for ESM that still fits within the conventional two-tier analysis pipeline and achieves higher statistical power than the alternatives: one performs regression time series analysis separately for each individual subject, and then conducts group analysis with the individual effect estimates. For one group of subjects, a linear mixed-effects (LME) model is preferred if no other explanatory variables are present. In all other scenarios, statistical inferences on the HDR shape can be achieved through a hybrid combination of multivariate testing (MVT) and dimensional reduction approaches with a multivariate model (MVM). Simulations are shown in terms of controllability for false positive rate (FPR) and power achievement among various testing methods. The strategy was applied to a dataset from a real experiment to compare among different testing strategies in terms of power assessment. In addition, we showcase that the MVM flexibility allows any number of explanatory variables including between- and within-subject factors as well as between-subjects covariates.

ACKNOWLEDGMENTS

Our work benefited significantly from the statistical computational language and environment *R*, its many packages, and the great support of the *R* community. All the plots were created in *R* with the base graphics library. Special thanks are due to Helios de Rosario for his help in technical details of using the *R* package *phia*. The research and writing of the paper were supported by the NIMH and NINDS Intramural Research Programs of the NIH/HHS, USA.

REFERENCES

- Alvarez, R. P., Biggs, A., Chen, G., Pine, D. S., and Grillon, C. (2008). Contextual fear conditioning in humans: cortical-hippocampal and amygdala contributions. *J. Neurosci.* 28, 6211–6219. doi: 10.1523/JNEUROSCI.1246-08.2008
- Badillo, S., Desmidt, S., Ginisty, C., and Ciuciu, P. (2014). “Multi-subject Bayesian joint detection and estimation in fMRI,” in *PRNI* (Tübingen), 1–4.
- Badillo, S., Varoquaux, G., and Ciuciu, P. (2013). “Hemodynamic estimation based on consensus clustering,” in *IEEE Pattern Recognition in Neuroimaging (PRNI)*, (Philadelphia, PA), 211–215.
- Barbé, K., Van Moer, W., and Nagels, G. (2012). Fractional-order time series models for extracting the haemodynamic response from functional magnetic resonance imaging data. *IEEE Trans. Biomed. Eng.* 59, 2264–2272. doi: 10.1109/TBME.2012.2202117
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., and Martin, A. (2003). FMRI responses to video and point-light displays of moving humans and manipulable objects. *J. Cogn. Neurosci.* 15, 991–1001. doi: 10.1162/089892903770007380
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300.
- Bernal-Rusiel, J. L., Greve, D. N., Reuter, M., Fischl, B., Sabuncu, M. R., and Alzheimer's Disease Neuroimaging Initiative, (2013). Statistical analysis of longitudinal neuroimage data with Linear Mixed Effects models. *Neuroimage* 66, 249–260. doi: 10.1016/j.neuroimage.2012.10.065

- Birn, R. M., Saad, Z. S., and Bandettini, P. A. (2001). Spatial heterogeneity of the nonlinear dynamics in the fMRI BOLD response. *Neuroimage* 14, 817–826. doi: 10.1006/nimg.2001.0873
- Buxton, R. B., Uludağ, K., Dubowitz, D. J., and Liu, T. T. (2004). Modeling the hemodynamic response to brain activation. *Neuroimage* 23(Suppl. 1), S220–S233. doi: 10.1016/j.neuroimage.2004.07.013
- Calhoun, V. D., Stevens, M. C., Pearson, G. D., and Kiehl, K. A. (2004). fMRI analysis with the general linear model: removal of latency-bias by incorporation of hemodynamic derivative terms. *Neuroimage* 22, 252–257. doi: 10.1016/j.neuroimage.2003.12.029
- Chaari, L., Vincent, T., Forbes, F., Dojat, M., and Ciuciu, P. (2013). Fast joint detection-estimation of evoked brain activity in event-related fMRI using a variational approach. *IEEE Tran. Med. Imaging* 32, 821–837. doi: 10.1109/TMI.2012.2225636
- Chen, G., Adelman, N. E., Saad, Z. S., Leibenluft, E., and Cox, R. W. (2014). Applications of multivariate modeling to neuroimaging group analysis: a comprehensive alternative to univariate general linear model. *Neuroimage* 99, 571–588. doi: 10.1016/j.neuroimage.2014.06.027
- Chen, G., Saad, Z. S., Britton, J. C., Pine, D. S., and Cox, R. W. (2013). Linear mixed-effects modeling approach to fMRI group analysis. *Neuroimage* 73, 176–190. doi: 10.1016/j.neuroimage.2013.01.047
- Chen, G., Saad, Z. S., Nath, A. R., Beauchamp, M. S., and Cox, R. W. (2012). fMRI group analysis combining effect estimates and their variances. *Neuroimage* 60, 747–765. doi: 10.1016/j.neuroimage.2011.12.060
- Church, J. A., Coalson, R. S., Lugar, H. M., Petersen, S. E., and Schlaggar, B. L. (2008). A developmental fMRI study of reading and repetition reveals changes in phonological and visual mechanisms over age. *Cereb. Cortex* 18, 2054–2065. doi: 10.1093/cercor/bhm228
- Ciuciu, P., Poline, J. B., Marrelec, G., Idier, J., Pallier, C., and Benali, H. (2003). Unsupervised robust nonparametric estimation of the hemodynamic response function for any fMRI experiment. *IEEE Trans. Med. Imaging* 22, 1235–1251. doi: 10.1109/TMI.2003.817759
- Cohen, M. S. (1997). Parametric analysis of fMRI data using linear systems methods. *Neuroimage* 6, 93–103. doi: 10.1006/nimg.1997.0278
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014
- Degras, D., and Lindquist, M. A. (2014). A hierarchical model for simultaneous detection and estimation in multi-subject fMRI studies. *Neuroimage* 98, 61–72. doi: 10.1016/j.neuroimage.2014.04.052
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., et al. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14240–14245. doi: 10.1073/pnas.0606302103
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647. doi: 10.1002/mrm.1910330508
- Fox, J., Friendly, M., and Weisberg, S. (2013). Hypothesis tests for multivariate linear models using the *car* package. *The R. Journal* 5, 2013, 39–52.
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., and Turner, R. (1998a). Event-related fMRI: characterizing differential responses. *Neuroimage* 7, 30–40. doi: 10.1006/nimg.1997.0306
- Friston, K. J., Josephs, O., Rees, G., and Turner, R. (1998b). Nonlinear event-related responses in fMRI. *Magn. Reson. Med.* 39, 41–52. doi: 10.1002/mrm.1910390109
- Geier, C. F., Garver, K. E., and Luna, B. (2007). Circuitry underlying temporally extended spatial working memory. *Neuroimage* 35, 904–915. doi: 10.1016/j.neuroimage.2006.12.022
- Gonzalez-Castillo, J., Saad, Z. S., Handwerker, D. A., Inati, S. J., Brenowitz, N., and Bandettini, P. A. (2012). Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5487–5492. doi: 10.1073/pnas.1121049109
- Greene, A. J., Gross, W. L., Elsing, C. L., and Rao, S. M. (2007). Hippocampal differentiation without recognition: an fMRI analysis of the contextual cueing task. *Learn. Mem.* 14, 548–553. doi: 10.1101/lm.609807
- Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823
- Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21, 1639–1651. doi: 10.1016/j.neuroimage.2003.11.029
- Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: the early beginnings. *Neuroimage* 62, 852–855. doi: 10.1016/j.neuroimage.2012.03.016
- Huynh, H., and Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomised block and split-plot designs. *J. Educ. Stat.* 1, 69–82. doi: 10.2307/1164736
- Ji, Y. (2010). *Data-driven fMRI data analysis based on parcellation*. Ph.D. thesis, The University of Nottingham, England.
- Lindquist, M. A., Loh, J. M., Atlas, L. Y., and Wager, T. D. (2009). Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. *Neuroimage* 45, S187–S198. doi: 10.1016/j.neuroimage.2008.10.065
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869–878. doi: 10.1038/nature06976
- Logothetis, N. K., and Wandell, B. A. (2004). Interpreting the BOLD signal. *Annu. Rev. Physiol.* 66, 735–769. doi: 10.1146/annurev.physiol.66.082602.092845
- Mériaux, S., Roche, A., Dehaene-Lambertz, G., Thirion, B., and Poline, J. B. (2006). Combined permutation test and mixed-effect model for group average analysis in fMRI. *Hum. Brain Mapp.* 27, 402–410. doi: 10.1002/hbm.20251
- Magri, C., Schridde, U., Murayama, Y., Panzeri, S., and Logothetis, N. K. (2012). The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. *J. Neurosci.* 32, 1395–1407. doi: 10.1523/JNEUROSCI.3985-11.2012
- Marrelec, G., Benali, H., Ciuciu, P., Pélégriani-Issac, M., and Poline, J. B. (2003). Robust Bayesian estimation of the hemodynamic response function in event-related BOLD fMRI using basic physiological information. *Hum. Brain Mapp.* 19, 1–17. doi: 10.1002/hbm.10100
- McGregor, K. M., Nocera, J. R., Sudhyadhom, A., Patten, C., Manini, T. M., Kleim, J. A., et al. (2013). Effects of aerobic fitness on aging-related changes of interhemispheric inhibition and motor performance. *Front. Aging Neurosci.* 5:66. doi: 10.3389/fnagi.2013.00066
- Motulsky, H. J. (2014). Common misconceptions about data analysis and statistics. *J. Pharmacol. Exp. Ther.* 351, 200–205. doi: 10.1124/jpet.114.219170
- Nichols, T. E., and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25. doi: 10.1002/hbm.1058
- Pedregosa, F., Eickenberg, M., Ciuciu, P., Thirion, B., and Gramfort, A. (2015). Data-driven HRF estimation for encoding and decoding models. *Neuroimage* 104, 209–220. doi: 10.1016/j.neuroimage.2014.09.060
- Poline, J. B., and Brett, M. (2012). The general linear model and fMRI: does love last forever? *Neuroimage* 62, 871–880. doi: 10.1016/j.neuroimage.2012.01.133
- Rencher, A. C., and Christensen, W. F. (2012). *Methods of Multivariate Analysis, 3rd Edn.* Hoboken, NJ: John Wiley & Sons, Inc.
- Reynell, C., and Harris, J. J. (2013). The BOLD signal and neurovascular coupling in autism. *Dev. Cogn. Neurosci.* 6, 72–79. doi: 10.1002/9781118391686
- Saad, Z. S., Chen, G., Reynolds, R. C., Christidis, P. P., Hammett, K. R., Bellgowan, P. S. F., et al. (2006). Functional imaging analysis contest (FIAC) analysis according to AFNI and SUMA. *Hum. Brain Mapp.* 27, 417–424. doi: 10.1002/hbm.20247
- Schlaggar, B. L., Brown, T. T., Lugar, H. M., Visscher, K. M., Miezin, F. M., and Petersen, S. E. (2002). Functional neuroanatomical differences between adults and school-age children in the processing of single words. *Science* 296, 1476–1479. doi: 10.1126/science.1069464
- Shan, Z. Y., Wright, M. J., Thompson, P. M., McMahon, K. L., Blokland, G. G. A. M., de Zubicaray, G. I., et al. (2013). Modeling of the hemodynamic responses in block design fMRI studies. *J. Cereb. Blood Flow Metab.* 34, 316–324. doi: 10.1038/jcbfm.2013.200
- Shuster, L. I., Moore, D. R., Chen, G., Ruschello, D. M., and Wonderlin, W. F. (2014). Does experience in talking facilitate speech repetition? *Neuroimage* 87, 80–88. doi: 10.1016/j.neuroimage.2013.10.064
- Steffener, J., Tabert, M., Reuben, A., and Stern, Y. (2010). Investigating hemodynamic response variability at the group level using basis functions. *Neuroimage* 49, 2113–2122. doi: 10.1016/j.neuroimage.2009.11.014
- Stephen, K. E., Weiskopf, N., Drysdale, P. M., Robinson, P. A., and Friston, K. J. (2007). Comparing hemodynamic models with DCM. *Neuroimage* 38, 387–401. doi: 10.1016/j.neuroimage.2007.07.040

- Taylor, P. A., Jacobson, J. L., van der Kouwe, A., Molteno, C. D., Chen, G., Wintermark, P., et al. (2015). A DTI-based tractography study of structural brain changes associated with prenatal alcohol exposure in newborn. *Hum. Brain Mapp.* 36, 170–186. doi: 10.1002/hbm.22620
- Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., and Poline, J.-B. (2006). Dealing with the shortcomings of spatial normalization: multi-subject parcellation of fMRI datasets. *Hum. Brain Mapp.* 27, 678–693. doi: 10.1002/hbm.20210
- Vincent, T., Badillo, S., Risser, L., Chaari, L., Bakhous, C., Forbes, F., et al. (2014). Flexible multivariate hemodynamics fMRI data analyses and simulations with PyHRF. *Front. Neurosci.* 8:67. doi: 10.3389/fnins.2014.00067
- Wang, P., Wang, Z., and He, L. (2012). Complex-number representation of informed basis functions in general linear modeling of functional magnetic resonance imaging. *J. Neurosci. Methods* 205, 28–35. doi: 10.1016/j.jneumeth.2011.12.016
- Weissman, D. H., Roberts, K. C., Visscher, K. M., and Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nat. Neurosci.* 9, 971–978. doi: 10.1038/nn1727
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., and Nichols, T. E. (2014). Permutation inference for the general linear model. *Neuroimage* 92, 381–397. doi: 10.1016/j.neuroimage.2014.01.060
- Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., and Smith, S. M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *Neuroimage* 21, 1732–1747. doi: 10.1016/j.neuroimage.2003.12.023
- Worsley, K. J., and Taylor, J. E. (2006). Detecting fMRI activation allowing for unknown latency of the hemodynamic response. *Neuroimage* 29, 649–654. doi: 10.1016/j.neuroimage.2005.07.032
- Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., et al. (2002). A general statistical analysis for fMRI data. *Neuroimage* 15, 1–15. doi: 10.1006/nimg.2001.0933
- Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* 12, 900–918. doi: 10.1038/jcbfm.1992.127
- Zhang, T., Li, F., Beckes, L., and Coan, J. A. (2013). A semi-parametric model of the hemodynamic response for multi-subject fMRI data. *Neuroimage* 75, 136–145. doi: 10.1016/j.neuroimage.2013.02.048
- Zhang, T., Li, F., Beckes, L., Brown, C., and Coan, J. A. (2012). Nonparametric inference of the hemodynamic response using multi-subject fMRI data. *Neuroimage* 63, 1754–1765. doi: 10.1016/j.neuroimage.2012.08.014

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Chen, Saad, Adleman, Leibenluft and Cox. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX A

List of Acronyms used in the Paper

AN(C)OVA	Analysis of (co)variance
ASM	Adjusted-shape method
AUC	Area under the curve
ESM	Estimated-shape method
EXC	Effect-by-component interaction
FPR	False positive rate
FSM	Fixed-shape method
GLM	General linear model
HDR	Hemodynamic response
IRF	Impulse response function
L2D	Euclidian (L^2) distance
LME	Linear mixed-effects
MAN(C)OVA	Multivariate analysis of (co)variance
MVM	Multivariate modeling
MVT	Multivariate testing
UVM	Univariate modeling
UVT	Univariate testing
XMV	Multivariate testing for interaction
XUV	Univariate testing for interaction.

APPENDIX B

FORMULATION OF MULTIVARIATE TESTING IN THE PRESENCE OF ONE OR MORE WITHIN-SUBJECT FACTORS

As discussed in Chen et al. (2014), all the within-subject factors are flattened into \mathbb{R}^1 under the multivariate model (MVM) formulation (1). Once the regression coefficient matrix A is estimated through solving the MVM system (1) with the least squares principle, each general linear test (GLT) can be expressed as a function of A ,

$$H_0 : L_{u \times q} A_{q \times m} R_{m \times v} = 0_{u \times v}, \quad (A1)$$

where the *hypothesis matrix* L , through premultiplying, specifies the weights among the rows of A that are associated with the between-subjects variables (groups or subject-specific quantitative covariates), and the *response transformation matrix* R , through postmultiplying, formulates the weighting among the columns of A that correspond to the m response variables. It is assumed that L and R are full of row- and column-rank respectively, and $u \leq q$, $v \leq m$. The matrix L (or R) plays a role of contrasting or weighted averaging among the groups of a between-subjects factor (or the levels of a within-subject factor).

The conventional multivariate test (MVT) can be performed through any of the four multivariate statistics (Wilks' λ , Pillai-Bartlett trace, Lawley-Hotelling trace, and Roy's largest root) with $R = I_m$ once the hypothesis matrix L in (A1) is constructed (Appendix B in Chen et al., 2014). For instance, suppose that we consider an m -variate model with the following explanatory variables: three genotypes of subjects, age and their interactions. Via effect coding with the first genotype as reference, the model

matrix X in (1) is of $q = 6$ columns: one for the intercept, two for the three genotypes, one for age, and two for their interactions. Accordingly, the $q = 6$ rows in A represent the overall mean, the respective effects for the second and third genotypes relative to the overall mean, the age effect associated with the overall mean, and the respective age effects for the second and third genotypes relative the average age effect. MVT for the main effect of genotypes, the genotype-by-age interaction, and the age effect for the first genotype can be obtained under (A1) respectively with

$$L_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, L_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$L_3 = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix}, R_1 = R_2 = R_3 = I_m.$$

Similarly, both univariate and within-subject multivariate tests can be formulated by obtaining both the hypothesis matrix L and the response transformation matrix R in (A1) (Appendix C in Chen et al., 2014). In addition, all the *post-hoc* t - and F -tests (options -gltCode and -glfCode respectively in 3dMVM) are also constructed as MVT under the platform (A1). For instance, the effect under a specific level and the contrast between two levels of a within-subject factor through -gltCode are evaluated essentially by a one-sample and a paired t -test respectively, while the main effect of a within-subject factor through -glfCode is assessed by a within-subject multivariate test.

When $R = \mathbf{1}_{m \times 1}$, the hypothesis (A1) solely focuses on the between-subjects explanatory variables (columns in the model matrix X of MVM; 1) while the effects among the levels of the within-subject factors are averaged (or collapsed). Therefore, the AUC approach (4) can be conceptually tested under the multivariate framework (A1), respectively for one group,

$$L_4 = \mathbf{1}, R_4 = \mathbf{1}_{m \times 1},$$

and two groups,

$$L_5 = (0, 1), R_5 = \mathbf{1}_{m \times 1},$$

even though they would be readily performed through the conventional one- and two-sample t -tests.

When applied to the effect-by-component interaction (9a or 9b) with ESM (EXC in Table 1), the MVM framework offers both univariate (XUV) and multivariate (XMV) approaches, which are tested under the same formulation, respectively for one group (A1),

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m,$$

$$L_6 = \mathbf{1}, R_6 = \begin{bmatrix} I_{m-1} \\ -\mathbf{1}_{1 \times (m-1)} \end{bmatrix},$$

and two groups,

$$H_0 : \alpha_{11} - \alpha_{21} = \alpha_{12} - \alpha_{22} = \dots = \alpha_{1m} - \alpha_{2m},$$

$$L_7 = (0, 1), R_7 = R_6.$$

For XMV, standard multivariate testing statistics (Wilks' λ , Pillai-Bartlett trace, Lawley-Hotelling trace, Roy's largest root) are constructed through the eigenvalues of the "ratio" $\mathbf{H}(\mathbf{H} + \mathbf{E})^{-1}$ between the SSPH matrix \mathbf{H} for the hypothesis (A1) against the SSPE matrix \mathbf{E} for the errors in the full model (Rencher and Christensen, 2012). In contrast, the univariate approach XUV is tested through the formulation of an F -statistic with the numerator and denominator sums of squares being as $\text{tr}(\mathbf{H}(\mathbf{R}^T \mathbf{R})^{-1})$ and $\text{tr}(\mathbf{E}(\mathbf{R}^T \mathbf{R})^{-1})$ under the sphericity assumption (Fox et al., 2013), and the F -value can be adjusted through the Greenhouse and Geisser (1959) or Huynh and Feldt (1976) correction if the sphericity assumption is violated.

All the applications so far in the literature have been focused on either MVT or UVT. In other words, a strict MVT applies to the situations of truly multivariate nature while a purely UVT is adopted to the conventional AN(C)OVA or GLM. However, if we treat the components from ESM as simultaneous response variables, the presence of one or more within-subject factors (e.g., two task conditions in the experimental data of this paper) necessitates a partial MVT. Here we demonstrate a strategy to formulate partial MVT with the construction of \mathbf{L} and \mathbf{R} using

a template of two-way within-subject ANOVA with factors A and B of a and b levels respectively. Suppose that we want to model the levels of factor A as a simultaneous response variables (e.g., components or effect estimates from ESM) while factor B is considered as an explanatory variable (e.g., conditions). MVT for the effect of B can be achieved through the following specifications in (A1),

$$\mathbf{L} = \mathbf{I}_q, \mathbf{R} = \mathbf{I}_a \otimes \mathbf{R}^{(B)}.$$

Similarly, if the levels of factor B are modeled as b simultaneous response variables while factor A is considered as an explanatory variable, we have the following MVT specifications for the effect of A ,

$$\mathbf{L} = \mathbf{I}_q, \mathbf{R} = \mathbf{R}^{(A)} \otimes \mathbf{I}_b.$$

The notations $\mathbf{R}^{(A)} = \begin{bmatrix} \mathbf{I}_{a-1} \\ -\mathbf{1}_{1 \times (a-1)} \end{bmatrix}$ and $\mathbf{R}^{(B)} = \begin{bmatrix} \mathbf{I}_{b-1} \\ -\mathbf{1}_{1 \times (b-1)} \end{bmatrix}$ above are conveniently the effect coding matrices for factors A and B respectively.



An empirical Bayes normalization method for connectivity metrics in resting state fMRI

Shuo Chen^{1*}, Jian Kang² and Guoqing Wang¹

¹ Department of Epidemiology and Biostatistics, University of Maryland, College Park, MD, USA, ² Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA

OPEN ACCESS

Edited by:

Jorge J. Riera,
Florida International University, USA

Reviewed by:

Baxter P. Rogers,
Vanderbilt University, USA
Wensong Wu,
Florida International University, USA

*Correspondence:

Shuo Chen,
Department of Epidemiology and
Biostatistics, University of Maryland,
2234M School of Public Health,
College Park, MD 20742, USA
shuochen@umd.edu

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 27 April 2015

Accepted: 24 August 2015

Published: 16 September 2015

Citation:

Chen S, Kang J and Wang G (2015)
An empirical Bayes normalization
method for connectivity metrics in
resting state fMRI.
Front. Neurosci. 9:316.
doi: 10.3389/fnins.2015.00316

Functional connectivity analysis using resting-state functional magnetic resonance imaging (rs-fMRI) has emerged as a powerful technique for investigating functional brain networks. The functional connectivity is often quantified by statistical metrics (e.g., Pearson correlation coefficient), which may be affected by many image acquisition and preprocessing steps such as the head motion correction and the global signal regression. The appropriate quantification of the connectivity metrics is essential for meaningful and reproducible scientific findings. We propose a novel empirical Bayes method to normalize the functional brain connectivity metrics on a posterior probability scale. Moreover, the normalization function maps the original connectivity metrics to values between zero and one, which is well-suited for the graph theory based network analysis and avoids the information loss due to the (negative value) hard thresholding step. We apply the normalization method to a simulation study and the simulation results show that our normalization method effectively improves the robustness and reliability of the quantification of brain functional connectivity and provides more powerful group difference (biomarkers) detection. We illustrate our method on an analysis of a rs-fMRI dataset from the Autism Brain Imaging Data Exchange (ABIDE) study.

Keywords: anticorrelation, connectivity, fMRI, network, normalization, resting state

1. Introduction

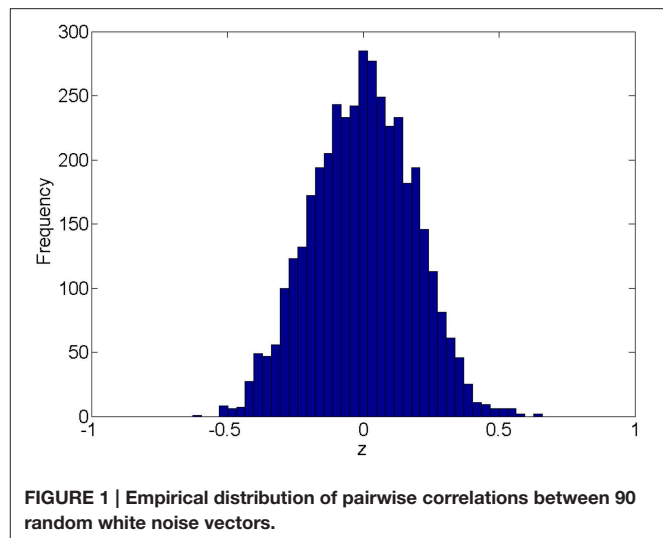
Resting-state fMRI (rs-fMRI) has been applied to study functional brain connectivity patterns and networks in the absence of external stimuli (Biswal et al., 1995; Beckmann et al., 2005; Fransson, 2005; De Luca et al., 2006; Fox et al., 2006). Many previous rs-fMRI studies have identified altered functional connectivity expressions and networks from different clinical populations (Dosenbach et al., 2007; Greicius, 2008; Fornito et al., 2012). To investigate the properties of the complex brain functional connectivity networks, the graph theory models have been developed and yielded many meaningful findings (Braun et al., 2009; Bullmore and Sporns, 2009; Rubinov and Sporns, 2010).

The functional connectivity analyses are often conducted based on connectivity metrics rather than the raw time courses from rs-fMRI data. There have been many functional connectivity metrics employed to measure the functional coherence of temporal profiles between two distinct brain areas, for example, Pearson correlation coefficients, mutual information coefficients, and spectral coherence (Zhou et al., 2009; Smith, 2012). Therefore, the functional connectivity strength is often quantified by a calculated statistic (most times a scalar), and hence the reproducibility and validity of the following group level statistical inferences are heavily impacted by the statistical

quantification method and choice of connectivity metric. However, the connectivity metrics could be sensitive to the changes of image acquisition and preprocessing procedures. For example, in the debate of whether global trend regression should be applied, it has been pointed out that such preprocessing step may shift the whole connectivity distribution (using the Pearson correlation coefficient metric) toward -1 and introduce false anticorrelations (Fox et al., 2009; Murphy et al., 2009; Weissenbacher et al., 2009; Chai et al., 2012). It brings up the practical trade-off between specificity of anticorrelation and the alignment of the scales of correlation value distributions across subjects. Although the agreement (of whether global signal regression should be used) has not been reached, it is clear that the scaling of the connectivity metrics can be influenced by many (preprocessing) factors and substantial noises (Murphy et al., 2013).

The brain functional connectivity often aims to identify the differentially expressed connections between brain areas for different cohorts. To provide valid and reproducible group level functional connectivity inferences for these studies, we are ought to assign proper values to the input connectivity metrics which are proportional to the true connectivity strength and comparable across subjects. Thus, the appropriate scaling and rescaling methods toward the raw connectivity metrics of the high-dimensional connectivity expressions are desired, which is often referred as a “normalization” step. The feature expression normalization has been widely used as a key standard preprocessing step for most of the high-throughput “omics” data, (e.g., the quantile normalization for gene expression microarray data) in order to mitigate the subjectwise systematic shift/noises and to improve the accuracy of differential expression detection by transforming the expression metrics to a comparable scale across subjects (Bolstad et al., 2003; Bullard et al., 2010; Robinson and Oshlack, 2010; Hansen et al., 2012). The normalization plays a crucial role in group level analysis of high-throughput data since the sensitivity, specificity, and reproducibility of differential expression detection rely on the proper quantification of the expression metrics. However, the normalization step has been rarely applied to brain functional connectomics data, though similar subject-wise systematic shift/noises may also exist in functional connectivity data. The appropriate normalization method is expected to be robust to the measurement shifts/noises and to provide a comparable connectivity expression metric across subjects. In addition, when studying the complex functional brain connectivity network, we often employ the graph theoretical models which require the scale of connectivity expression ranging between zero and one (“Binarization”) (Rubinov and Sporns, 2010; Smith, 2012). When the Pearson correlation coefficient is used, the correlation values below zero are often (hard) thresholded (“Thresholding”) (Rubinov and Sporns, 2010; Smith, 2012). However, thresholding or binarization of the continuous connectivity expression values could lead to substantial information loss (Harrell, 2001). Thus, a normalization approach which maps connectivity metrics to the support between zero and one is also desired.

To address the above unmet needs of functional connectivity analysis, we present a new empirical Bayes normalization



method for rs-fMRI connectivity analysis. The method has three main advantages: (1) it mitigates subjectwise systematic shift/noises and provides robust normalized metrics to ensure the connectivity metrics comparable between subjects; (2) the normalized metrics improve differential expression detection for true biomarker detection; (3) it quantifies the connectivity expression value ranging between zero and one which is well-suited for graph theoretical models. In this article, we use Pearson correlation for demonstration because it is most widely used and studied (Zalesky et al., 2012), though the proposed normalization method can be applied to any functional connectivity metrics.

2. Methods

In this section, we illustrate the normalization method based on functional connectivity expressions between 90 nodes, which represents the commonly used first 90 Anatomical Automatic Labeling (AAL) regions in brain connectivity studies (Tzourio-Mazoyer et al., 2002; Zalesky et al., 2010b).

2.1. Distribution of Connectivity

We first introduce the null distribution of 4005 pairwise correlations between time courses from 90 nodes. Each time course is a randomly simulated white noise vector (including 50 data points) with mean 0, and variance 1 and all time courses are generated independently. The resulting connectivity metric distribution is shown in **Figure 1**. The correlations range between (-0.63, 0.66) and are centered around 0. The 4005 sampling correlations (the calculated statistics) are used to quantify the connectivity expressions, and among those correlations there are many values close to 1 or -1 that are often considered as “false positively” correlated or anticorrelated. We denote the distribution in **Figure 1** as the null distribution.

2.2. Normalization Function

In practice, the distribution of connectivity expressions from rs-fMRI data is often mixed by the null distribution as well as the

distributions from the “true positive” correlated or anticorrelated components (a small hump close to 1 or -1). Thus, there are more than one component in the distribution of correlations (i.e., a mixture distribution). Moreover, at the group-level the modes or medians of the correlation distributions from different subjects may shift apart significantly from each other, which may be a result of systematic measurement errors (e.g., in the image acquisition and preprocessing steps). The systematic shifts could cause connectivity expression metrics not comparable across subjects and then lead to a failure of group level inferences such as biomarker detection. To address such concerns, we propose a normalization method to quantify the connectivity expression when adjusting the probability of “false positives” and systematic shifts across subjects.

We denote the connectivity expression value by z , and the probability distribution by $f(z)$. The connectivity expressions are high-throughput, lending itself to recognition of the pattern of “false positives” through a mixture model:

$$f(z) = p_0 f_0(z) + p_1 f_1(z), \quad (2.1)$$

where $p_0 = Pr\{\text{uncorrelated}(\text{null})\}$ and $f_0(z)$ is the probability density distribution (pdf) for the null component; and $p_1 = Pr\{\text{correlated}(\text{non} - \text{null})\}$ and $f_1(z)$ is the pdf for the non-null component. The mixture distribution is defined identically to the local false discovery rate model (*fdr*) proposed by Efron (2004). $f_0(z)$ and $f_1(z)$ are either parametric distributions such as normal distributions or non/semi-parametric (empirical distributions) (Wu et al., 2006; Strimmer, 2008). However, different from the interest of detecting the local false positive rate of the statistic z , our goal is to assign a normalized value $g(z)$ to each connectivity expression metric z (g is the mapping/normalization function).

In the mixture model, we can estimate the probability of z from the non-null component and the null component. Given $p_0, f_0(z), p_1, f_1(z)$, the posterior probability of a connectivity belonging to the non-null component at z is

$$g(z) = p_1 f_1(z) / f(z) = 1 - p_0 f_0(z) / f(z), \quad (2.2)$$

which equals to one minus local false discovery rate *fdr*(z). We use $g(z)$ as the normalization function of the connectivity metric z , which represents the estimated posterior probability of z being truly connected or anticorrelated.

The normalization function $g(z)$ generally yields a higher probability value when z is larger, but rather than a linear relationship it depends on the parameters and distributions of $\{p_0, f_0(z), p_1, f_1(z)\}$. However, in practice, the prior parameters and distributions $\{p_0, f_0(z), p_1, f_1(z)\}$ are unknown and are often estimated from the observed data of z parametrically or nonparametrically.

Our normalization method is called an empirical Bayes method because the normalized connectivity expression is the posterior probability of z from f_1 , and the model parameters $\{p_0, f_0(z), p_1, f_1(z)\}$ are estimated directly rather by sampling from the full conditionals. Fortunately, the estimation techniques for such type of empirical Bayes mixture model have been well-developed and thoroughly discussed (Efron, 2004; Wu

et al., 2006; Strimmer, 2008; Schwartzman et al., 2009). For derivation and discussion of the detailed estimation procedure, we refer the readers to the original papers. Provided with the estimated $\{\widehat{p}_0, \widehat{f}_0(z), \widehat{p}_1, \widehat{f}_1(z)\}$, the estimated normalization function becomes

$$g_s(z) = \widehat{p}_1 \widehat{f}_1(z) / (\widehat{p}_0 \widehat{f}_0(z) + \widehat{p}_1 \widehat{f}_1(z)). \quad (2.3)$$

The normalization function $g_s(z)$ is estimated based on a single subject/image s ($s = 1, \dots, N$, and N is the total number of subjects), as it is determined by the distribution of connectivity expression metrics of each individual. The normalized connectivity expressions are comparable across subjects because they are probability metrics. In general, only high-throughput expression data can include sufficient data points to obtain reliable prior parameter and distribution estimates ($\{\widehat{p}_0, \widehat{f}_0(z), \widehat{p}_1, \widehat{f}_1(z)\}$), hence we would apply the normalization method only when the pairwise connectivity metrics are calculated from at least 70 ROIs. The normalization procedure is conducted prior to the group level statistical inferences such as statistical tests and regressions to ensure the connectivity expression metrics being appropriately scaled and comparable across subjects. The statistical inferences based on normalized connectivity expression metrics could be less affected by the systematic shifts and random measurement errors, and hence are expected to be more robust and reproducible. We will demonstrate the properties of the normalization function in the simulation and data example sections. As the direct assessment of the normalization effect on connectivity metrics (calculated statistics) could be challenging, we examine the normalization method by comparing the statistical inferences based on normalized connectivity metrics and raw (non-normalized) connectivity metrics.

3. Simulations

In this section, we simulate a case-control rs-fMRI study to examine the performance of our normalization method. We generate 30 subjects for each group and within each subject we simulate 4005 correlation coefficients between 90 nodes/regions. We assume that the correlations between the first 30 ROIs are differentially expressed (the control group exhibits higher connections than the case group).

The beta distribution is employed to simulate correlation coefficients because it is more flexible and better resembles the real distribution of correlation coefficients from rs-fMRI data than other distributions (e.g., Gaussian distribution) (Ji et al., 2005; Jantschi and Sorana, 2011). We generate z_1 from the non-null distribution by a transformed Beta distribution: $x_1 \sim \text{Beta}(\alpha_1 = 3, \beta_1 = 3)$ and $z_1 = 1.55x_1 - 0.55$ for correlation coefficients with higher connectivity expression levels; and z_0 from the null distribution by $x_0 \sim \text{Beta}(\alpha_0 = 18, \beta_0 = 18)$ and then $z_0 = x_0 * 2 - 1$. z_1 represent 435 highly expressed correlation coefficients between the first 30 nodes for each subject in the control group, and z_0 represents the rest of correlations for subjects in control group and all correlations for subjects in the case group. In this way, all simulated correlations

range from $[-1, 1]$, and **Figure 2** demonstrates the simulated data for case and control group. Additionally, we use different set of parameters to represent various patterns of correlation distribution (e.g., Murphy et al., 2009) including: (i) more dispersed null component $x_0 \sim \text{Beta}(\alpha_0 = 9, \beta_0 = 9)$ (\mathcal{P}_1); (ii) right skewed connected component $x_1 \sim \text{Beta}(\alpha_1 = 2, \beta_1 = 3)$ (\mathcal{P}_2); (iii) left skewed connected component $x_1 \sim \text{Beta}(\alpha_1 = 3, \beta_1 = 2)$ (\mathcal{P}_3).

In addition, we simulate another scenario by adding systematic shifts across subjects by:

$$\begin{aligned}\mu_s &\sim \text{uniform}(-0.2, 0.2), \\ \tilde{z}_s &= z_s + N(\mu_s, \sigma^2),\end{aligned}$$

where \tilde{z}_s represent the correlations for subject s with systematic shift and the values over -1 or 1 are set to -1 and 1 (**Figure 4A**). We use σ^2 to indicate the magnitude of the shifts.

We apply our normalization method to the simulated correlations with the main goal of differentially expressed connectivity discovery. The R (<http://CRAN.R-project.org/>) package “*locfdr*” is used to estimate the mixture model; and the normalization function $g_s(z)$ is calculated for each individual (see Example in Supplementary Material). **Figure 3** shows that the mixture model is well estimated as well as the shape of the normalization function. Comparing to the original correlation or the variance stabilizing transformation methods (e.g., Fisher’s Z, probit, or logit transformed correlations), the the posterior probability based normalization function incorporates the “false positive” belief with observed connectivity expressions by empirical Bayes framework. The normalized correlations are not related to the original correlations linearly, but monotonely increasing. $g(z)$ increases steeply between around 0.4 and 0.6 because the posterior belief of “true positive” rises drastically. If there are both “true positive” correlation and anticorrelation components, then three components will be detected and estimated and two normalization functions are provided separately for positive and negative correlations (see details in Section 5).

In addition, we compare the raw correlations (without and with subject systematic shifts) with the normalized connectivity expressions to investigate the effects of normalization in connectivity metric quantification and differentially expressed connectivity detection. Different levels of subject systematic shifts (different σ^2) are also included. We first evaluate the effects of normalization on random shifts. If there is a random shift from the random measurement error, **Figure 4A** demonstrates the histograms of the original correlations (red) and systematically (with randomness) shifted correlations (blue) for a subject in the control group. **Figure 4B** illustrates the impact of the systematic shifts on the non-normalized and the normalized connectivity expression. The red histogram in **Figure 4B** shows the difference of original correlations and shifted correlations. Thus, if there is a systematic shift the connectivity will be affected with consistent bias, which may cause invalid group level inferences. The blue histogram in **Figure 4B** shows the differences of normalized original correlations and normalized shifted correlations which are distributed around 0. Clearly, the normalized connectivity

metric is almost invariant to the systematic shifts, therefore the normalization algorithm improves the robustness of the connectivity metrics to systematic shifts/noises.

We then examine the performance of our normalization method on differential expression detection (the main aim). We conduct the two sample Wilcoxon signed-rank non-parametric tests ($\alpha = 0.05$) on the 4005 connectivity metrics z and normalized connectivity $g(z)$ under both non-shifted and shifted scenarios. As we evaluate simultaneous multiple tests, the FDR (with $q = 0.1$ as the threshold) is applied to adjust multiple testing in the simulation study.

Ideally, the test results reveal the 435 “true positives” with 0 “false positives.” **Figure 5** shows the testing results by different methods and scenarios. **Figure 5A** reflects the true differentially expressed connectivity expressions between the two groups for the first 30 nodes (red) and the rest are at the level (blue). **Figure 5B** shows the testing results between the two groups based on the non-normalized correlations. **Figure 5C** are the testing results based on the probit (variance stabilizing) transformed correlations (the logit transformation performance is very similar). **Figure 5D** are the testing results based on the empirical Bayes normalized correlations. **Figures 5E,F** are the test results of non-normalized and normalized correlations under the scenario with systematic shifts. Based on all the differentially expressed connectivity/biomarker discovery results, the normalized connectivity metrics have much lower type I and II errors. **Table 1** summarizes the detailed results with comparison to the truth over 100 times of simulations. The number of false positive testing results of non-normalized correlations is about 17 times of the normalized correlations, and the number of false negative testing results is more than about 20 times; the difference is even larger in the shifted scenario. The performance of probit variance stabilizing transformed correlations are similar to the original correlations. The levels of random shift (σ^2) affect the performance of the differential detection, however after the empirical Bayes normalization the shift almost has no impact on the result findings. Therefore, the simulation study results indicate that our normalization method can effectively scale the connectivity to appropriate level and improves the power to identify the true differentially expressed connectivity with low false positive rate. When the null is more dispersed and connected component is right skewed, the two mixture components are more mixed and thus the false positives and false negatives increase. Yet, our method outperforms the non-normalized correlations for differentially expressed feature detection. Overall, the empirical Bayes normalization model provides a more robust pathway for connectivity expression quantification and enables biomarker discovery with both high sensitivity and specificity.

4. Data Example

This data set was collected at Brain Mapping Center in University of California, Los Angeles (UCLA), one of the data collecting sites in the Autism Brain Imaging Data Exchange (ABIDE) (Rudie et al., 2012, 2013; Di Martino et al., 2014). The imaging was performed on Siemens magneto Trio scanners. The imaging data

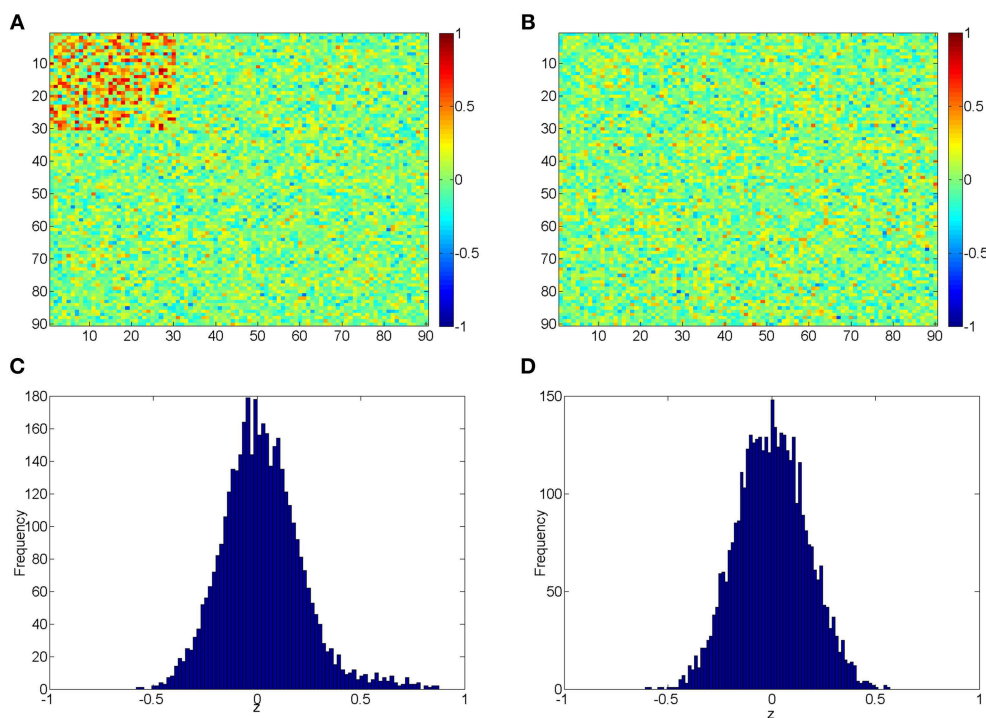


FIGURE 2 | (A,B) Heatmaps of the simulated correlations of the control and case groups; **(C,D)** Histograms of the simulated correlations of the control and case groups.

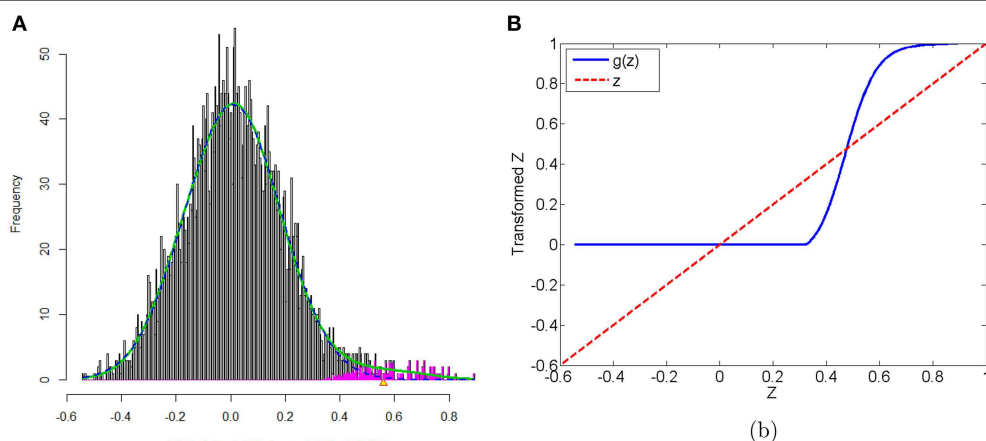


FIGURE 3 | (A) Mixture model estimation procedure using “*locfdr*” in R; **(B)** Normalization function vs. Original correlations: the blue line is the normalization function that maps the raw connectivity metrics to normalized metrics; the red line is used as a reference representing no normalization is applied.

were obtained using a gradient echo T2*-weighted echo planar imaging sequence, echo time $TE = 28$ ms, repetition time $TR = 3$ s, 64×64 matrix with 34 slices 4.0 mm tick, resulting in whole brain coverage with a voxel size of $3 \text{ mm} \times 3 \text{ mm} \times 4 \text{ mm}$. During the MRI scanning, initially 33 participants (typical controls, TC) and 49 patients with the Autism spectrum disorders (ASD) were asked to lie as still as possible, keep their eyes open, try not to fall asleep, and think about whatever they want. A white background with a black central fixation cross was presented

during the resting state scan, although participants were not asked to fixate, they were verified that they had not fallen asleep at the end of the scan. Participants with large motions were removed from the dataset, resulting in 32 participants in the TC group and 41 in the ASD group.

The rs-fMRI data are performed slice time correction and motion correction. The data are registered to a standard MNI space with voxel size 2 mm^3 and is normalized to be percent signal change. The masks of the white matter (WM), the gray

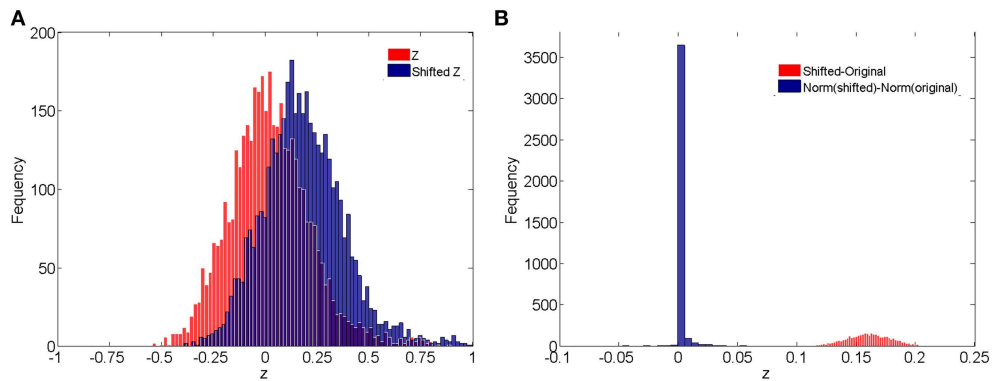


FIGURE 4 | (A) Histograms of correlations and correlations with systematic shifts for a subject in the control group; **(B)** Histograms of the change of the shift.

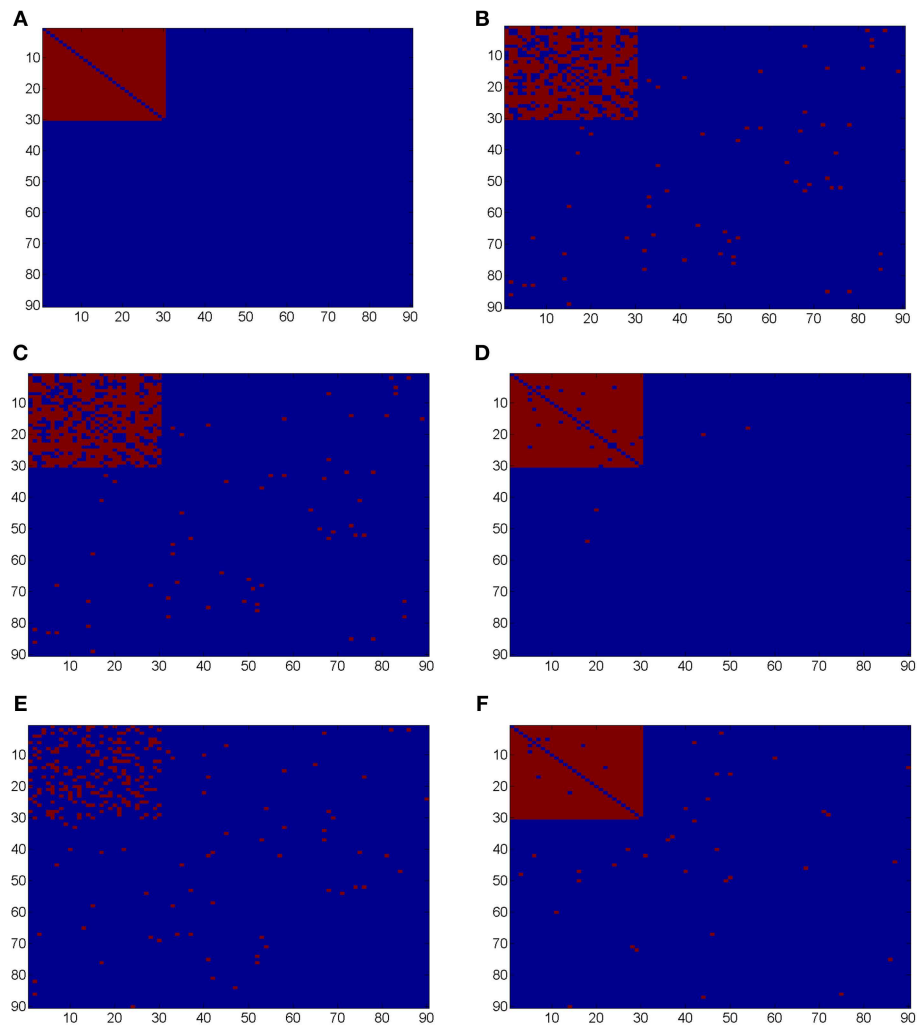


FIGURE 5 | (A) Heatmaps of truth the connectivity between the first 30 nodes are differentially expressed between the two groups; **(B–D)** Heatmaps of the test results using Wilcoxon signed-rank test and FDR control with $q = 0.1$ (red = reject and blue = fail to reject) of the original correlations z , the probit (variance stabilizing) transformed correlation, and the normalized correlations $g_s(z)$, respectively, under the scenario of no systematic shifts; **(E, F)** Heatmaps of the test results of the original correlations z and the normalized correlations $g_s(z)$ under the scenario of with systematic shifts.

TABLE 1 | Results of differential expression tests with normalized and unnormalized correlations (without and with systematic shift): mean and standard deviation of 100 simulations.

Methods	False positives(%)	Std	False negatives(%)	Std
Correlation	32.4 (0.9)	6.1	127.1 (29.2)	9.6
Probit transformed correlation	34.7 (0.9)	6.0	129.4 (29.7)	10.2
Normalized correlation	2.1 (0.06)	0.3	7.3 (1.7)	1.1
Correlation \mathcal{P}_1^a	52.9 (1.48)	8.9	186.5 (52.24)	15.4
Normalized correlation \mathcal{P}_1	3.6 (0.1)	0.6	28.1 (6.5)	3.9
Correlation \mathcal{P}_2	68.3 (1.9)	0.6	236.4 (54.3)	31.6
Normalized correlation \mathcal{P}_2	11.8 (0.3)	3.2	43.8 (10.1)	8.2
Correlation \mathcal{P}_3	16.5 (0.4)	4.4	83.5 (19.2)	8.7
Normalized correlation \mathcal{P}_3	1.3 (0.04)	0.3	2.5 (0.7)	0.5
Correlation + shift ($\sigma^2 = 0.3$)	37.1 (1.04)	7.4	306.5 (70.4)	19.5
Normalized correlation + shift ($\sigma^2 = 0.3$)	17.6 (0.9)	2.2	3.5 (0.8)	0.5
Correlation + shift ($\sigma^2 = 1$)	62.3 (1.74)	9.8	326.4 (74.7)	20.1
Normalized correlation + shift ($\sigma^2 = 1$)	19.3 (0.53)	2.9	4.1 (0.9)	0.7

^aPlease refer to the parameters in paragraph two of the Section 3.

matter (GM), and the cerebrospinal fluid (CSF) are crated in the standard MNI space. The mean time series from the WM and the CSF are calculated. The time series from the GM are regressed out the mean time series of the WM, the CSF and the six movement parameters. A linear trend is removed from all the signal. The fMRI time series are filtered using a bandpass with passing band (0.009–0.08 Hz) and spatially smoothed with 6 mm FWHM Gaussian kernel. We then use the first 90 AAL ROIs as nodes, and take the average of all voxels' temporal profiles within each ROI as region level signal for all subjects (Zalesky et al., 2010b). Four-thousand-five Pearson correlation coefficients are calculated between the 90 nodes, and then Fisher's z transformation are applied. In this analysis, we focus on the differential connectivity expressions between TC and ASD by using normalized connectivity metrics.

We apply the normalization algorithm to all 4005 connectivity metrics for each individual, and no subject in this data set is detected with anticorrelation component of the mixture model. **Figure 6** shows the distribution of correlations for one subject as well as the corresponding empirical Bayes normalization function. Next, we conduct Wilcoxon signed-rank tests toward all 4005 original correlations and normalized correlations between 90 ROIs for TC vs. TSD. We then perform local *fdr* for multiple testing control. Unlike the simulation study, the ground truth of the false positives and false negatives of the data example is unknown. Comparing to the simulation testing results, it seems that the difference between test results of original and normalized correlations has the similar pattern: the normalized connectivity test results include small *p*-values scattered randomly. Because 4005 tests are performed simultaneously, the multiple testing correction methods including local *fdr* and Network Based Statistics (NBS) performed for both empirical Bayes normalized correlations and original correlations (Efron, 2004; Zalesky et al., 2010a). No significant feature or network is identified after the correction for the original correlations (*q*-value 0.1 as threshold

for local *fdr* and permutation *p*-value 0.05 for NBS). In contrast, the analysis based on empirical Bayes normalized connectivity metrics shows significant connectivity differences between the ASD and TC groups, and 44 connectivity features have *fdr q*-values less than 0.1. We demonstrate the results in **Figure 7**. The ASD group show higher function connectivity between pairs of ROIs for all the 44 features than the TC group. Most of these significantly expressed connectivity are between distant ROIs, which are across the the functional subsystems of primary sensory, subcortical, limbic, paralimbic, and association areas defined by Mesulam (1998) and Supekar et al. (2013). We further perform bootstrap analysis to evaluate the reliability of the findings. From 3000 resamples, the 44 features are detected on average 78.6% (with sd 11.3%). As comparison, we detect no connectivity between or within any of these subsystems showing greater connectivity in the TD group, compared with the ASD group. These results suggest that hyper-connectivity in ASD spans multiple functional subsystems of the human brain. The revealed results are consistent with the recent findings of brain hyper-connectivity of ASD children by Supekar et al. (2013), which include multiple studies from three image data acquisition sites in the U.S.

We note that the results can only be identified by using the empirical Bayes normalized connectivity metrics, but not by the original connectivity metrics. Therefore, the normalization step is essential for rs-fMRI based brain connectivity study, and our empirical Bayes normalization method provides a sound pathway to successfully fulfill the task.

5. Discussion

In this article, we have presented a novel empirical Bayes method for rs-fMRI connectivity metric normalization, and the simulation study and the data example have shown that the quantification and statistical inferences based on the normalized inputs are more powerful and reliable. The normalization step

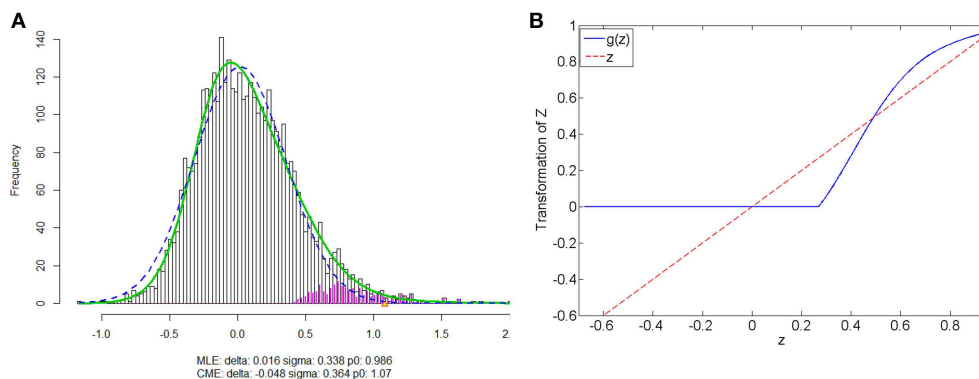


FIGURE 6 | Subject one in data example: (A) is the mixture model estimation procedure using “*locfdr*” in R for Fisher’s *z* transformed correlation; **(B)** the normalization function vs. the original correlations: the blue line is the normalization function that maps the raw connectivity metrics to normalized metrics; the red line is used as a reference representing no normalization is applied.

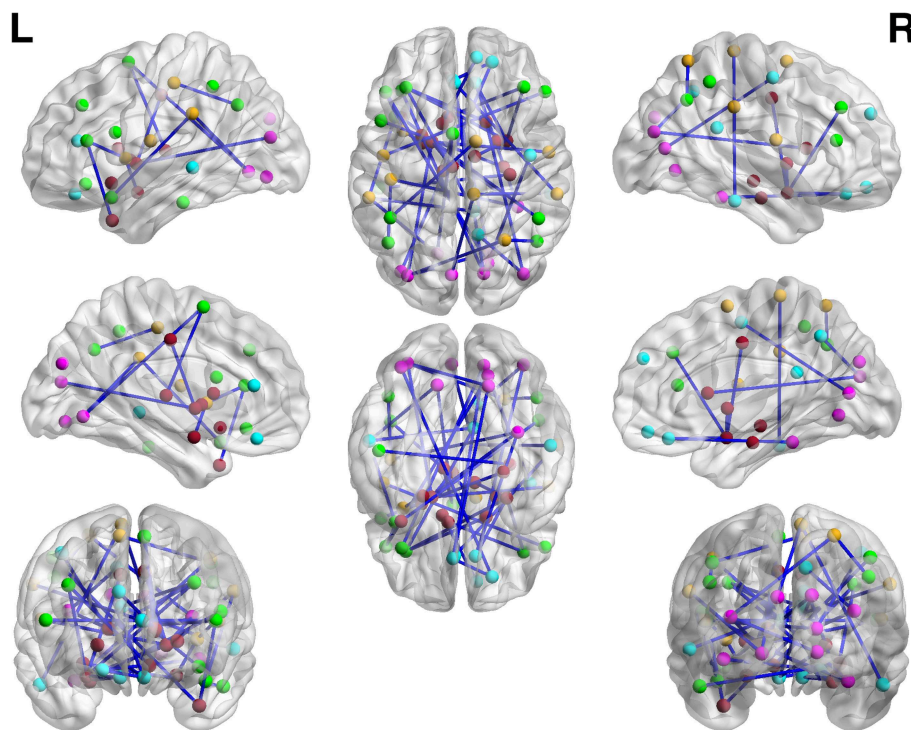


FIGURE 7 | The regions showed higher correlations in children with ASD, compared to the TD group ($q < 0.05$, corrected for multiple comparisons). No pairs of regions showed higher connectivity in the TD than the ASD group.

has been widely used in high-throughput biomedical data analysis with the goal to remove systematic measurement error generated in the complex data acquisition and preprocessing steps and to improve the validity and reproducibility of the following statistical analyses. It has been discussed that a preprocessing step of global signal regression could shift the distributions of the correlations and influence the statistical inferences (Fox et al., 2009; Murphy et al., 2009; Weissenbacher et al., 2009). There may be many other latent factors to affect the quantification of the connectivity metrics as well. Therefore,

we feel that normalization toward connectivity metrics should be introduced.

5.1. Quantification of Brain Functional Connectivity Metrics

Different from the high-throughput “omics” data, the brain functional connectivity is not measured directly but rather calculated by some statistics/metrics based on a pair of time courses from fMRI data. It is unclear how the calculated statistics/metrics can appropriately reflect the true connectivity

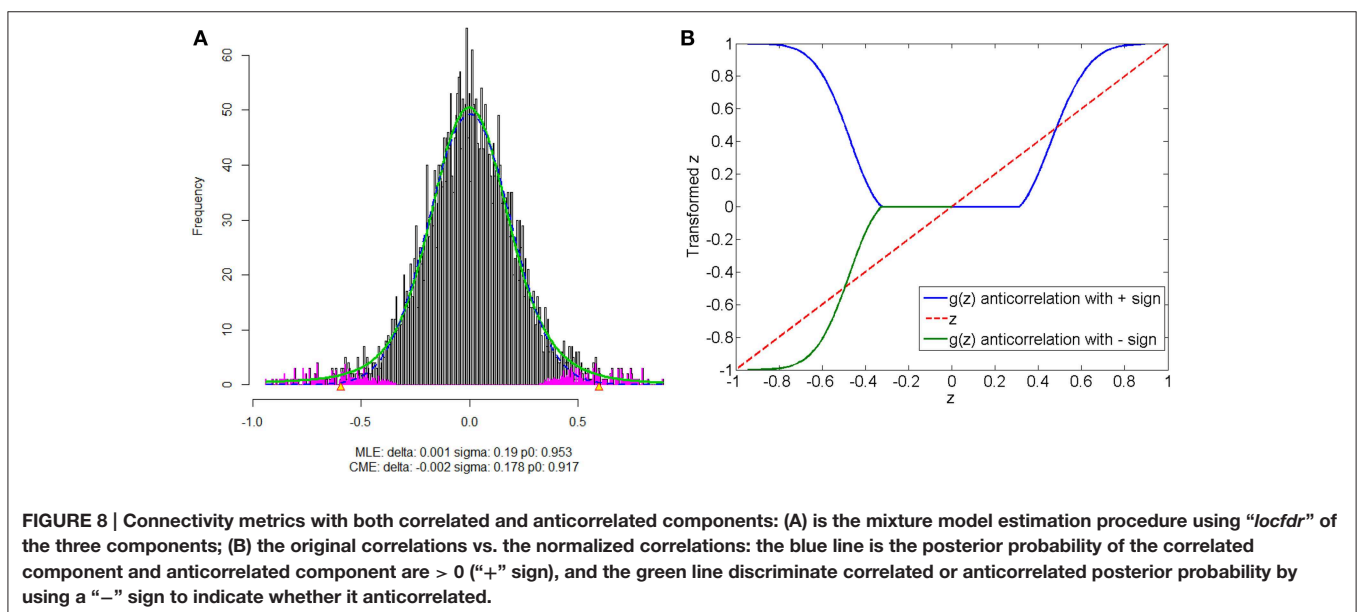
strength and are comparable across subjects, regardless what statistic is chosen (e.g., correlation coefficient or mutual information coefficient). It is possible to obtain extremely large absolute value correlations between two white noise vectors, which gives rise to the false positive discovery. From the statistical perspective, most connectivity statistics can be proved to follow a known distribution asymptotically and accordingly the p -values are calculated with both type I and II errors. Comparing with the conventional normalization method such as quantile normalization, the empirical Bayes mixture model lends itself to incorporating the false positive concept into quantification of the functional connectivity expression and provides a (posterior) probability based scale. The data driven (rather than a deterministic linear/nonlinear transformation) quantification method could provide a more comparable scale for group level connectivity inferences. For example, a 0.1 difference in original correlations could be mapped to around 0.5 difference in the normalized correlations at the interaction between two components due to the increase of posterior probability of true positive. The amplified difference tend to improve the subtle difference detection, because it can better represent connectivity strength. The computational techniques for the mixture model estimation have been developed for local fdr estimation by Efron (2004) and Wu et al. (2006), which provides us a convenient tool to calculate the subject-specific normalization function. The only assumption of our method is that the majority ($p_0 > 0.9$) of connectivity expressions are from the null distribution, which needs to further verified with more rs-fMRI studies. The assumption is generally valid, and all connectivity metric distributions of the data sets we tested follow such pattern. If the assumption is violated, Wu et al. (2006) provides promising numerical solution using nonparametric curve fitting methods. Moreover, another obvious advantage of the normalization method is that it maps the correlations to the range of $[0, 1]$ by the empirical Bayes posterior probability normalization function,

which avoids the information loss due to hard thresholding of correlations in complex network analysis using graph theoretical models (Rubinov and Sporns, 2011).

The appropriate brain connectivity metric normalization method improves the power to detect the truly differentially expressed features and yield less false positive findings. In the simulation study, we compare the test results based on different connectivity metrics with reference to ground truth, and it shows the empirical Bayes normalized correlation has the lowest type I and II errors and is more robust to systematic shifts. When applying our method to the data example, the analysis results based on normalized connectivity metrics detect hyper-connectivity between pairs of regions from distant functional subsystems for the ASD group with comparing to TC group. Such features are not detected by using the non-normalized correlations. The findings align with the results by Supekar et al. (2013) which performs between region connectivity analysis for several autism studies from different sites. Supekar et al. (2013) also provides explanation of these findings from the perspectives of neuroscience and the link to clinical symptoms of ASD. The practical brain connectivity study using neuroimaging technology often involves multiple steps of numerical analysis which are subject to many unavoidable errors and noises, and we feel that the empirical Bayes normalization improves both power and reliability of statistical analysis.

5.2. Anticorrelations

The anticorrelations in rs-fMRI data have drawn attention of many neuroimaging researchers (Fox et al., 2009; Murphy et al., 2009; Weissenbacher et al., 2009; Chai et al., 2012). The discussion has not reached to the agreement whether the anticorrelations are “true positive” or “false positive.” The proposed normalization method provides a pathway to automatically detect the “true positive” anticorrelation



component by classifying the “true positives” and “false positives” based on the empirical distribution of connectivity metric. **Figure 8** shows that correlated and anticorrelated components can be identified, if existing, we could assign either “+” or “−” sign to anticorrelated connectivity metric depending on different following analyses. Generally, “−” sign suits the regression analysis or statistical tests better, because anticorrelation could be considered as the opposite of correlation. When applying the graph theoretical model based network analysis using normalized connectivity, two separate analyses should be conducted for correlations and anticorrelations (with “+” sign) if both components are detected, with the normalized connectivity metric range of [0, 1] (in **Figure 8B**). Thus, the results include two parts of inferences: properties of correlated networks and anticorrelated networks. Although in our data example there is no anticorrelation component detected, that normalization method can be also applied to deal with anticorrelations in practical data analysis. Yet, out normalization method could be combined with pre-processing steps (e.g., global signal regression), as the normalized connectivity is probability and shift-invariant.

6. Conclusion

In summary, a new rs-fMRI connectivity metric normalization method has been developed and applied to functional

brain connectivity analysis. The better connectivity normalization/quantification methods yield generally higher reproducibility. Although we utilize the Pearson correlation coefficient as connectivity metric and rs-fMRI for demonstration, we are optimistic that the developed method are ready to be applied to the task-induced fMRI connectivity study and other connectivity metrics because the empirical Bayes framework is flexible to fit various distributions of connectivity metrics.

Acknowledgments

The authors would thank the researchers of the ABIDE project for sharing their clinical and rs-fMRI data at http://fcon_1000.projects.nitrc.org/indi/abide/. Chen's research is supported in part by UMD Tier1A seed grant. Kang's research was partially supported by NIH grant 1R01MH105561. The authors also want thank Dr. Luiz Pessoa from University of Maryland, College Park and Dr. F. DuBois Bowman from Columbia University for constructive discussions.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnins.2015.00316>

References

- Beckmann, C. F., DeLuca, M., Devlin, J. T., and Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1001–1013. doi: 10.1098/rstb.2005.1634
- Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541.
- Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19, 185–193. doi: 10.1093/bioinformatics/19.2.185
- Bullard, J. H., Purdom, E., Hansen, K. D., and Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* 11:94. doi: 10.1186/1471-2105-11-94
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Braun, U., Plichta, M. M., Esslinger, C., Sauer, C., Haddad, L., Grimm, O., et al. (2012). Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. *Neuroimage* 59, 1404–1412. doi: 10.1016/j.neuroimage.2011.08.044
- Chai, X. J., Castañón, A. N., Ongür, D., and Whitfield-Gabrieli, S. (2012). Anticorrelations in resting state networks without global signal regression. *Neuroimage* 59, 1420–1428. doi: 10.1016/j.neuroimage.2011.08.048
- De Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., and Smith, S. M. (2006). fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage* 29, 1359–1367. doi: 10.1016/j.neuroimage.2005.08.035
- Di Martino, A., Yan, C. G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659–667. doi: 10.1038/mp.2013.78
- Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11073–11078. doi: 10.1073/pnas.0704320104
- Efron, B. (2004). Large-scale simultaneous hypothesis testing. *J. Am. Stat. Assoc.* 99, 96–104. doi: 10.1198/016214504000000089
- Fornito, A., Zalesky, A., Pantelis, C., and Bullmore, E. T. (2012). Schizophrenia, neuroimaging and connectomics. *Neuroimage* 62, 2296–2314. doi: 10.1016/j.neuroimage.2011.12.090
- Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., and Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10046–10051. doi: 10.1073/pnas.0604187103
- Fox, M. D., Zhang, D., Snyder, A. Z., and Raichle, M. E. (2009). The global signal and observed anticorrelated resting state brain networks. *J. Neurophysiol.* 101, 3270–3283. doi: 10.1152/jn.90777.2008
- Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum. Brain Mapp.* 26, 15–29. doi: 10.1002/hbm.20113
- Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. *Curr. Opin. Neurol.* 21, 424–430. doi: 10.1097/WCO.0b013e328306f2c5
- Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, NY: Springer.
- Hansen, K. D., Irizarry, R. A., and Wu, Z. (2012). Removing technical variability in RNA-seq data using conditional quantile normalization. *Biostatistics* 13, 204–216. doi: 10.1093/biostatistics/kxr054
- Ji, Y., Wu, C., Liu, P., Wang, J., and Coombes, K. R. (2005). Applications of beta-mixture models in bioinformatics. *Bioinformatics* 21, 2118–2122. doi: 10.1093/bioinformatics/bti318
- Jantschi, L., and Sorana, D. B. (2011). Distributing correlation coefficients of linear structure-activity/property models. *Leon. J. Sci.* 10, 27–48.

- Mesulam, M. M. (1998). From sensation to cognition. *Brain* 121, 1013–1052.
- Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., and Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* 44, 893–905. doi: 10.1016/j.neuroimage.2008.09.036
- Murphy, K., Birn, R. M., and Bandettini, P. A. (2013). Resting-state fMRI confounds and cleanup. *Neuroimage* 80, 349–359. doi: 10.1016/j.neuroimage.2013.04.001
- Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11:r25. doi: 10.1186/gb-2010-11-3-r25
- Rubinov, M., and Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *Neuroimage* 56, 2068–2079. doi: 10.1016/j.neuroimage.2011.03.069
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Rudie, J. D., Hernandez, L. M., Brown, J. A., Beck-Pancer, D., Colich, N. L., Gorrindo, P., et al. (2012). Autism-associated promoter variant in *MET* impacts functional and structural brain networks. *Neuron* 75, 904–915. doi: 10.1016/j.neuron.2012.07.010
- Rudie, J. D., Brown, J. A., Beck-Pancer, D., Hernandez, L. M., Dennis, E. L., Thompson, P. M., et al. (2013). Altered functional and structural brain network organization in autism. *Neuroimage* 2, 79–94. doi: 10.1016/j.nicl.2012.11.006
- Smith, S. M. (2012). The future of FMRI connectivity. *Neuroimage* 62, 1257–1266. doi: 10.1016/j.neuroimage.2012.01.022
- Schwartzman, A., Dougherty, R. F., Lee, J., Ghahremani, D., and Taylor, J. E. (2009). Empirical null and false discovery rate analysis in neuroimaging. *Neuroimage* 44, 71–82. doi: 10.1016/j.neuroimage.2008.04.182
- Strimmer, K. (2008). A unified approach to false discovery rate estimation. *BMC Bioinformatics* 9:303. doi: 10.1186/1471-2105-9-303
- Supekar, K., Uddin, L. Q., Khouzam, A., Phillips, J., Gaillard, W. D., Kenworthy, L. E., et al. (2013). Brain hyperconnectivity in children with autism and its links to social deficits. *Cell Rep.* 5, 738–747. doi: 10.1016/j.celrep.2013.10.001
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Weissenbacher, A., Kasess, C., Gerstl, F., Lanzenberger, R., Moser, E., and Windischberger, C. (2009). Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. *Neuroimage* 47, 1408–1416. doi: 10.1016/j.neuroimage.2009.05.005
- Wu, B., Guan, Z., and Zhao, H. (2006). Parametric and nonparametric FDR estimation revisited. *Biometrics* 62, 735–744. doi: 10.1111/j.1541-0420.2006.00531.x
- Zalesky, A., Fornito, A., and Bullmore, E. T. (2010a). Network-based statistic: identifying differences in brain networks. *Neuroimage* 53, 1197–1207. doi: 10.1016/j.neuroimage.2010.06.041
- Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., et al. (2010b). Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* 50, 970–983. doi: 10.1016/j.neuroimage.2009.12.027
- Zalesky, A., Fornito, A., and Bullmore, E. (2012). On the use of correlation as a measure of network connectivity. *Neuroimage* 60, 2096–2106. doi: 10.1016/j.neuroimage.2012.02.001
- Zhou, D., Thompson, W. K., and Siegle, G. (2009). MATLAB toolbox for functional connectivity. *Neuroimage* 47, 1590–1607. doi: 10.1016/j.neuroimage.2009.05.089

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Chen, Kang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Parcellation Based Nonparametric Algorithm for Independent Component Analysis with Application to fMRI Data

Shanshan Li^{1*}, Shaojie Chen², Chen Yue² and Brian Caffo²

¹ Department of Biostatistics, Indiana University Fairbanks School of Public Health, Indiana University, Indianapolis, IN, USA,

² Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

OPEN ACCESS

Edited by:

Bertrand Thirion,
Institut National de Recherche en
Informatique et Automatique, France

Reviewed by:

Xi-Nian Zuo,
Chinese Academy of Sciences, China
Alexis Roche,
Siemens Healthcare/Centre
Hospitalier Universitaire Vaudois,
Switzerland
Ronald Phlypo,
Grenoble Images Parole Signal
Automatique-Lab, France

*Correspondence:

Shanshan Li
sl50@iu.edu

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 14 May 2015

Accepted: 11 January 2016

Published: 29 January 2016

Citation:

Li S, Chen S, Yue C and Caffo B
(2016) A Parcellation Based
Nonparametric Algorithm for
Independent Component Analysis
with Application to fMRI Data.
Front. Neurosci. 10:15.
doi: 10.3389/fnins.2016.00015

Independent Component analysis (ICA) is a widely used technique for separating signals that have been mixed together. In this manuscript, we propose a novel ICA algorithm using density estimation and maximum likelihood, where the densities of the signals are estimated via p-spline based histogram smoothing and the mixing matrix is simultaneously estimated using an optimization algorithm. The algorithm is exceedingly simple, easy to implement and blind to the underlying distributions of the source signals. To relax the identically distributed assumption in the density function, a modified algorithm is proposed to allow for different density functions on different regions. The performance of the proposed algorithm is evaluated in different simulation settings. For illustration, the algorithm is applied to a research investigation with a large collection of resting state fMRI datasets. The results show that the algorithm successfully recovers the established brain networks.

Keywords: blind source separation, density estimation, functional MRI, p-spline bases, signal processing

1. INTRODUCTION

This manuscript puts forward two innovations. Firstly, we demonstrate a fast, likelihood motivated and straightforward method for applying independent components analysis (ICA). Secondly, we propose a parcellation based adjustment when the source signals distribute differently across regions. Our work is routed in the context of understanding human brain networks, and we use functional magnetic resonance imaging (fMRI) data for illustration in this manuscript.

We approach our study of fMRI by simultaneously analyzing all voxels. This is in contrast to regional or seed-based approaches (Buckner et al., 2005; Wang et al., 2006; Allen et al., 2007) that restrict attention to carefully chosen locations. Such approaches require strong assumptions on the choice of seeds or parcellation used to define region. Hence voxel-wise approaches are important complementary procedures. Given the volume of voxels under study (usually on the order of fifty thousand non-background ones), flexible yet parsimonious models approaches are required. However, even with parsimonious models, whole brain voxel-level techniques are more empirical and exploratory than their more hypothesis driven regional and seed-based counterparts. Thus, exploratory factor-analytic models are common approaches in voxel-level investigations.

Independent components analysis (ICA) is a factor-analytic approach that has been frequently utilized for the analysis of functional neuroimaging data, because of its success in discovering

important brain networks in many applications (McKeown et al., 1998a,b; Calhoun et al., 2001a; Guo and Pagnoni, 2008). Two key benefits of ICA are its exploratory nature and its often considered reasonable underlying generative model. Specifically, it models collected signals, X , as linear weighted combinations of independent sources, S_1, S_2, \dots, S_p . Thus, we can write the noise-free ICA model as $X = AS$, where $S = [S_1, S_2, \dots, S_Q]$ and A is a $Q \times Q$ full rank matrix, the so-called mixing matrix. The goal of ICA is to recover the underlying signals S_1, S_2, \dots, S_Q from their observed mixtures X_1, X_2, \dots, X_Q . Note that, in the context of fMRI, the independent components S_1, S_2, \dots, S_Q are often interpreted as brain networks and A is the mixing matrix characterizing the temporal pattern of the corresponding brain networks.

Various algorithms for ICA have been proposed in the literature. See Hyvärinen et al. (2001); Comon and Jutten (2010); Risk et al. (2014) for comprehensive reviews. One common procedure is to postulate a parametric family for the source distributions and then obtain the independent components (ICs) by optimizing a contrast function that measures the distribution property of the output (Samworth and Yuan, 2012). The contrast functions can be selected based on different measures, such as entropy, mutual independence, high-order decorrelations, divergence between the joint distribution of the output and some model, etc. (Cardoso, 1998). These include the popular FastICA algorithm (Hyvärinen and Oja, 2000), the JADE algorithm (Cardoso, 1999), the Pearson ICA algorithm (Karvanen et al., 2000), and a few other algorithms proposed by Comon (1994); Amari and Cardoso (1997); Li and Adali (2010). An alternative procedure is to assume smooth densities for the source distributions and use nonparametric or semiparametric approaches to estimate those density functions. The mixing matrix can then be derived using maximum likelihood method. For example, Bach and Jordan (2003) developed a nonparametric estimation approach based on canonical correlations in a reproducing kernel Hilbert space. Hastie and Tibshirani (2002) expressed the source distribution by an exponentially tilted Gaussian function and used the fixed-point algorithm for estimation of the mixing matrix. Boscolo et al. (2004) used kernel density estimation techniques to model the underlying densities and quasi-Newton method for optimization. Guo and Pagnoni (2008) used Gaussian mixture models for the source distribution and provided an expectation-maximization (EM) framework for estimation, assuming Gaussian noise in the model. Eloyan et al. (2013) estimated the source distribution by using mixture density estimates, and proposed a constrained EM algorithm for estimation.

The benefit of the likelihood-based ICA algorithm is that, as a byproduct of the ICA algorithm, one obtains the fully specified likelihood of the ICA model which can be used for further statistical inference. For example, based on the fully specified likelihood, one can conduct Bayesian analysis or perform likelihood based model selection. However, the existing likelihood-based ICA algorithms are mostly semi-parametric and are usually computationally intensive. In this manuscript, we aim to develop a likelihood-based algorithm that is exceedingly simple and truly blinded to the source distributions.

We propose to estimate the density function of the ICs via histogram smoothing, following a well-known approach in the penalized spline literature. At its core, likelihood-based ICA requires estimation of the mixing matrix and flexible density estimation for the ICs. Our approach, like many other likelihood-based approaches, iteratively estimates these components separately using block maximization. In contrast to other approaches, we use an exceedingly simple density estimation technique via histogram smoothing. Specifically, we assume the bin counts of the frequency histogram follow a Poisson distribution and express the mean counts as sum of B-spline bases via generalized linear model. To smooth the histogram, we follow Eilers and Marx (1996) to construct a penalized likelihood with a difference penalty on coefficients of adjacent B-splines. Apart from its simplicity, a benefit of this approach is speed. Density estimation and evaluation for tens of thousands of voxels is time consuming, and worse, is performed within an iterative algorithm. Using histogram smoothing, the voxel-level calculation reduces to estimating a histogram, a very fast process.

We briefly mention that, in our primary area of application, fMRI, we focus entirely on noise-free group spatial independent component analysis. By assuming noise-free model, noise in the data is absorbed into the estimated ICs and the mixing matrix. By using spatial ICA model, the fMRI data is decomposed into spatial maps multiplied by their respective time courses, where the maps are drawn from spatial distributions that are statistically independent (Calhoun et al., 2001a). The spatial independence assumption is well suited to the sparse nature of the spatial pattern for typical brain activation (McKeown and Sejnowski, 1998; Guo and Pagnoni, 2008). The time courses estimated from spatial ICA describe the temporal characteristics of functional networks, i.e., areas of temporal correlation in the BOLD signal. For multi-subject fMRI data, we assume common spatial maps for all subjects and subject-specific mixing matrices, therefore, we can concatenate all subjects' data in the temporal domain, and apply ICA to the aggregated data matrix. The group mixing matrix is the concatenated time course for all subjects. Individual mixing matrices can be backreconstructed by partitioning the group mixing matrix into submatrices corresponding to each subject.

The remainder of the paper is organized as follows. Section 2 describes the p-spline based ICA algorithm and considers relaxation of the *i.i.d* signal assumption. Section 3 shows the performance of the proposed algorithm in simulation study. Section 4 provides the application of the proposed algorithm to the 1000 Functional Connectome Project (https://www.nitrc.org/projects/fcon_1000/), while Section 5 gives a discussion.

2. METHODS

2.1. Description of ICA Methodology

Independent component analysis models collected signals as linear weighted combinations of independent sources. Notationally, let X_i be a $T \times V$ matrix for subject $i = 1, \dots, I$. In the context of fMRI, T indicates scans while V indicates

voxels. Assume the number of ICs is Q . The ICA model specifies $X_i = A_i S$, where A_i is a $T \times Q$ mixing matrix and S is a $Q \times V$ matrix of ICs. By assuming common spatial maps across subjects, we can stack the individual matrices in the temporal domain. Let $X = [X_1^T, X_2^T, \dots, X_I^T]^T$ be the $TI \times V$ group data matrix, and $A = [A_1^T, A_2^T, \dots, A_I^T]^T$ be the $TI \times Q$ group mixing matrix. Spatial group ICA simply specifies the standard model

$$X = AS. \quad (1)$$

We use parentheses to index matrices so that $X(t, v)$ is element (t, v) of X and define $X(t, \cdot)$ as row t of X and $X(\cdot, v)$ as column v . Then, model (1) could be rewritten as $X(t, v) = \sum_{q=1}^Q A(t, q)S(q, v)$ and $X = \sum_{q=1}^Q A(\cdot, q)S(q, \cdot)$.

We assume that $E[X] = \mu_x = 0$ and hence $E[S] = \mu_s = 0$. If this assumption were not made, the ICA model would imply $X - \mu_x = A(S - \mu_s)$, which is exactly an ICA model with a centered data matrix and the ICs having mean 0. Hence, X is demeaned prior to analyses and μ_s is assumed to be zero. Similarly, since $A(\cdot, q)S(q, \cdot) = \{A(\cdot, q)/c\} * \{cS(q, \cdot)\}$, ICs are only identified up to scalar multiplication. Thus, we assume that $\text{Var}\{S(q, v)\} = 1$ for $q = 1, \dots, Q$ and $v = 1, \dots, V$.

ICA gets its name by assuming that $S(q, \cdot) \perp\!\!\!\perp S(q', \cdot)$ when $q \neq q'$, where $\perp\!\!\!\perp$ implies statistical independence. However, standard variations of ICA also assumes that $\{S(q, v)\}_{v=1}^V$ is an *i.i.d* collection, which we also adopt for now. The *i.i.d* assumption will be relaxed later in the next subsection. As a consequence of these assumptions, $X(\cdot, v) \perp\!\!\!\perp X(\cdot, v')$ when $v \neq v'$; yet note that $X(t, \cdot)$ is not (necessarily) independent of $X(t', \cdot)$.

Typically, $Q < TI$ and Equation (1) is overdetermined. A two-stage dimension reduction is often performed to reduce the computational load and avoid overfitting (Calhoun et al., 2001a; Beckmann and Smith, 2005; Guo and Pagnoni, 2008; Eloyan et al., 2013; Risk et al., 2014). Specifically, in the first stage, an SVD is performed in the temporal domain within subject, where the first R eigenvectors are retained. The dimension for the group data matrix then becomes $RI \times V$. In the second stage, an SVD is performed on the group data matrix obtained from the first stage and the first Q eigenvectors are retained to force a determined linear system for the group ICA model. This discards information in the data. However, one hopes that by selecting the first Q singular vectors, the most relevant features of the data will be retained. The choice of R and Q could be based on various criteria, including variance explained, information-theoretic criteria, and practical considerations. This is not a major concern in this article.

2.2. ICA Through Fast Nonparametric Density Estimation

ICA estimates S by seeking an unmixing matrix, say \hat{B} , such that $\hat{B}X$ is a good approximation to the original sources S . Let $B = A^{-1}$ be the estimand of interest. Notationally following Hyvärinen et al. (2001), if f_q is the density for $S(q, v)$ for $v = 1, \dots, V$, and $f = (f_1, \dots, f_Q)$, then standard multivariate random variable transformation results imply that the joint density of

$X(\cdot, v)$ is

$$\begin{aligned} g\{X(\cdot, v)\} &= |\det(B)| \prod_{q=1}^Q f_q\{S(q, v)\} \\ &= |\det(B)| \prod_{q=1}^Q f_q\{B(q, \cdot)X(\cdot, v)\}, \end{aligned}$$

therefore the joint log-likelihood including all contributions for $v = 1, \dots, V$ is

$$\mathcal{L}(B, f) = \sum_{v=1}^V \sum_{q=1}^Q \log[f_q\{B(q, \cdot)X(\cdot, v)\}] + V \log|\det(B)|.$$

It is generally not possible to solve the joint likelihood for the parameters in f_q and B simultaneously. Instead, an iterative optimization is often performed. Specifically, given the current estimate of B at iteration k , say $\hat{B}^{(k)}$, one can get an estimate for S via $\hat{S}^{(k)} = \hat{B}^{(k)}X$. Given $\hat{S}^{(k)}(q, \cdot)$, density estimation techniques can be used to obtain $\hat{f}_q^{(k)}$, the estimate of f_q at iteration k .

We suggest the use of histogram smoothing as the density estimation technique, where the bin counts of the frequency histogram are assumed to follow a Poisson distribution and a penalized likelihood is constructed to produce a smooth density estimate. The details of histogram smoothing can be found in Eilers and Marx (1996), and we provide a sketch below. (Readers not familiar with statistical smoothing may skip the rest of this paragraph). Notationally, let $c^{(k)}(q, 0) < c^{(k)}(q, 1) < \dots < c^{(k)}(q, J)$ be equidistant histogram cutpoints, where $c^{(k)}(q, 0) = -\epsilon + \min \hat{S}^{(k)}(q, \cdot)$ and $c^{(k)}(q, J) = \epsilon + \max \hat{S}^{(k)}(q, \cdot)$. The number ϵ is added to avoid numerical boundary effects. Let $n^{(k)}(q, j) = \sum_{v=1}^V I\{c^{(k)}(j-1) < \hat{S}^{(k)}(q, \cdot) \leq c^{(k)}(j)\}$, for $j = 1, \dots, J$, be the count of values between cutpoints $j-1$ and j for row q of $\hat{S}^{(k)}$. Define the midpoints of intervals $[c^{(k)}(q, j-1), c^{(k)}(q, j)]$ by $m^{(k)}(q, j)$ for $j = 1, \dots, J$. We obtain a density estimate via the log-linear model $n^{(k)}(q, j) \sim \text{Poisson}\{\lambda^{(k)}(q, j)\}$, where $\log\{\lambda^{(k)}(q, \cdot)\} = \sum_{l=1}^L D^{(k)}\{m^{(k)}(q, \cdot), l\}\beta^{(k)}(q, l)$. Here the log function is presumed to act component-wise on vectors, $D^{(k)}$ is a B-spline basis design matrix, L is the number of knots for B-splines, and $\beta^{(k)}(q, \cdot)$ is a vector of coefficients. To avoid overfitting the B-spline model, and to avoid sensitivity to the degrees of freedom, we choose a large value for the degrees of freedom and put a squared penalty on the coefficients. Let $\mu^{(k)}(q, j)$ denote the expectation of $n^{(k)}(q, j)$, then the penalized log likelihood takes the form (Eilers and Marx, 1996)

$$\begin{aligned} \mathcal{L} &= \sum_{j=1}^J n^{(k)}(q, j) \ln \mu^{(k)}(q, j) - \sum_{j=1}^J \mu^{(k)}(q, j) \\ &\quad - \delta \sum_{l=3}^L \frac{\{\Delta^2 \beta^{(k)}(q, l)\}^2}{2}, \end{aligned}$$

where δ is a parameter controlling the smoothness of the fit, Δ denotes the difference operator, $\Delta^2 \beta^{(k)}(q, l) = \beta^{(k)}(q, l) - 2\beta^{(k)}(q, l-1) + \beta^{(k)}(q, l-2)$. The resulting model is then a generalized linear

mixed model on the counts. The B-spline basis is evaluated at the midpoint of the cutpoint interval. However, via interpolation, the smoother gives an estimate for all values, thus yielding a continuous function, say $\hat{f}_q^{(k)}(s)$, which is the density estimate.

Using generalized linear mixed models to penalize smoothing has become standard practice and is well described in Ruppert et al. (2003). Histogram smoothing as a density estimate appears to be less commonly used. However, we note that this p-spline based density smoother has very attractive properties (Eilers and Marx, 1996). First, it results in a proper density. Secondly, it elegantly handles boundary issues, unlike other density estimators (such as kernel density estimator). Thirdly, the estimated density conserves the first few empirical moments (means and variances) of the histogram, depending on the order of the B-splines. More details regarding these properties can be found in Eilers and Marx (1996). Note that, conservation of moments is an important property that guarantees the identifiability of the ICA model. We choose a cubic B-spline which then conserves the first two moments of the histogram.

Furthermore, due to the convenient differentiation properties of B-spline bases and the simple exponential (Poisson) model, the first and second derivatives of $\hat{f}_q^{(k)}$ are immediately available, where $\hat{f}_q^{(k)} = \exp\{\hat{f}_q^{(k)}\} \beta^{(k)}(q, \cdot) dD^{(k)}$. Thus, derivatives of $\mathcal{L}(B)$ are available in closed form, making gradient- and Hessian-based optimization algorithms easy to implement. This is useful for the stage of the algorithm for obtaining the next iterate of B . Accordingly, we use a Newton-Raphson method to update the mixing matrix. Specifically, let \mathcal{L}' and \mathcal{L}'' denote the first and second derivatives of the log likelihood. At the k th iteration, we update B by

$$B^{(k+1)} = B^{(k)} - \mathcal{L}''(B^{(k)})^{-1} \mathcal{L}'(B^{(k)}). \quad (2)$$

The starting values of B should satisfy the condition that the underlying ICs are the same for all subjects. Following Eloyan et al. (2013), we decompose the full matrix X using the population value decomposition $X = U\Sigma V^T$ (Crainiceanu et al., 2011), and the starting values of the B_i are chosen as the i th block of the rows of $U\Sigma$. Thus, given a starting value for B , histogram smoothing is used to obtain f_q , then given the update for f_q , the natural gradient algorithm is used to obtain B and these steps are iterated until convergence. Let P denote $B^{(k)}(B^{(k+1)})^{-1}$. We use the Amari metric between $B^{(k+1)}$ and $B^{(k)}$ as our convergence criterion (Amari, 1998), where the metric is defined as

$$d\{B^{(k)}, B^{(k+1)}\} = \frac{1}{2Q} \sum_{i=1}^Q \left(\sum_{j=1}^Q \frac{|P_{ij}|}{\max_j |P_{ij}|} - 1 \right) + \frac{1}{2Q} \sum_{j=1}^Q \left(\sum_{i=1}^Q \frac{|P_{ij}|}{\max_i |P_{ij}|} - 1 \right).$$

The Amari metric is useful, as it is invariant to permutation of the ordering of the ICs, a necessary condition for a convergence metric to be useful.

2.3. ICA Based on Parcellation

Most ICA algorithms (such as the commonly used fastICA, JADE, etc.) assume that $\{S(q, v)\}_{v=1}^V$ is an *i.i.d* collection for all $q = 1, \dots, Q$. Intrinsically, this is to assume that the values of the ICs are independent draws from a density. The *i.i.d* assumption is made for simplicity, but it may not hold for fMRI data. Calhoun et al. (2001b) considered possible violations of the independence assumption for task-based fMRI data. They found that the ICA algorithm was successful when the correlation in the signal was small, but it may fail when the signals are highly dependent. However, for most task-based fMRI and resting-state fMRI data, the correlation between voxels is negligible. Therefore, we do not pursue the approach to deal with violation of the independence assumption here. Instead, we consider relaxation of the identically distributed assumption.

Specifically, we propose to account for the difference in the activity across the brain by allowing different density distribution in different regions. To this end, we adopt the functional parcellation of the brain activity map proposed by Yeo et al. (2011). The parcellation includes 17 functional networks in the cerebral cortex, that is, $I = 18$ ROIs for the whole brain. We assume the signals are *i.i.d* within region but could be differently distributed across region. Under this assumption, the density function f_q can be written as the sum of the region-specific density function, that is,

$$f_q(s) = \sum_{i=1}^I I(s \in R_i) f_{iq}(s),$$

where R_i denotes the i th ROI, f_{iq} is the density function on R_i . Thus, f_{iq} takes positive values on the i th region and zero elsewhere. The density estimate of f_{iq} can be obtained using the same procedure as proposed in Section 2.2, confined to the i th region. The estimate for f_q can be constructed by taking the sum of \hat{f}_{iq} . The rest of the ICA algorithm follows the proposed procedure in Section 2.2.

The proposed ICA algorithm can be summarized as follows:

1. Choose an initial value of for the mixing matrix B .
2. Alternate until convergence of B using the Amari metric.
 - a. Let $S = BX$.
 - b. For each IC q , calculate the density function $f_{iq}(s)$ on the i th ROI, $i = 1, 2, \dots, I$, using the p-spline based density estimation algorithm.
 - c. Get $f_q(s) = \sum_{i=1}^I I(s \in R_i) f_{iq}(s)$.
 - d. Update the mixing matrix B using the Newton-Raphson method, see Equation (2).

Note that, in the special case that $f_{1q} = f_{2q} = \dots = f_{Iq}$, the above algorithm reduces to the algorithm proposed in Section 2.2 assuming *i.i.d* signals across the entire brain.

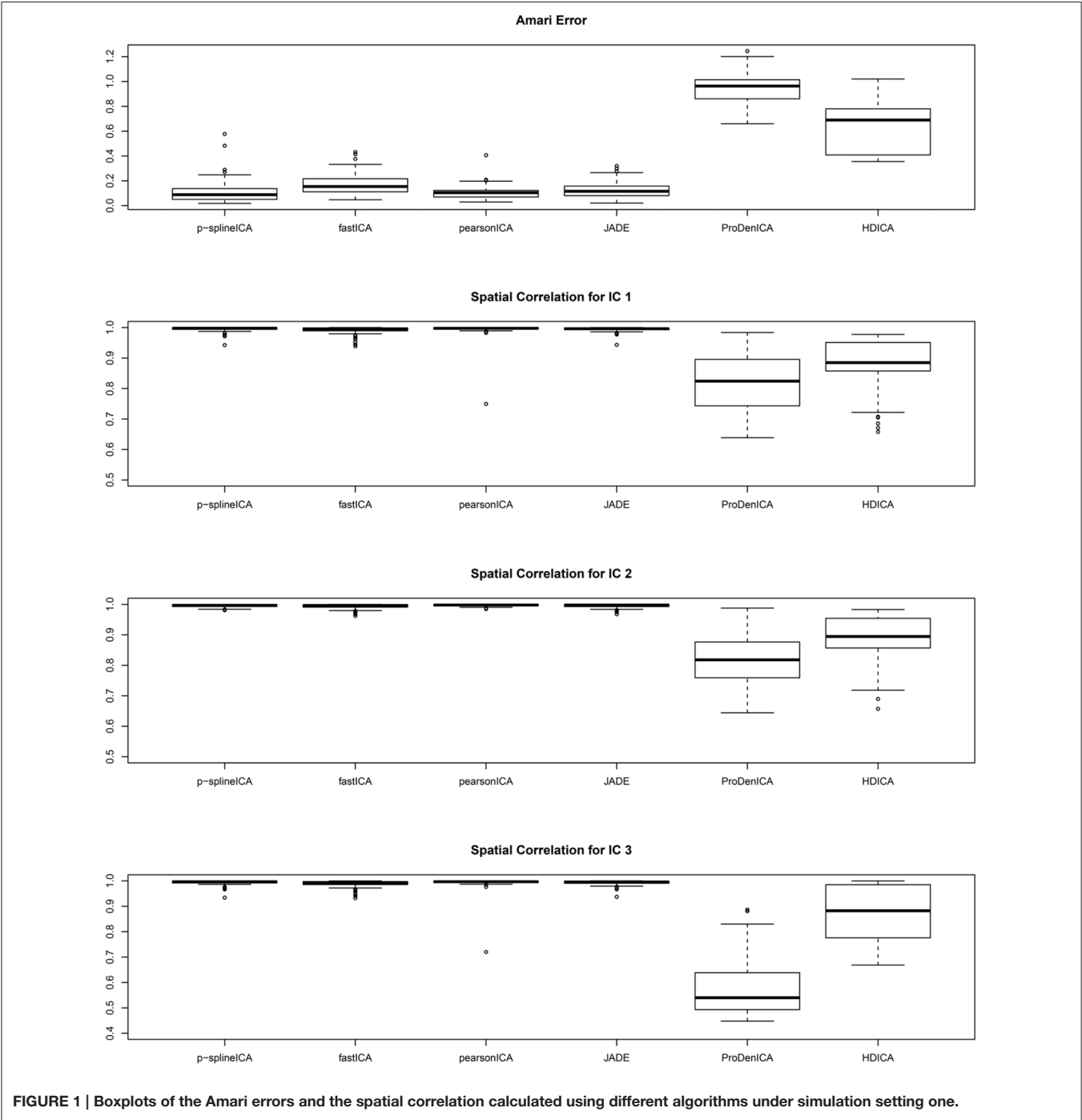
3. SIMULATION

We conduct simulation studies to evaluate the performance of the proposed ICA algorithm. We consider four settings where

TABLE 1 | The average computation time (in seconds) per simulation replication using different algorithms in the simulation study.

Simulation	p-spline ICA	fastICA	Pearson ICA	JADE	ProDenICA	HDICA
Scenario 1	6.21	< 0.01	0.01	0.01	3.12	308.34
Scenario 2	5.89	< 0.01	0.01	< 0.01	1.67	76.59
Scenario 3	7.27	0.01	0.01	0.01	1.74	123.84

data are generated using different distributions. We compare the results of the proposed algorithm with fastICA (Hyvärinen et al., 2001), JADE (Cardoso, 1999), Pearson ICA (Karvanen et al., 2000), ProDenICA (Hastie and Tibshirani, 2002), and HDICA (Eloyan et al., 2013). We implement the algorithms fastICA, ProDenICA, JADE, PearsonICA using the R packages “fastICA” (Marchini et al., 2013), “ProDenICA” (Hastie and Tibshirani, 2010), “JADE” (Nordhausen et al., 2014), and “PearsonICA”



(Karvanen, 2006). The proposed p-spline based ICA algorithm and the HDICA (Eloyan et al., 2013) are also implemented in R.

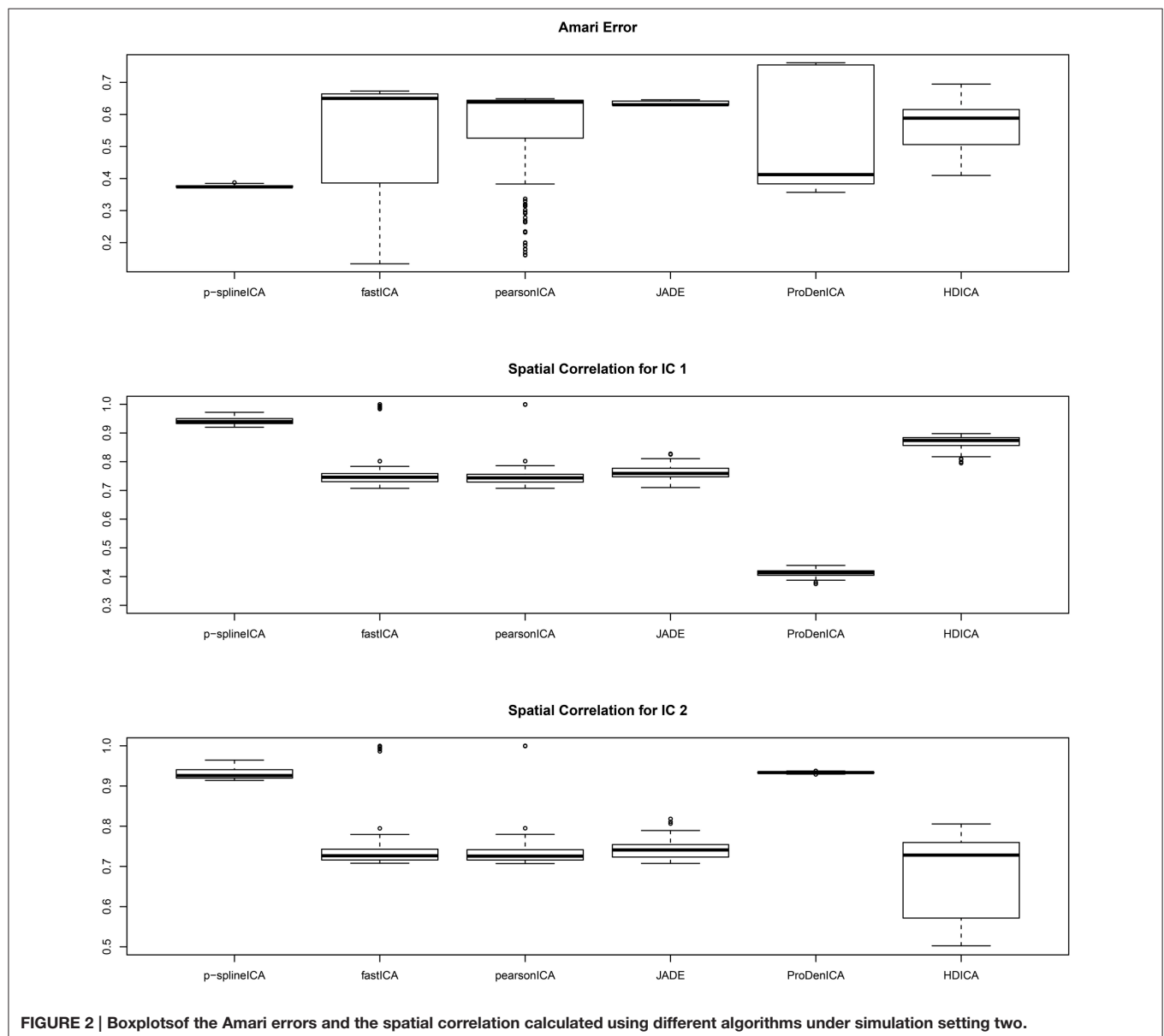
The computation environment is a multi-core Linux cluster with more than 680 cores running in the average of 2.5 GHz speed and 4.4 TB of memory. On average, the contrast-function based algorithms (fastICA, PearsonICA, JADE) perform much faster than the likelihood-based algorithms (p-spline ICA, ProDenICA, HDICA). (See **Table 1** for a summary of the computation time using different algorithms.) However, since those are essentially two different sets of algorithms, we restrict the comparison of the computational intensity within the category of likelihood-based algorithm.

In the first set of simulation studies, we assume there are $Q = 3$ independent components, and they are generated by $S(1, \cdot) \sim \text{Weibull}(1, 1)$, $S(2, \cdot) \sim \text{Gamma}(1, 1)$, and $S(3, \cdot) \sim$

$\text{Gamma}(2, 2)$, respectively. Standard Gaussian noises are added to the generated ICs. The mixing matrix is given by

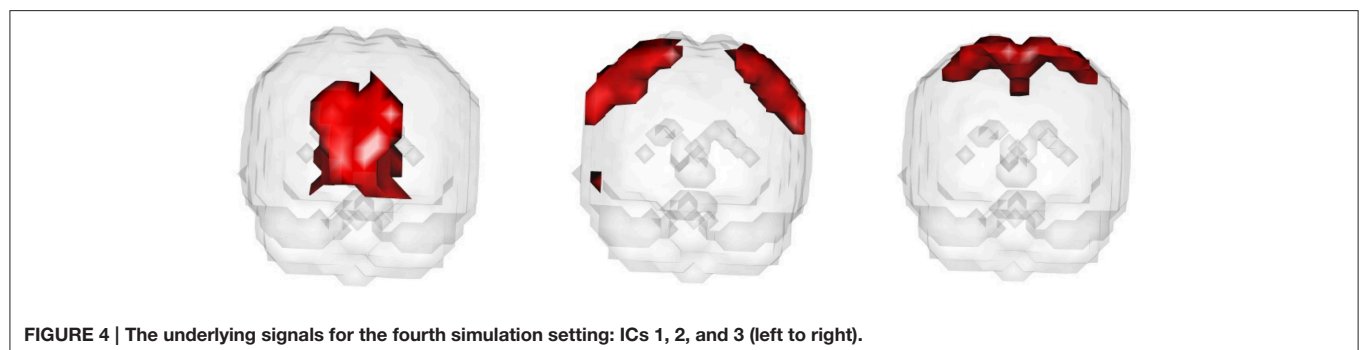
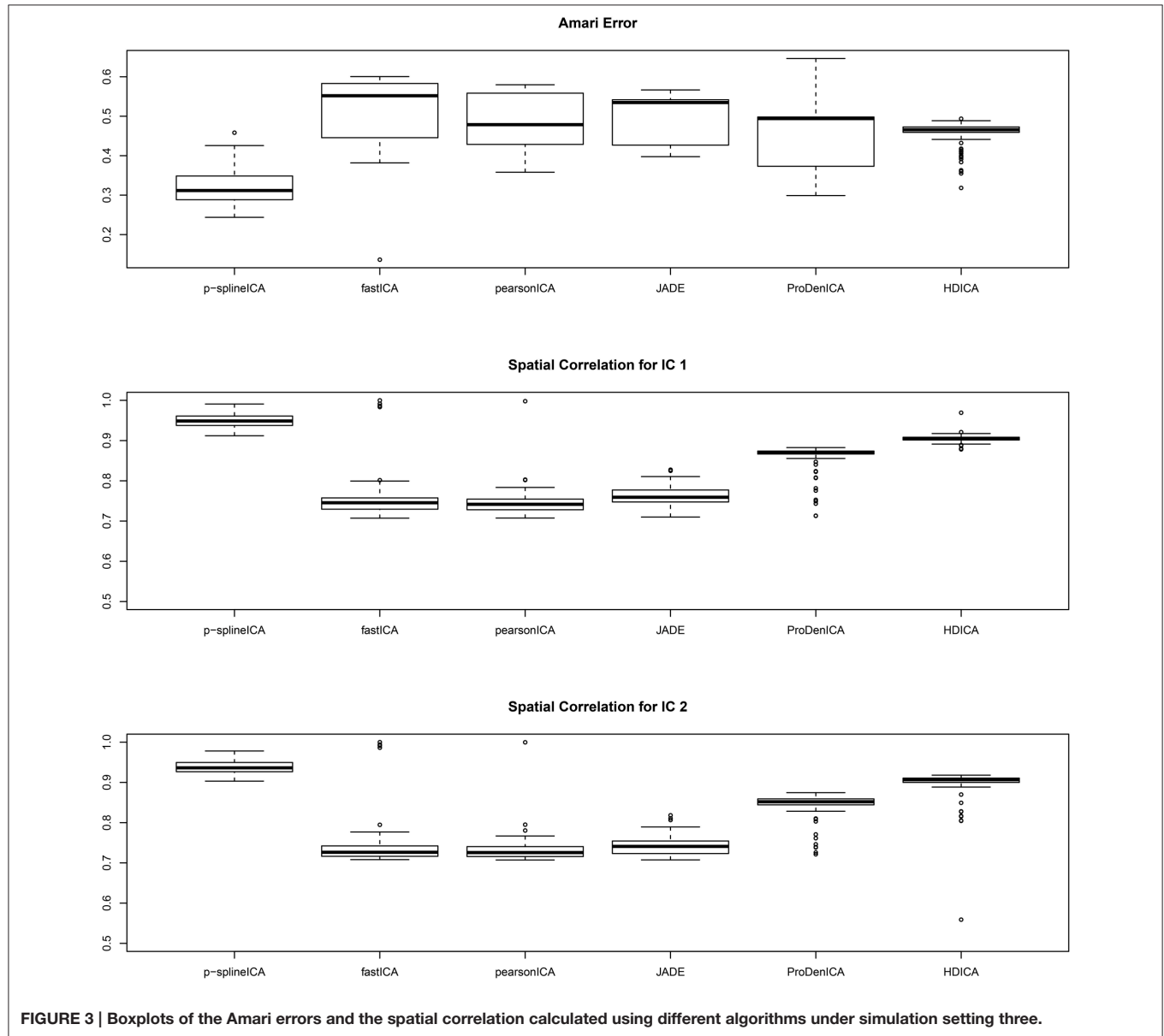
$$A = \begin{pmatrix} 2 & 1 & 2 \\ 3 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix}.$$

Figure 1 summarizes the simulation results based on 200 replications. The spatial correlation is the absolute correlation between the estimated spatial map and the true spatial map without noise. The Amari error is computed to evaluate the accuracy of the estimated mixing matrix (Amari, 1998). It is seen from the boxplots of the spatial correlation and the Amari errors that the proposed ICA algorithm performs equally well as fastICA, JADE, and PearsonICA, and all



these algorithms perform substantially better than ProDenICA algorithm. ProDenICA fails probably due to the extreme values introduced by the noise (See more discussion in Risk et al.,

2014). This shows that the ProDenICA is sensitive to extreme values, while our algorithm is robust to extreme values. The average computation time per replication is 6.21 s using the



p-spline ICA, 3.12 s using ProDenICA, and 308.34 s using HDICA.

In the second setting, we assume the number of source signals $Q = 2$, and we generate the signals based on parcellation. Specifically, we partition the real line into 10 intervals, with cutoffs at the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, and 90th percentiles of the normal distribution. For the first IC, the density function is uniformly distributed within each interval, but the overall shape is approximately normal. For the second IC, the density function follows Laplace distribution within each interval, and the overall shape is approximately normal. The mixing matrix is given by

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}.$$

The boxplots of the spatial correlation and the Amari errors based on 200 replications are summarized in **Figure 2**. Under the second scenario, the underlying signals have region specific densities, and the overall density functions for both components are approximately normal. All the competing algorithms considered in the simulation studies show substantial bias. These algorithms fail to recover the true signals because they heavily depend on the non-gaussianity assumption. On the contrary, the proposed algorithm accounts for the effect of parcellation and recovers the true signals with relatively high accuracy. The proposed algorithm substantially outperforms all the competing algorithms under the second setting. The average computation time per replication is 5.89 s using p-spline ICA, 1.67 s using ProDenICA, and 76.59 s using HDICA.

In the third setting, we generate multi-subject data with number of subject $I = 3$. The source signals are the same as those in the second setting, and the mixing matrices for the three subjects are given by

$$A_1 = \begin{pmatrix} 1 & 0.25 \\ 0.5 & -0.5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 2 \\ 0.5 & -0.5 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 2 \end{pmatrix}.$$

The simulation results are summarized in **Figure 3**, where in each simulation replication, the Amari error is calculated as the average of the Amari errors for all three subjects. The results show that, for multi-subject data, the proposed algorithm successfully recovers both the common spatial signals and the individual mixing matrices. In addition, for similar reasons as in the second setting, the proposed algorithm substantially outperforms all the competing algorithms.

In the fourth setting, we generate the ICs and mixing matrices by mimicking signals from real fMRI data. Specifically, we run fastICA on 10 subjects from the NITRC 1000 Connectome dataset to get twenty ICs (networks). Three of the twenty networks are chosen as the true signals, and they are shown in **Figure 4**. The time courses are also signals from real data, obtained in a similar way as in Calhoun et al. (2009). They are shown in **Figure 5**. We first apply a two-stage dimension reduction using the method as described in Section 2.1. Then we apply the proposed atlas-based ICA algorithm using the brain parcellation proposed by Ye et al. (2011). The correlation matrix

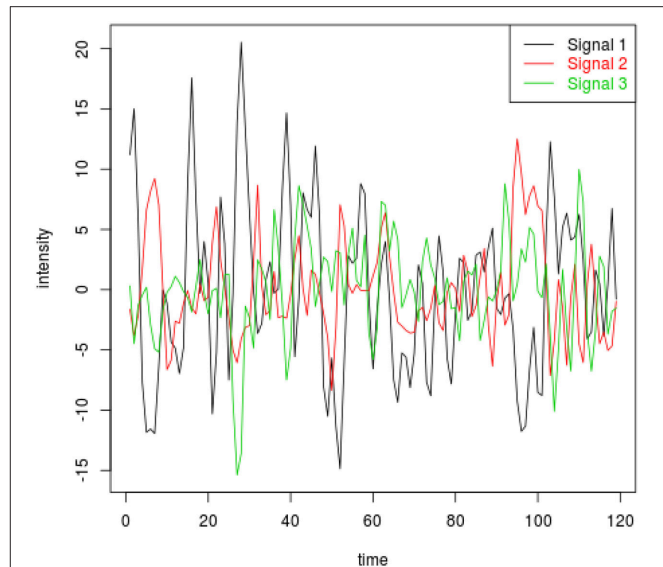


FIGURE 5 | Time courses corresponding to ICs 1, 2, and 3 for the fourth simulation setting.

between the true signals and the estimated signals using the proposed algorithm is

$$\begin{pmatrix} -0.999 & -0.025 & 0.007 \\ 0.028 & 0.999 & -0.006 \\ -0.004 & -0.003 & 0.999 \end{pmatrix}.$$

The results indicate that our proposed p-spline based ICA algorithm is successful in recovering signals from real fMRI data.

4. APPLICATION

We apply our proposed algorithm to the 1000 Functional Connectomes Project dataset, which consists of thousands of resting state scans combined across multiple sites with the goal of facilitating discovery and analysis of brain networks (Biswal et al., 2010). It is one of the largest freely available fMRI datasets. The fMRI scans were collected when the subjects stayed in the scanner for 2.2–20 min in resting state. Scanning parameters used to acquire the data from each site are detailed elsewhere (for complete information see https://www.nitrc.org/projects/fcon_1000/).

As the quality and scanning parameters vary across sites, we focus on data from the largest site, Cambridge, which contains $I = 50$ subjects. For the subjects used in this analysis, the number of time points is $T = 119$. We use the MNI template to remove the background noise and to retain voxels that are in the actual brain. For each subject, we have a $T \times V$ dimensional matrix X_i . The group data matrix X is obtained by concatenating I subjects' data in the temporal domain.

Following Biswal et al. (2010), we assume there are $Q = 20$ independent components in this application. An SVD is

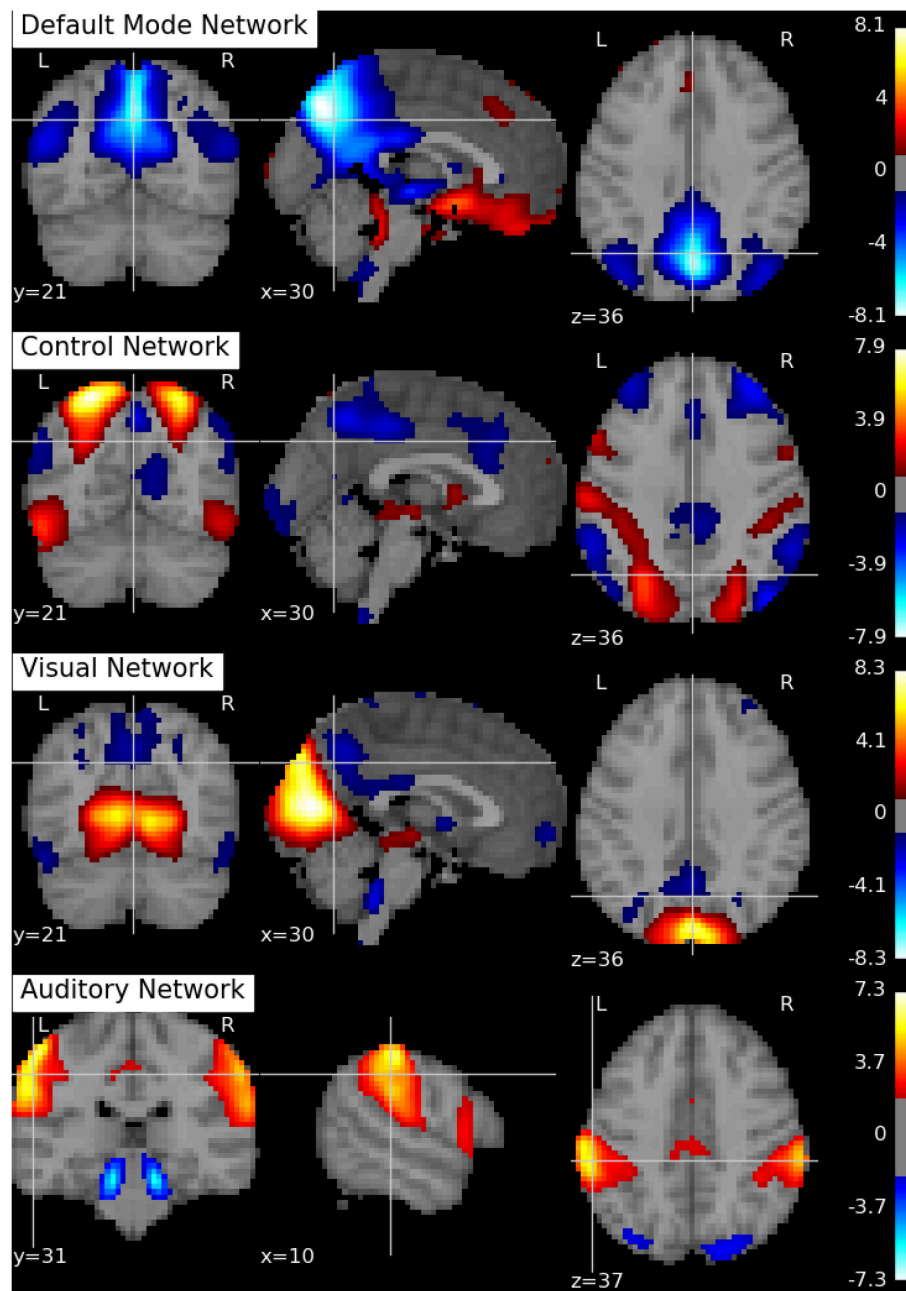


FIGURE 6 | Brain networks obtained by the p-spline ICA algorithm using data from the 1000 Functional Connectome Project.

performed to reduce the dimension of the aggregated data matrix to $Q \times V$. The ICA algorithms are then applied to the reduced data matrix and the Python toolbox Nilearn (Abraham et al., 2014) is used for visualization of the estimation results. Specifically, the estimated ICs using the proposed p-spline based ICA algorithm are shown in **Figure 6**. Several main brain networks including the default mode network (DMN) and the control network are successfully identified by the proposed algorithm. As a comparison, the results from fastICA are shown in **Figure 7**. The ICs estimated by fastICA and the p-spline ICA are matched by

correlation. Of the 20 pairs, the highest correlation is 0.99, the lowest correlation is 0.52, and the median correlation is 0.93. Specifically, the following is a list of these correlations for the major brain networks: visual network (0.99), auditory network (0.98), DMN (0.96), and control network (0.92).

As suggested by an anonymous reviewer, we investigate the impact of the dimension of the reduced space on the final results. Specifically, we select different values of R and Q (the number of eigenvectors in the subject-level and group-level dimension reduction), and rerun the ICA algorithm on the dimension

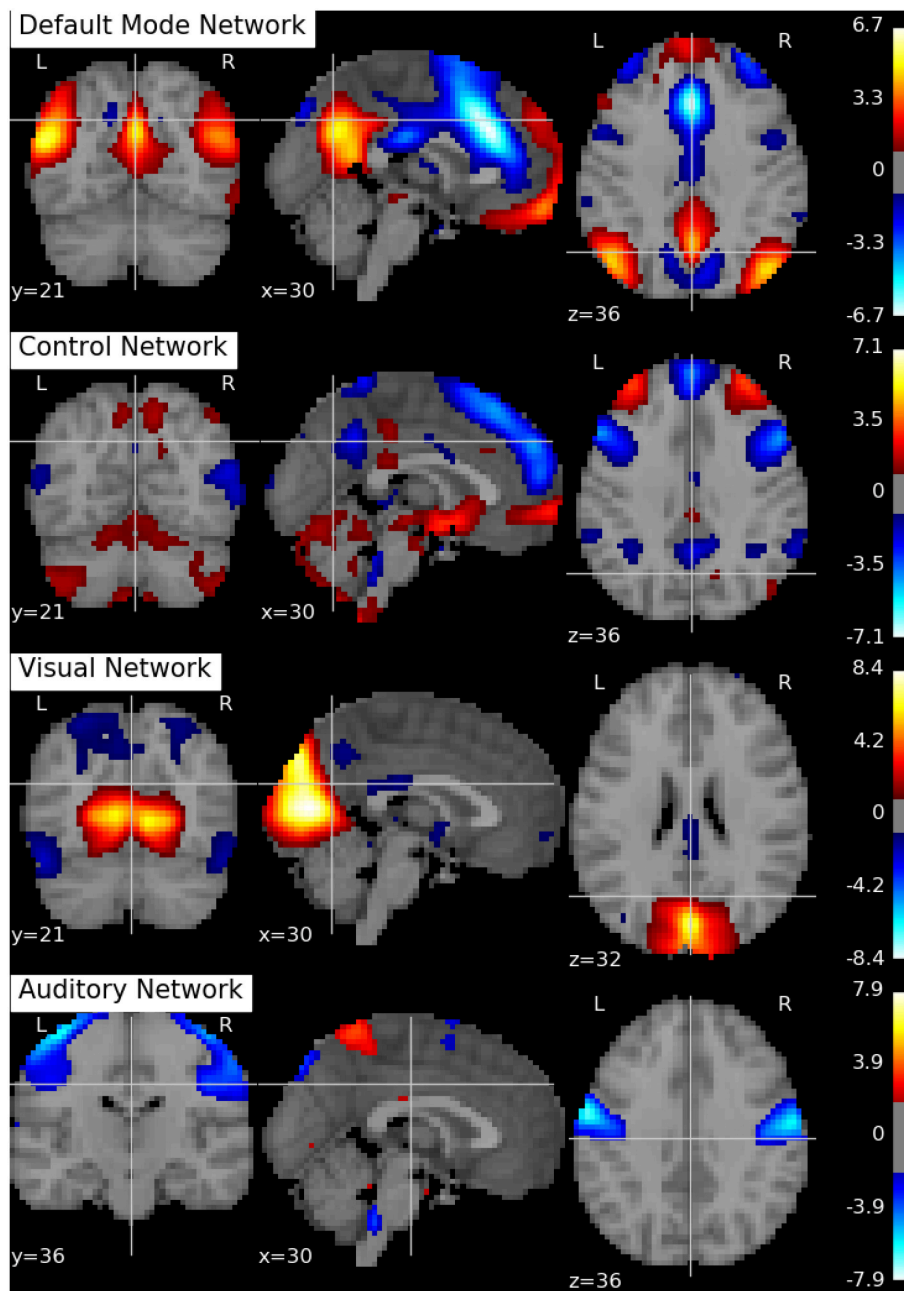


FIGURE 7 | Brain networks obtained by the fastICA algorithm using data from the 1000 Functional Connectome Project.

reduced dataset. We set $R = 15, 20, 30$ and $Q = 15, 20, 30$, respectively. Similarly as in Li et al. (2007), we find that the IC estimates are well separated when $Q = 15, 20$. When $Q = 30$, the estimation of the major networks shows degradation and a few of the other estimated components seem to be noise. Specifically, the correlations for the major brain networks using $R = 20, Q = 15$, and $R = 20, Q = 20$ are as follows: visual network (0.96), auditory network (0.73), DMN (0.86) and control network (0.84). In addition, the correlations for the major brain networks using $R = 20, Q = 30$, and $R = 20, Q = 20$

are as follows: visual network (0.78), auditory network (0.61), DMN (0.88) and control network (0.69). In summary, we find that, although the estimation results depend on the number of components, the major networks appear to be robust against the choices of number of components.

5. DISCUSSION

Independent component analysis is a factor-analytic approach that is commonly used in analyzing fMRI data. In this

manuscript, we present a novel and simple ICA algorithm that is fast, likelihood based and straightforward to program. The algorithm is nonparametric, data-driven, and is blind to the particular distribution of the underlying signals. As a byproduct of the algorithm, we obtain the likelihood function of the ICA model which can be used for further statistical inference. It should be noted that, the likelihood function in our algorithm is a profile likelihood, since we are mainly interested in the mixing matrix estimates and the parameters over the spline basis are nuisance parameters. Indeed, one could also study the coefficients on the spline basis in a full likelihood, but this is not the goal of this manuscript, hence the variance of the estimator of the mixing matrix depends on the variance of the nuisance parameters.

The proposed algorithm is extended to allow for region specific IC density functions, on the rationale that most signals of interest are reasonably confined to a subset of the entire anatomical brain space (Guo and Pagnoni, 2008). When the source signals distribute identically across brain, the estimation accuracy of the parcellation-based estimator is similar to that of the full-brain estimator, because it becomes equivalent to the full-brain estimator. However, when the source signals distribute differently across brain, the full-brain estimator may result in substantial bias while the parcellation-based estimator can successfully recover the source signals. It should be noted that the parcellation based adjustment can be applied to other ICA algorithms as well. Indeed, for any gradient-based ICA, one can do the adjustment by taking a weighted sum over the updates of each of the parcellations, where the weights account for the number of samples in the parcellations. This flexibility ensures the generalizability of the proposed parcellation based adjustment.

Simulation studies show that our proposed algorithm works well in both the simple and complex situations, and it

substantially outperforms the existing ICA algorithms when the identically distributed assumption of the source signals is violated. In applying the proposed algorithm to the fMRI data, we choose to account for the difference in brain activities across regions by using the brain parcellation proposed by Yeo et al. (2011). Our data application results show that the proposed algorithm successfully identifies the main brain networks in the 1000 Functional Connectomes Project dataset.

There are a few directions for future research. Firstly, the test-retest reliability of the intrinsic brain networks is an important issue and has been studied extensively in recent years. For example, Zuo et al. (2010) found that a few functionally relevant components (such as the default mode, auditory-motor and executive control) show the highest reliability across all components. It would be interesting to compare different ICA algorithms in identifying and characterizing those functionally relevant components. Secondly, there are a variety of existing brain parcellation schemes, including those proposed by Tzourio-Mazoyer et al. (2002); Fischl et al. (2004); Beckmann et al. (2009); Yeo et al. (2011). It would be interesting to study the optimal choice of parcellation under different scientific scenarios. Thirdly, as pointed out by an anonymous reviewer, pre-whitening, although a standard pre-processing procedure, may result in loss of information and bias in estimation (Cardoso, 1994). It would be interesting to investigate alternative pre-processing procedures to avoid the bias introduced by pre-whitening.

ACKNOWLEDGMENT

This work was partly supported by a grant from the Simons Foundation (#354180, Shanshan Li).

REFERENCES

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., et al. (2014). Machine learning for neuroimaging with scikit-learn. *Front. Neuroinform.* 8:14. doi: 10.3389/fninf.2014.00014
- Allen, G., Barnard, H., McColl, R., Hester, A. L., Fields, J. A., Weiner, M. F., et al. (2007). Reduced hippocampal functional connectivity in Alzheimer disease. *Arch. Neurol.* 64, 1482–1487. doi: 10.1001/archneur.64.10.1482
- Amari, S.-I. (1998). Natural gradient works efficiently in learning. *Neural Comput.* 10, 251–276. doi: 10.1162/089976698300017746
- Amari, S.-I., and Cardoso, J.-F. (1997). Blind source separation-semiparametric statistical approach. *IEEE Trans. Signal Process.* 45, 2692–2700. doi: 10.1109/78.650095
- Bach, F. R., and Jordan, M. I. (2003). Kernel independent component analysis. *J. Mach. Learn. Res.* 3, 1–48. doi: 10.1109/icassp.2003.1202783
- Beckmann, C. F., and Smith, S. M. (2005). Tensorial extensions of independent component analysis for multisubject fMRI analysis. *Neuroimage* 25, 294–311. doi: 10.1016/j.neuroimage.2004.10.043
- Beckmann, M., Johansen-Berg, H., and Rushworth, M. F. (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J. Neurosci.* 29, 1175–1190. doi: 10.1523/JNEUROSCI.3328-08.2009
- Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734–4739. doi: 10.1073/pnas.0911855107
- Boscolo, R., Pan, H., and Roychowdhury, V. P. (2004). Independent component analysis based on nonparametric density estimation. *IEEE Trans. Neural Netw.* 15, 55–65. doi: 10.1109/TNN.2003.820667
- Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., et al. (2005). Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *J. Neurosci.* 25, 7709–7717. doi: 10.1523/JNEUROSCI.2177-05.2005
- Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001a). A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* 14, 140–151. doi: 10.1002/hbm.1048
- Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001b). Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. *Hum. Brain Mapp.* 13, 43–53. doi: 10.1002/hbm.1024
- Calhoun, V. D., Liu, J., and Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage* 45, 163–172. doi: 10.1016/j.neuroimage.2008.10.057
- Cardoso, J.-F. (1994). "On the performance of orthogonal source separation algorithms," in *Proceedings of EUSIPCO*, Vol. 94 (Edinburgh), 776–779.
- Cardoso, J.-F. (1998). Blind signal separation: statistical principles. *Proc. IEEE* 86, 2009–2025. doi: 10.1109/5.720250
- Cardoso, J.-F. (1999). High-order contrasts for independent component analysis. *Neural Comput.* 11, 157–192. doi: 10.1162/089976699300016863

- Comon, P. (1994). Independent component analysis, a new concept? *Signal Process.* 36, 287–314. doi: 10.1016/0165-1684(94)90029-9
- Comon, P., and Jutten, C. (2010). *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Cambridge, MA: Academic Press.
- Crainiceanu, C. M., Caffo, B. S., Luo, S., Zipunnikov, V. M., and Punjabi, N. M. (2011). Population value decomposition, a framework for the analysis of image populations. *J. Am. Stat. Assoc.* 106, 775–790. doi: 10.1198/jasa.2011.ap10089
- Eilers, P., and Marx, B. (1996). Flexible smoothing with B-splines and penalties (with comments and rejoinder). *Stat. Sci.* 11, 89–121. doi: 10.1214/ss/1038425655
- Eloyan, A., Crainiceanu, C. M., and Caffo, B. S. (2013). Likelihood-based population independent component analysis. *Biostatistics* 14, 514–527. doi: 10.1093/biostatistics/kxs055
- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., et al. (2004). Automatically parcellating the human cerebral cortex. *Cereb. Cortex* 14, 11–22. doi: 10.1093/cercor/bhg087
- Guo, Y., and Pagnoni, G. (2008). A unified framework for group independent component analysis for multi-subject fMRI data. *NeuroImage* 42, 1078–1093. doi: 10.1016/j.neuroimage.2008.05.008
- Hastie, T., and Tibshirani, R. (2002). *Independent Component Analysis Through Product Density Estimation*. Technical Report, Department of Statistics, Stanford University.
- Hastie, T., and Tibshirani, R. (2010). *ProDenICA: Product Density Estimation for ICA using tilted Gaussian Density Estimates*. R package version 1.0.
- Hyvärinen, A., Karhunen, J., and Oja, E. (2001). *Independent Component Analysis*. Hoboken, NJ: John Wiley and Sons. doi: 10.1002/0471221317
- Hyvärinen, A., and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Netw.* 13, 411–430. doi: 10.1016/S0893-6080(00)00026-5
- Karvanen, J. (2006). *Package PearsonICA*. R package version 1.2-4.
- Karvanen, J., Eriksson, J., and Koivunen, V. (2000). “Pearson system based method for blind separation,” in *Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation (ICA2000)* (Helsinki), 585–590.
- Li, X.-L., and Adali, T. (2010). Independent component analysis by entropy bound minimization. *IEEE Trans. Signal Process.* 58, 5151–5164. doi: 10.1109/TSP.2010.2055859
- Li, Y.-O., Adali, T., and Calhoun, V. D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. *Human Brain Mapp.* 28, 1251–1266. doi: 10.1002/hbm.20359
- Marchini, J. L., Heaton, C., and Ripley, B. D. (2013). *fastICA: FastICA Algorithms to Perform ICA and Projection Pursuit*. R package version 1.2-0.
- McKeown, M. J., Jung, T. P., Makeig, S., Brown, G., Kindermann, S. S., Lee, T. W., et al. (1998a). Spatially independent activity patterns in functional MRI data during the stroop color-naming task. *Proc. Natl. Acad. Sci. U.S.A.* 95, 803–810. doi: 10.1073/pnas.95.3.803
- McKeown, M. J., Makeig, S., Brown, G. G., Jung, T. P., Kindermann, S. S., Bell, A. J., et al. (1998b). Analysis of fMRI data by blind separation into independent spatial components. *Hum. Brain Mapp.* 6, 160–188.
- McKeown, M. J., and Sejnowski, T. J. (1998). Independent component analysis of fMRI data: examining the assumptions. *Hum. Brain Map.* 6, 368–372.
- Nordhausen, K., Cardoso, J.-F., Miettinen, J., Oja, H., Ollila, E., and Taskinen, S. (2014). *JADE: JADE and other BSS Methods as Well as Some BSS Performance Criteria*. R package version 1.9-92.
- Risk, B. B., Matteson, D. S., Ruppert, D., Eloyan, A., and Caffo, B. S. (2014). An evaluation of independent component analyses with an application to resting-state fMRI. *Biometrics* 70, 224–236. doi: 10.1111/biom.12111
- Ruppert, D., Wang, M., and Carroll, R. (2003). *Semiparametric Regression*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge, MA: Cambridge University Press. doi: 10.1017/cbo9780511755453
- Samworth, R. J., and Yuan, M. (2012). Independent component analysis via nonparametric maximum likelihood estimation. *Ann. Stat.* 40, 2973–3002. doi: 10.1214/12-AOS1060
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Wang, L., Zang, Y., He, Y., Liang, M., Zhang, X., Tian, L., et al. (2006). Changes in hippocampal connectivity in the early stages of Alzheimer's disease: evidence from resting state fMRI. *Neuroimage* 31, 496–504. doi: 10.1016/j.neuroimage.2005.12.033
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi: 10.1152/jn.00338.2011
- Zuo, X.-N., Kelly, C., Adelstein, J. S., Klein, D. F., Castellanos, F. X., and Milham, M. P. (2010). Reliable intrinsic connectivity networks: test-retest evaluation using ICA and dual regression approach. *Neuroimage* 49, 2163–2177. doi: 10.1016/j.neuroimage.2009.10.080

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Li, Chen, Yue and Caffo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Identification of Voxels Confounded by Venous Signals Using Resting-State fMRI Functional Connectivity Graph Community Identification

Klaudius Kalcher^{1,2*†}, Roland N. Boubela^{1,2†}, Wolfgang Huf^{1,2}, Christian Našel³ and Ewald Moser^{1,2,4}

¹ Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria, ² MR Centre of Excellence, Medical University of Vienna, Vienna, Austria, ³ Department of Radiology, Tulln Hospital, Karl Landsteiner University of Health Sciences, Tulln, Austria, ⁴ Brain Behaviour Laboratory, Department of Psychiatry, University of Pennsylvania Medical Center, Philadelphia, PA, USA

OPEN ACCESS

Edited by:

Brian Caffo,
Johns Hopkins University, USA

Reviewed by:

Hideao Fukuyama,
Kyoto University, Japan
Xin Di,
New Jersey Institute of Technology,
USA

*Correspondence:

Klaudius Kalcher
klaudius.kalcher@meduniwien.ac.at

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 15 August 2015

Accepted: 25 November 2015

Published: 16 December 2015

Citation:

Kalcher K, Boubela RN, Huf W,
Našel C and Moser E (2015)
Identification of Voxels Confounded by
Venous Signals Using Resting-State
fMRI Functional Connectivity Graph
Community Identification.
Front. Neurosci. 9:472.
doi: 10.3389/fnins.2015.00472

Identifying venous voxels in fMRI datasets is important to increase the specificity of fMRI analyses to microvasculature in the vicinity of the neural processes triggering the BOLD response. This is, however, difficult to achieve in particular in typical studies where magnitude images of BOLD EPI are the only data available. In this study, voxelwise functional connectivity graphs were computed on minimally preprocessed low TR (333 ms) multiband resting-state fMRI data, using both high positive and negative correlations to define edges between nodes (voxels). A high correlation threshold for binarization ensures that most edges in the resulting sparse graph reflect the high coherence of signals in medium to large veins. Graph clustering based on the optimization of modularity was then employed to identify clusters of coherent voxels in this graph, and all clusters of 50 or more voxels were then interpreted as corresponding to medium to large veins. Indeed, a comparison with SWI reveals that $75.6 \pm 5.9\%$ of voxels within these large clusters overlap with veins visible in the SWI image or lie outside the brain parenchyma. Some of the remaining differences between the two modalities can be explained by imperfect alignment or geometric distortions between the two images. Overall, the graph clustering based method for identifying venous voxels has a high specificity as well as the additional advantages of being computed in the same voxel grid as the fMRI dataset itself and not needing any additional data beyond what is usually acquired (and exported) in standard fMRI experiments.

Keywords: fMRI, BOLD, graph analysis, graph clustering, physiological signals, brain, veins

1. INTRODUCTION

Any interpretation of fMRI results as indirect measures of neuronal activation rests on the assumption that magnetization changes caused by changes in blood oxygenation are due to brain activity in the immediate vicinity. Whether and to what extent this assumption holds, however, has been the matter of much debate from the first days of fMRI onwards. While the discussion of

this “brain or vein” question has generated a wealth of research on identifying veins and signals originating from them, the majority of fMRI studies still ignore the issue and take no measures to assess or reduce the influence of signals from major vessels (Menon, 2012).

Quantification of the influence of draining veins on fMRI results has provided evidence that it is most pronounced at low magnetic field strengths, and the relative influence of microvasculature to the MR signal increases as the field strength increases (Duong et al., 2003): while at 1.5 T the signal originates virtually entirely in the macrovasculature (Lai et al., 1993), the ratio shifts in favor of the microvasculature as a major signal source at 3T, 4T, and 7T. Still, even at higher field strengths, protocols using gradient echo EPI are highly sensitive to signal changes originating from larger veins, to the point that no significant improvement can be observed between 3T and 7T results (Geißler et al., 2013).

Several efforts have been made to reduce the influence of venous signals on fMRI measurements, and a number of effective ways for increasing the specificity of signals for the microvasculature have emerged from them. The use of spin-echo instead of gradient echo sequences can drastically reduce extravascular signal contributions (Duong et al., 2003), but only at a steep cost in terms of signal-to-noise and contrast-to-noise ratio (Norris, 2012). Specific corrections to eliminate venous signals based on phase images have also been developed (Menon, 2002; Rowe and Logan, 2005; Curtis et al., 2014), but are rarely included in the sequences used by typical fMRI studies, and yield the risk of introducing errors through over-correction (Nencka and Rowe, 2007).

The development of these methods for identifying veins is to some extent related to that of methods for eliminating physiological influences, leading to converging developments. One of the most common approaches to address physiological signal contamination is RETROICOR (Glover et al., 2000), which uses externally measured respiration and cardiac signals for a regression-based correction of fMRI data. A later method termed CompCor (Behzadi et al., 2007) eliminates the need for externally measured signals by identifying potential regressors from either ventricular and white matter signals, or from signals in voxels with higher-than-average time course standard deviation—a feature typically seen in voxels containing larger veins. The most recent development, Highcor, merges this line of research with the work on phase-based venous suppression mentioned above by using the correlation between phase and magnitude image time series to identify venous voxels that can be used to extract regressors for physiological noise reduction (Curtis and Menon, 2014).

The reasons for identifying venous voxels are more complex than the elimination of global physiological noise, however, as any BOLD effects necessarily carry the potential for signal changes further downstream along draining veins. It is thus desirable to not only reduce global physiological noise as in these regression approaches, but also to reliably identify voxels with potential venous signal contributions to help interpreting signal changes seen in and around them. The most direct way of localizing macrovascular effects might be the

creation of venous maps using susceptibility weighted imaging (SWI; Reichenbach et al., 2000; Haacke et al., 2004). Still, extravascular signal influences in fMRI measurements might extend beyond the delineation of veins in the SWI image, and imperfect coregistration of SWI to fMRI images can further limit the precision of this method of localization. An immediate measure proposed to minimize macrovascular influences is the elimination of voxels with time series coefficients of variation much larger than the local average of their surrounding voxels [a concept related to the second of the two approaches used by Behzadi et al. (2007) in CompCor, see above], which empirically corresponds to regions next to large vessels, as done in the minimal preprocessing pipeline (Glasser et al., 2013) of the Human Connectome Project (Van Essen et al., 2013). This method has the advantages of requiring neither separate protocols nor specialized measurement techniques, and estimates affected voxels directly from standard fMRI time series; however, there exists no evaluation of the correspondence of the time series standard deviation with venous effects.

The question of which voxels are influenced by large vessels is not a purely theoretical exercise. While in the beginnings of fMRI, the brain-vs.-vein debate was settled by a general acknowledgment that with the imaging resolutions then available, the practical relevance of knowing whether a signal originates in the microvasculature in the cortex or in the draining vein at its surface was limited. There are multiple reasons why this argument should no longer be used as an excuse for avoiding the question. First, many veins run in sulci and might cause a signal change whose causal origin is on the gyrus on one side of the sulcus be misattributed to the gyrus on the other side. While earlier imaging techniques might not have allowed to make such distinctions regardless of whether the measured signal originated in the parenchyma or in a draining vein, many current analysis techniques like surface projections for surface-based analyses rely on correct attribution of signals at this level. The second, somewhat related, reason is that improvements in spatial resolution at higher field strengths and with the use of more sophisticated acceleration techniques (Simultaneous Image Refocused EPI, Multiband EPI) have led to the possibility of imaging at sub-millimeter resolutions (Feinberg et al., 2010), but this improvement in nominal spatial image resolution can only lead to interpretable gains if the physiological basis for the measured effect matches this granularity. Indeed, Turner (2002) suggested that due to dilution effects, draining vein effects might not be seen at more than a few millimeters distance from the gray matter region drained where the effect originated. But while a spatial gap between neuronal origin and the immediate source of the measured BOLD signal of 4 mm might be considered of limited relevance when imaging at a spatial resolution of 3 mm, the existence of unavoidable spatial discrepancies of this magnitude would render advances into higher resolutions entirely pointless. Finally, it is possible to show that at least in some cases, draining vein effects might occur at much larger distances from their neuronal origins, as is the case for signal changes in the basal vein of Rosenthal (BVR) next to the amygdala (Boubela et al., 2015).

In the analysis of the BVR signals, it also became apparent that one distinguishing feature of venous voxels was their resting-state functional connectivity pattern (Boubela et al., 2015), exhibiting very strong positive and negative correlations to other voxels in the macrovasculature. Thus, in an approach similar to functional parcellation methods of the brain (Eickhoff et al., 2015) this resting-state connectivity structure between voxels could be used to distinguish voxels in the macrovasculature from others. One such approach consists in using graph learning tools on the connectivity graph (where vertices correspond to voxels or sets of voxels, and edges link vertices with correlated time courses together). Previously, connectivity analyses have rarely been performed on a voxel-wise level, among others for computational reasons: if, for example, 150,000 voxels lie within the brain mask, the complete voxel-by-voxel correlation matrix would consist of $2.25 \cdot 10^{10}$ entries, taking up about 167 GB (in practice, the number of voxels is typically reduced by restricting analysis to gray matter voxels or by resampling the data to a coarser resolution). Not all of these entries actually need to be stored to perform analyses, in particular when analyzing a relatively sparse connectivity graph, and efficient tools for tackling similar problems on large datasets have been developed in other fields.

In this work, graph based cluster analysis was performed to show how these tools can be applied to solve a practical problem of fMRI data analysis. Voxel-by-voxel correlations are computed for all in-brain voxels to create a voxelwise connectivity graph. Resampling to a coarser resolution as well as limiting analysis to a subset of voxels (e.g., gray matter voxels) are avoided as they would hamper the specific research question: resampling would lead to a loss of specificity in that affected voxels would be averaged with adjacent voxels to form the larger voxels of the coarser grid, and a gray matter mask might exclude parts of the venous structure, effectively hindering the identification of venous voxels based on their connectivity to voxels within other veins. Based on previous observations (Boubela et al., 2015), the largest clusters with the strongest (positive and negative) correlations among their voxel's time-series emerging from a clustering of this graph could be expected to reflect medium to large veins.

2. MATERIALS AND METHODS

2.1. Subjects

Fifteen healthy subjects (8 females/7 males, mean age 35.3, SD 13.3) were recruited at Medical University of Vienna. Exclusion criteria were prior psychiatric or neurologic illnesses, as well as the usual exclusion criteria for MR studies. All subjects gave written informed consent prior to the scan and the study was approved by the local institutional review board (Ethikkommission der Medizinischen Universität Wien).

2.2. Data Acquisition Protocols

All MRI scans were performed on a 3 Tesla TIM Trio using the standard 32-channel head coil and whole-body gradients (Siemens Medical Solutions, Erlangen, Germany). First, a high-resolution anatomical image was acquired using MPRAGE with $1 \times 1 \times 1.1 \text{ mm}^3$ resolution, and 160 sagittal slices (TE = 4.21 ms,

TR = 2300 ms, flip angle 90° , inversion time 900 ms). Second, BOLD fluctuations at rest were measured with a short-TR multi-band EPI-sequence (Feinberg et al., 2010) using $1.7 \times 1.7 \times 2 \text{ mm}^3$ resolution, 2 mm slice gap (matrix size 128×128 , 32 axial slices, TE = 31 ms, TR = 333 ms, flip angle 30° , multiband factor 8, bandwidth = 1776 Hz/Pixel) collecting 1200 volumes. Finally, susceptibility weighted images (SWI) were acquired at $0.6 \times 0.6 \times 2.0 \text{ mm}$ resolution (matrix size 384×384 , 52 slices per slab, 1 slab, TE = 29 ms, TR = 42 ms, flip angle 15°) to visualize medium to large venous vessels.

2.3. Preprocessing

To keep closely to the original images, only minimal preprocessing was applied to functional data, including only skull stripping using FSL BET, motion removal using FSL MCFLIRT, and band-pass filtering. For the latter, the pass-band used was 0.01–0.2 Hz, to avoid as far as possible influence from high-frequency respiratory or cardiac fluctuations. SWI images were segmented using FSL FAST for vein delineation, coregistered to the EPI weighted images, with the vein masks generated from segmentation also being transformed into EPI space using the resulting transformation parameters, using trilinear interpolation to ensure that all voxels in EPI space with some overlap with veins from the SWI mask have non-zero values. This vein map in EPI space was then binarized to generate a vein mask for the EPI images.

2.4. Graph Generation

Pairwise Pearson correlation coefficients were computed between all voxels within the brain mask, using GPUs for the calculation of the correlation coefficients (Boubela et al., under revision) and splitting the dataset into tiles to allow for the computations to fit within GPU memory (6 GB). For each subject, this correlation matrix was thresholded to generate the adjacency matrix of a graph using a correlation threshold such that $S < 4$, with

$$S = \frac{\log E}{\log K}$$

(where E is the number of edges and K the average node degree). This thresholding criterion results in a rather sparse graph of only the strongest correlations, which are more likely to reflect adjacent voxels along the same vein or otherwise highly congruent voxel signals (as opposed to the more subtle long-distance connections of brain networks of neuronal origin; see below in the discussion for more details on the effect of the sparsity criterion). For each subject, the largest correlation threshold fulfilling $S < 4$ was computed iteratively by decreasing the threshold in steps of 0.01, starting at 1. The threshold was applied to the absolute value of the correlation coefficients to take into account both positive and negative correlations exceeding a certain correlation strength. The resulting connectivity graph had all in-brain voxels as vertices, and each correlation between two voxels that was above the threshold resulted in edges between the two corresponding vertices, with the correlation coefficient used as edge weight. Graphs were represented using the package *igraph* (version 1.0.1) in R (version 3.1.1). Self-loops and

multiple edges were eliminated using the `igraph` function `simplify`.

2.5. Graph Cluster Identification

Community identification on the graph was performed using the method based on modularity optimization by Newman (2006) as implemented in the `igraph` function `cluster_fast_greedy`. The optimization of graph modularity means that the resulting clusters are defined by their voxels having maximum connectedness among each other and minimum connectedness to voxels outside their own cluster. All voxels from all clusters that individually contained 50 or more voxels were then pooled into a single mask, which thus contained all voxels with time-courses strongly correlated (either positively or negatively) with those of a large number of different voxels. It should be noted that Newman's method is intended to detect communities in connected networks and that its application on sparse networks as used here might result in some cases in entire connected components being categorized as clusters. Nonetheless, for our purposes, this is still sufficient to detect groups of voxels with highest relative connectedness to each other considering the general sparsity of the graph.

2.6. Validation

To show that these voxels correspond mostly to vasculature, the overlap with the vein mask from SWI was computed (in EPI space); the proportion of voxels from the graph clustering map that overlaps with the SWI vein mask can be seen as a measure for the specificity of the graph clustering method, though it should be kept in mind that it is not the true specificity because the segmented SWI is not the ground truth for the identification of venous voxels: coregistration imperfections can lead to spatial deviations in the localization of these voxels, and not all low signal intensities in SWI originate from veins since other factors like iron levels (higher in the basal ganglia than in the rest of the brain) or proximity to air cavities or bone affect susceptibility. The latter observation also implies that it is impossible to make any meaningful quantification of the sensitivity of the graph clustering method by using SWI, as it means that an accurate map of venous voxels should not indiscriminately include all voxels with low signal intensity in SWI.

3. RESULTS

Overall, of the voxels within the brain masks (between 142,800 and 172,300 for the different subjects, mean 157,600, SD 10,540), $17,730 \pm 5069$ voxels (or $11.2 \pm 3.0\%$) were identified by the graph clustering algorithm as being part of large highly coherent networks (see **Table 1**).

Single-subject images of the graph clustering masks overlaid over SWI are shown in **Figure 1**. The spatial distribution of the voxels identified by the method seems to exhibit a consistent pattern. Most of the voxels within the brain follow the path of veins visible in the SWI underlay. Another set of voxels delineates areas of low signal quality in orbitofrontal regions subject to susceptibility artifacts or at the edge of the brain, in either case such voxels could be discarded for fMRI analyses interested in neuronal effects.

This observation can also be quantified by comparing the mask gained from graph clustering with a binarized mask gained from segmenting the SWI image (see **Figure 2**), and the average proportion of voxels of the mask identified via graph clustering overlapping with veins in SWI is 0.67 ± 0.05 . A further significant proportion of voxels not directly overlapping veins lies on the edge of the brain mask, as identified by eroding the brain mask with a $5 \times 5 \times 5$ -voxel kernel (see **Figure 4**) using the R package `mmand` (Clayden, 2014), outside of what can be recognized as the brain itself (BET seems to be rather conservative in skullstripping), possibly reflecting signals from superficial veins, raising the overlap proportion to 0.77 ± 0.06 (see also **Table 1**).

In some brain areas (e.g., in the medial prefrontal region in **Figure 1**), the locations of the vein recognizable in SWI on one hand and the voxels of the graph clustering mask on the other can be observed not to overlap perfectly. Still, the similarity of the shape between the two features, only shifted by 1–2 voxels, strongly suggests that they are caused by the same underlying structure (i.e., the same vein). Such discrepancies are not necessarily worrying. Since the graph clustering mask is generated from the EPI voxel timecourses themselves as opposed to the SWI images acquired in a separate measurement, they can be seen as yielding potentially valuable complementary information on the effect of a vein on the EPI measurement.

Comparing the time series standard deviations of voxels within and outside the clustering brain mask reveals that voxels

TABLE 1 | Quantitative overview of the voxels identified by graph clustering, and comparison to segmentation of SWI.

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
Correlation threshold	0.63	0.72	0.75	0.79	0.88	0.95
Voxels in brain mask	142,800	148,900	154,400	157,600	167,100	172,300
Voxels in clustering mask	10,470	13,600	18,400	17,730	20,940	26,760
Idem, in % of brain mask	6.5%	9.4%	11.1%	11.3%	12.4%	17.9%
Overlap with SWI veins	58.8%	63.7%	68.4%	67.1%	70.8%	72.5 %
Overlap with SWI veins or brain edge	64.5%	72.9%	78.3%	76.6%	79.7 %	87.6%

The minimum, first quartile, median, mean, third quartile, and maximum across all subjects is given for the correlation threshold obtained for the binarization of the correlation matrix to a graph adjacency matrix, the absolute numbers of voxels within the brain mask and within the graph clustering mask, the percentage of voxels of the brain mask that were within the clustering mask, as well as for the percentage of voxels of the graph clustering map that lie within the mask generated from SWI segmentation (which can be interpreted as a measure of specificity), without and with the additional inclusion of brain edge voxels.

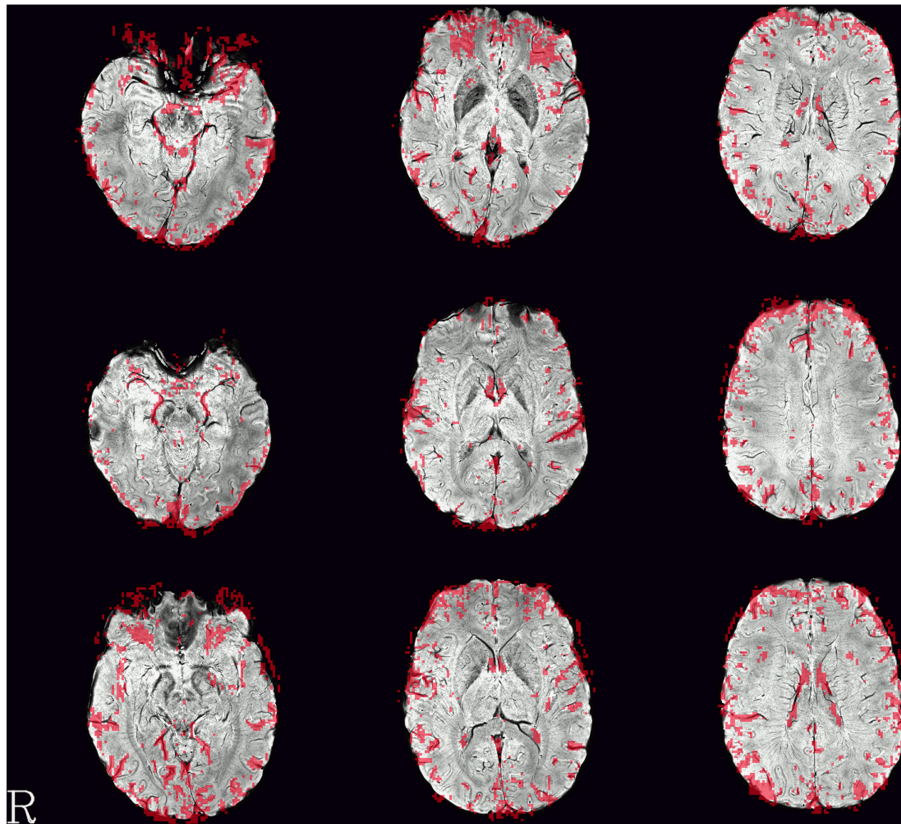


FIGURE 1 | Example single-subject graph clustering masks overlaid on the respective subject's SWI. Red areas indicate voxels within the venous voxel mask gained from graph clustering, each row showing data from a different subject. Structures seen in the graph clustering maps tend to reflect veins seen in SWI. Note that the delineation of those veins in SWI and the voxels in the EPI influenced by the signal as identified from the BOLD EPI itself can deviate slightly. Slices were selected to contain larger veins running along the slice orientation in the SWI.

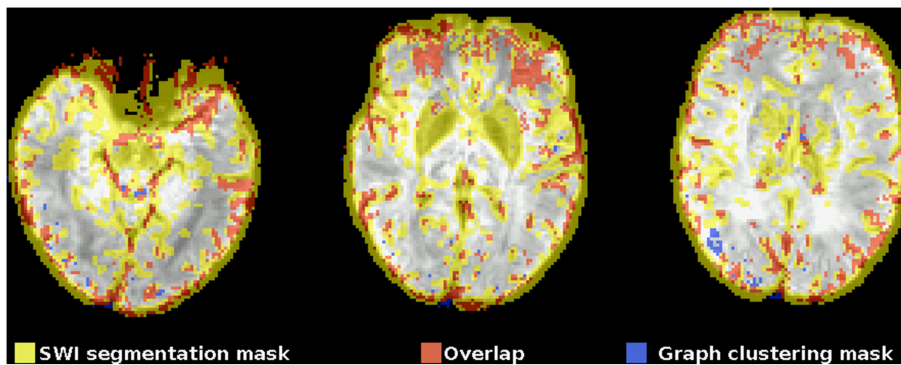


FIGURE 2 | Example comparison between the segmented SWI and graph clustering mask of one subject (the subject shown in the top row in Figure 1). The underlay is the SWI, coregistered, and resampled to the EPI space. Voxels within the mask gained from the segmented SWI are yellow, voxels from the graph clustering mask that are within the SWI mask are red, voxels from the graph clustering mask that do not overlap with the SWI mask are blue. Note how the SWI segmentation mask tends to be rather unspecific to veins in regions with susceptibility-related low signal intensities of different origin, most notably in the basal ganglia due to their high iron levels.

within the mask indeed have on average significantly higher standard deviations ($p < 2 \cdot 10^{-16}$ in all individual subjects), in congruence with the underlying assumption of the method used by Glasser et al. (2013). However, the overlap between the

signal standard deviations in macro- and microvasculature is very pronounced in all subjects, suggesting that a threshold based on the signal standard deviation alone might not be sufficient to discriminate between the two types of voxels (see Figure 3).

4. DISCUSSION

The results presented here show that graph based brain network analysis on a voxelwise basis can yield important insights on the origins of the underlying signals. In particular, network clustering yields a set of voxels defined by strong connections among each other and weak connectivity to voxels outside of this set that contains mostly voxels in medium to large veins (as identified by SWI) as well as extra-cerebral voxels. The very structure of these clusters, voxels spread across the whole brain with a large number of connections with correlation strengths above 0.7–0.95 (depending on the subject), strongly suggests that these voxels contain little information related to local neuronal activity, and rather reflect blood oxygenation changes on a larger (physiologic) scale. These voxels thus violate the underlying assumption of most fMRI analyses that BOLD signal changes in a voxel can be interpreted as an indirect measure of local neuronal activations, and should thus be excluded in this type of analysis.

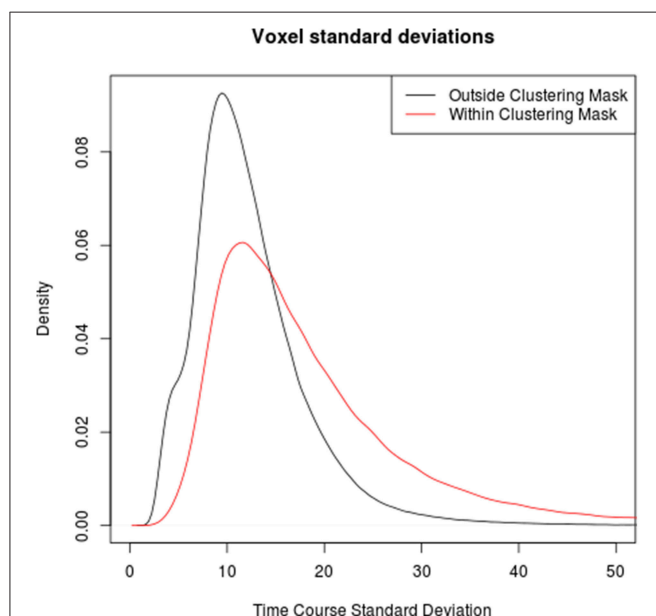


FIGURE 3 | Comparison between the distributions of time series standard deviation in voxels outside (black) and within (red) the clustering mask over all subjects. Note that while time series standard deviations in voxels identified as veins tends to be higher than in those outside veins, the boundary is not clear-cut, highlighting the limits of venous voxel identification based on time series standard deviations.

It is worth noting that some correlation patterns typically arising in fMRI datasets are not reflected in the results computed here and do not seem to confound the correlation-based vein identification. The first of those is that in task fMRI, large vessels often correlate with task signals in active brain regions as they are draining the blood from those regions. One might thus expect that conversely, there should be a correlation between the venous signals investigated here and the signals from voxels within the parenchyma that are drained by these veins, making it difficult to distinguish venous voxels from parenchyma voxels. This does not seem to occur here, as indicated by the overlap of most voxels in the clustering mask with venous voxels identified by SWI. One reason for this might be that the correlation strength between the signals within the vein and each of the individual regions drained by the vein is substantially below the relatively high correlation thresholds that emerge from the high S threshold used in network creation. For medium to large veins, each gray matter voxel ultimately drained by them contributes only a small part to the signal in voxels within the vein, thus leading to lower correlation strengths between parenchyma voxels and veins of this scale in accordance with theoretical models of downstream dilution of effects in veins. This might explain why the clustering mask includes only larger vessels, and fails to identify some of the smaller vessels appearing in the SWI image. The reason for the absence of the correlations between signals from large veins and large areas of activation that typically occur in task fMRI might be that in task fMRI, there is an artificially high coherence of a particular (set of) brain region(s) with the signal in the vein due to the task-induced structure in these activations. In resting-state data as used in this study, however, the patterns along which all regions draining into a particular vein contribute to its signal are less coherent among each other, with different regions potentially contributing differently to the venous signal, and thus having lower individual correlations with it.

The second type of correlation pattern that might be expected to be visible in a functional connectivity graph are resting-state networks previously described in the literature, such as the default-mode network (DMN) or the left and right fronto-parietal networks. The reason for them not appearing in the cluster mask is that in unblurred datasets as those used in this study to compute the correlation graphs, the correlations between parenchyma voxels in these networks are much lower than those between venous voxels, and the high correlation threshold used to construct the graphs ensures that only the latter are reflected in it. This subtlety of voxel time course correlation patterns is easily lost when using only blurred datasets, but can be visualized

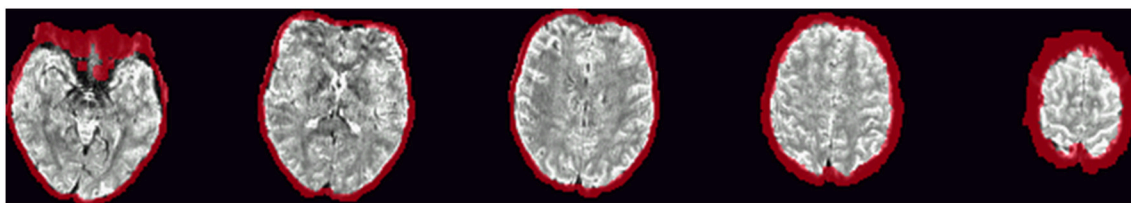


FIGURE 4 | Example of the definition of “brain edge” voxels, marked in red, and defined by erosion of the brain mask by a $5 \times 5 \times 5$ -voxel kernel.

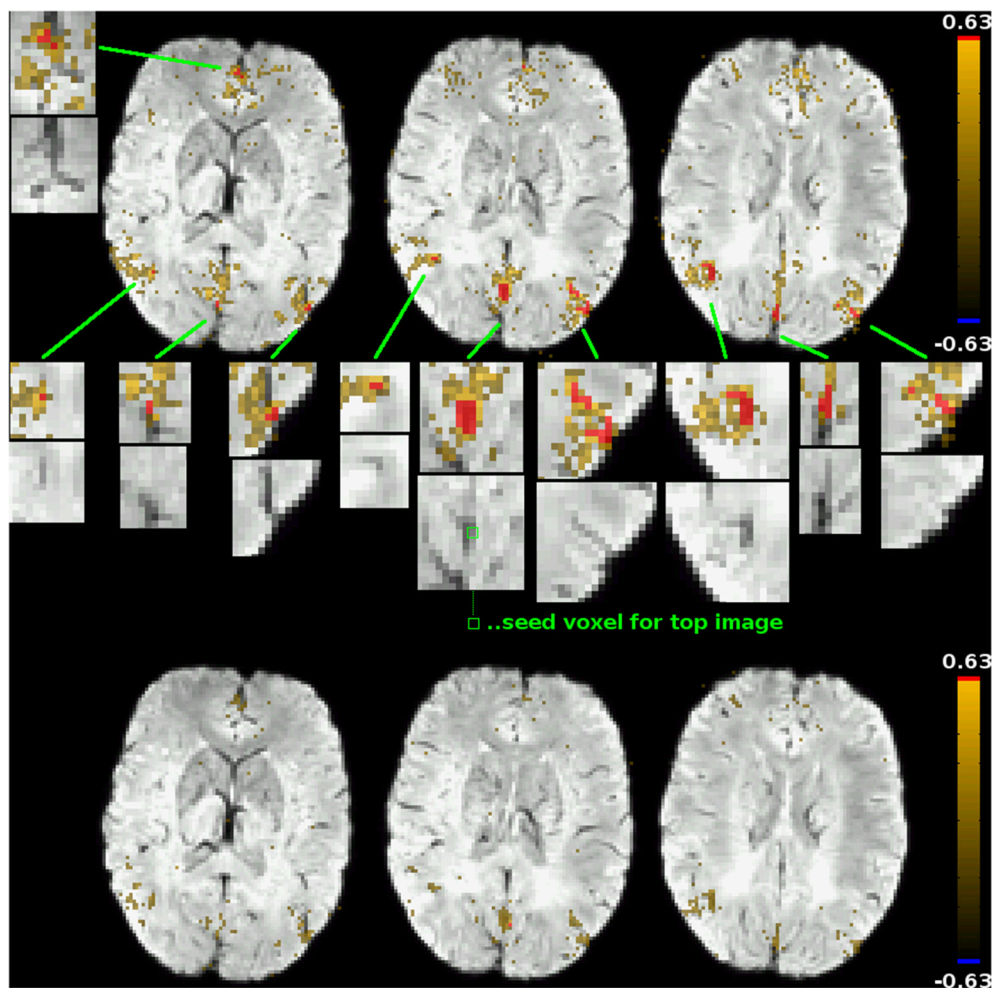


FIGURE 5 | Functional connectivity maps for two adjacent voxels in the PCC: one within a vein as identified by SWI (top subfigure; seed location marked out in the central zoomed insert), one outside of large veins (bottom subfigure; seed location identifiable as a red dot in the connectivity map). The underlay is the subject's SWI, coregistered to EPI space. Voxels exceeding that subject's correlation threshold for graph binarization (0.63) are marked out in red, maps are thresholded at a correlation coefficient of 0.25. Only for the venous seed do correlations exceeding this threshold exist; they are shown in zoomed inserts, along with the corresponding picture detail from the underlay alone. All voxels with a correlation coefficient exceeding the binarization threshold can be attributed to veins identifiable in the underlay. Correlations to other voxels of the default-mode network, including non-venous voxels, can be seen for both seeds, but correlation strengths in these voxels do not exceed the binarization threshold.

effectively using the DMN as an example. **Figure 5** shows the functional connectivity of two voxels in the posterior cingulate cortex (PCC), a main constituent of the DMN, one of them a voxel clearly in a vein as identified by SWI (its functional connectivity map being shown in the top part of the figure), the other being an adjacent voxel outside the vein (its functional connectivity map is shown in the bottom part of the figure). For the venous voxel, the correlation coefficients in adjacent venous voxels exceed 0.63, which was the threshold for graph construction in that particular subject (based on the network sparsity criterion of $S = 4$; this was the lowest connectivity threshold for all subjects, see **Table 1**), as well as a number of other voxels in typical DMN regions in the medial prefrontal cortex as well as bilaterally in the parietal cortices. On closer inspection, one notices that all of those voxels can be related

to veins identifiable on the SWI image (see zoomed inserts). Correlations with other voxels in DMN regions, including voxels further from large visible veins, also exist, but with far lower correlation strengths. When using a seed outside of veins, as exemplified by the connectivity map in the lower part of the figure of an adjacent voxel in the parenchyma, correlations above the threshold can be found neither in other DMN regions nor even in adjacent voxels. Indeed, the highest correlation coefficient found anywhere in the brain for that particular seed is 0.49—far away from the threshold of 0.63. This is consistent with results typically obtained from blurred resting-state datasets: signal time courses from veins draining one part of a resting-state network can be seen as reflecting the signal time courses of the voxels in the gray matter that these veins drain in a way similar to how the time course of a voxel in a blurred dataset reflects the

time courses of the voxels in its neighborhood, and time courses of voxels in the draining veins of different parts of a resting-state network are more strongly correlated with each other in the same way as correlation strengths between voxels in different parts of a resting-state network are increased in spatially blurred datasets. Correlation strengths between parenchyma voxels of resting-state networks are lower, and thus, if the correlation threshold used to generate a binarized graph from the voxelwise correlation matrix is high enough, these parenchyma voxel correlations are not reflected in the graph analyses performed after binarization, and only the connections between venous voxels remain in the graph and, ultimately, define the graph modules.

Of course, this work is not the first attempt at identifying artifactual signals in fMRI datasets. Previous attempts include both the use of complementary measurements (including SWI) as well as the localization of its effects based on EPI measurements alone. The advantage of using EPI-based identification over complementary measurements is the more immediate relationship to the application of the results to the fMRI analyses in question. In the results presented here, this can be seen in the slight differences in localization of venous voxels between the clustering and SWI-based masks in the context of general correspondence between the two (see for example **Figure 1**). While the general similarity between the features seen in the SWI and clustering masks corroborates the theory of venous origins of the signals seen in the graph cluster voxels, the difference in location highlights that imperfect correspondence between two different modalities, emerging from imperfect coregistration, geometric distortion or other origins, necessarily limit the use of inference from complementary measurements for the identification of affected voxels in the EPI image. In this sense, vein identification methods using EPI and SWI complement each other as different, independent modalities, and both are important to provide a link between the results from EPI images to signal sources not directly visible in BOLD EPI such as most veins. While SWI yields the more anatomically accurate maps of the venous architecture in a subjects' brain, methods based on post-processing techniques applied on the BOLD EPI data themselves add to this a more direct view on the immediate effect of these veins on fMRI measurements.

Among methods based on EPI measurements, two categories can be distinguished, the first being methods using externally measured signals and correlating them with time courses from the EPI measurement, and the second being methods based on the analysis of EPI time courses by themselves. The first category typically uses high-frequency physiological nuisance signals (usually heart rate and respiration monitoring), which, however, has a different spatial distribution than the venous signals investigated here (Windischberger et al., 2002): high-frequency physiological noise tends to be concentrated near arteries and the CSF, which is subject to the same pulsations, while the low-frequency physiological signals investigated here tend to be localized in or near the venous macrovasculature. Physiological low-frequency signals are acquired and analyzed only in very few studies, but studies using them have shown them to be quite useful in identifying blood flow related

phenomena in fMRI datasets (Tong and Frederick, 2012; Tong et al., 2014). The use of peripheral measurements, however, has one practical and one more fundamental limitation. The practical limitation is the necessity of additional hardware and measurement overhead for their acquisition leading to such measurements not always being available for all fMRI datasets, and the potential for additional error sources in their acquisition. While this issue can be overcome in any given study if the necessary steps are taken prospectively, it cannot be employed when analyzing datasets acquired without measuring these peripheral physiological signals, as is often the case in investigations using data shared by other researchers (Biswal et al., 2010; Kalcher et al., 2012). A more fundamental issue, though, is the time delay involved between the recording of the physiological signals at the external measurement location (e.g., the fingertip or toe for pulse oxymetry) and the brain, or, to be more precise, different locations in the brain. With standard EPI sequences as currently used, with a TR of between 2 and 3 s, the effect of this issue is rather limited, but with the current development toward short-TR multiband EPI sequences with higher temporal resolutions, the difficulties arising from the delay between the peripheral acquisition of physiological signals and their effect in the brain become more pronounced.

Finally, the use of measures directly derived from the EPI time series has been mostly confined to computationally less complex methods (e.g., the voxelwise time course standard deviation, a variant of which has been used in the Human Connectome Project), in part due to the lack of tools for tackling the computational challenges posed by more sophisticated methods like the voxel-by-voxel graph clustering approach presented here. Readily available tools from the domain of big data analysis can be applied to overcome computational obstacles and open the way to more comprehensive analysis tools. The comparison of time course standard deviations within and outside the graph clustering mask (see **Figure 3**) confirms the rationale behind the Human Connectome Project's preprocessing step of eliminating voxels with higher than normal standard deviations, but at the same time suggests that a one-dimensional measure not taking into account the connection structure between voxels might not yield a clear-cut discrimination threshold, as values of this score for normal brain tissue voxels with relatively high signal standard deviation and venous voxels with relatively low signal standard deviations overlap substantially.

In contrast, voxelwise graph analysis can be a useful tool to identify voxels in the macrovasculature by their highly correlated low-frequency signals. This latter point should be highlighted, as the band-pass filter applied (0.01–0.2 Hz) eliminates the possibility that the correlated signals in those voxels can merely be attributed to large-scale physiological noise (e.g., of respiratory or cardiac origin), which would have a higher frequency signal spectrum. Instead, they might exhibit more problematic signal fluctuations in the low-frequency domain, easily misattributed to local low-frequency fluctuations. In addition, the presence of such fluctuations might also be indicative of a risk of seeing downstream activations due to venous drainage of activations at

more distant voxels, as it occurs during some emotional-visual tasks in the BVR (Boubela et al., 2015). The identification of voxels at risk is thus a powerful tool to increase specificity in the interpretation of fMRI BOLD activations.

REFERENCES

- Behzadi, Y., Restom, K., Liao, J., and Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37, 90–101. doi: 10.1016/j.neuroimage.2007.04.042
- Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734–4739. doi: 10.1073/pnas.0911855107
- Boubela, R. N., Kalcher, K., Huf, W., Seidel, E.-M., Derntl, B., Pezawas, L., et al. (2015). fMRI measurements of amygdala activation are confounded by stimulus correlated signal fluctuation in nearby veins draining distant brain regions. *Sci. Rep.* 5:10499. doi: 10.1038/srep10499
- Clayden, J. (2014). *mmand: Mathematical Morphology in Any Number of Dimensions*. London, UK: R package version 1.1.0.
- Curtis, A. T., Hutchison, R. M., and Menon, R. S. (2014). Phase based venous suppression in resting-state BOLD GE-fMRI. *Neuroimage* 100, 51–59. doi: 10.1016/j.neuroimage.2014.05.079
- Curtis, A. T., and Menon, R. S. (2014). Highcor: a novel data-driven regressor identification method for BOLD fMRI. *Neuroimage* 98, 184–194. doi: 10.1016/j.neuroimage.2014.05.013
- Duong, T. Q., Yacoub, E., Adriany, G., Hu, X., Ugurbil, K., and Kim, S.-G. (2003). Microvascular BOLD contribution at 4 and 7 T in the human brain: gradient-echo and spin-echo fMRI with suppression of blood effects. *Magn. Reson. Med.* 49, 1019–1027. doi: 10.1002/mrm.10472
- Eickhoff, S. B., Thirion, B., Varoquaux, G., and Bzdok, D. (2015). Connectivity-based parcellation: critique & implications. *Hum. Brain Mapp.* doi: 10.1002/hbm.22933. [Epub ahead of print].
- Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Gunther, M., et al. (2010). Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS ONE* 5:e15710. doi: 10.1371/journal.pone.0015710
- Geißler, A., Fischmeister, F. P. S., Grabner, G., Wurnig, M., Rath, J., Foki, T., et al. (2013). Comparing the microvascular specificity of the 3- and 7-T BOLD response using ICA and susceptibility-weighted imaging. *Front. Hum. Neurosci.* 7:747. doi: 10.3389/fnhum.2013.00474
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the human connectome project. *Neuroimage* 80, 105–124. doi: 10.1016/j.neuroimage.2013.04.127
- Glover, G. H., Li, T. Q., and Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44, 162–167. doi: 10.1002/1522-2594(200007)44:1<162::AID-MRM23>3.0.CO;2-E
- Haacke, E. M., Xu, Y., Cheng, Y.-C. N., and Reichenbach, J. R. (2004). Susceptibility weighted imaging (SWI). *Magn. Reson. Med.* 52, 612–618. doi: 10.1002/mrm.20198
- Kalcher, K., Huf, W., Boubela, R. N., Filzmoser, P., Pezawas, L., Biswal, B., et al. (2012). Fully exploratory network independent component analysis of the 1000 functional connectomes database. *Front. Hum. Neurosci.* 6:301. doi: 10.3389/fnhum.2012.00301
- Lai, S., Hopkins, A. L., Haacke, E. M., Li, D., Wasserman, B. A., Buckley, P., et al. (1993). Identification of vascular structures as a major source of signal contrast in high resolution 2D and 3D functional activation imaging of the motor cortex at 1.5T: preliminary results. *Magn. Reson. Med.* 30, 387–392. doi: 10.1002/mrm.1910300318
- Menon, R. S. (2002). Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI. *Magn. Reson. Med.* 47, 1–9. doi: 10.1002/mrm.10041
- Menon, R. S. (2012). The great brain versus vein debate. *Neuroimage* 62, 970–974. doi: 10.1016/j.neuroimage.2011.09.005
- Nencka, A. S., and Rowe, D. B. (2007). Reducing the unwanted draining vein BOLD contribution in fMRI with statistical post-processing methods. *Neuroimage* 37, 177–188. doi: 10.1016/j.neuroimage.2007.03.075
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8577–8582. doi: 10.1073/pnas.0601602103
- Norris, D. G. (2012). Spin-echo fmri: the poor relation? *Neuroimage* 62, 1109–1115. doi: 10.1016/j.neuroimage.2012.01.003
- Reichenbach, J. R., Barth, M., Haacke, E. M., Klarhöfer, M., Kaiser, W. A., and Moser, E. (2000). High-resolution MR venography at 3.0 tesla. *J. Comput. Assist. Tomogr.* 24, 949–957. doi: 10.1097/00004728-200011000-00023
- Rowe, D. B., and Logan, B. R. (2005). Complex fMRI analysis with unrestricted phase is equivalent to a magnitude-only model. *Neuroimage* 24, 603–606. doi: 10.1016/j.neuroimage.2004.09.038
- Tong, Y., and Frederick, B. D. (2012). Concurrent fNIRS and fMRI processing allows independent visualization of the propagation of pressure waves and bulk blood flow in the cerebral vasculature. *Neuroimage* 61, 1419–1427. doi: 10.1016/j.neuroimage.2012.03.009
- Tong, Y., Hocke, L. M., and Frederick, B. D. (2014). Short repetition time multiband echo-planar imaging with simultaneous pulse recording allows dynamic imaging of the cardiac pulsation signal. *Magn. Reson. Med.* 72, 1268–1276. doi: 10.1002/mrm.25041
- Turner, R. (2002). How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes. *Neuroimage* 16, 1062–1067. doi: 10.1006/nimg.2002.1082
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., et al. (2013). The WU-Minn human connectome project: an overview. *Neuroimage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041
- Windischberger, C., Langenberger, H., Sycha, T., Tschernko, E. M., Fuchsberger, G., Schmetterer, L., et al. (2002). On the origin of respiratory artifacts in BOLD-EPI of the human brain. *Magn. Reson. Imaging* 20, 575–582. doi: 10.1016/S0730-725X(02)00563-5

ACKNOWLEDGMENTS

This study was financially supported by the Austrian Science Fund (P22813, P23533).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kalcher, Boubela, Huf, Naßel and Moser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Voxel-Wise Functional Connectome Can Be Efficiently Derived from Co-activations in a Sparse Spatio-Temporal Point-Process

Enzo Tagliazucchi^{1,2,3*}, Michael Siniatchkin¹, Helmut Laufs^{2,4} and Dante R. Chialvo^{5,6}

¹ Institute for Medical Psychology, Christian-Albrechts University, Kiel, Germany, ² Department of Neurology and Brain Imaging Center, Goethe University Frankfurt am Main, Germany, ³ Department of Sleep and Cognition, Netherlands Institute for Neuroscience, Amsterdam, Netherlands, ⁴ Department of Neurology, University Hospital Schleswig-Holstein, Christian-Albrechts-University Kiel, Kiel, Germany, ⁵ Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Buenos Aires, Argentina, ⁶ Center for Multidisciplinary Complex Systems Studies and Brain Sciences (CEMSC3), Escuela de Ciencia y Tecnología, Universidad Nacional de San Martín, Buenos Aires, Argentina

OPEN ACCESS

Edited by:

Brian Caffo,
Johns Hopkins University, USA

Reviewed by:

Joshua T. Vogelstein,
Johns Hopkins University, USA
Jian Kang,
Emory University, USA

*Correspondence:

Enzo Tagliazucchi
tagliazucchi.enzo@gmail.com

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 24 June 2015

Accepted: 04 August 2016

Published: 23 August 2016

Citation:

Tagliazucchi E, Siniatchkin M, Laufs H
and Chialvo DR (2016) The
Voxel-Wise Functional Connectome
Can Be Efficiently Derived from
Co-activations in a Sparse
Spatio-Temporal Point-Process.
Front. Neurosci. 10:381.
doi: 10.3389/fnins.2016.00381

Large efforts are currently under way to systematically map functional connectivity between all pairs of millimeter-scale brain regions based on large neuroimaging databases. The exploratory unraveling of this “functional connectome” based on functional Magnetic Resonance Imaging (fMRI) can benefit from a better understanding of the contributors to resting state functional connectivity. In this work, we introduce a sparse representation of fMRI data in the form of a discrete point-process encoding high-amplitude events in the blood oxygenation level-dependent (BOLD) signal and we show it contains sufficient information for the estimation of functional connectivity between all pairs of voxels. We validate this method by replicating results obtained with standard whole-brain voxel-wise linear correlation matrices in two datasets. In the first one ($n = 71$), we study the changes in node strength (a measure of network centrality) during deep sleep. The second is a large database ($n = 1147$) of subjects in which we look at the age-related reorganization of the voxel-wise network of functional connections. In both cases it is shown that the proposed method compares well with standard techniques, despite requiring only data on the order of 1% of the original BOLD signal time series. Furthermore, we establish that the point-process approach does not reduce (and in one case increases) classification accuracy compared to standard linear correlations. Our results show how large fMRI datasets can be drastically simplified to include only the timings of large-amplitude events, while still allowing the recovery of all pair-wise interactions between voxels. The practical importance of this dimensionality reduction is manifest in the increasing number of collaborative efforts aiming to study large cohorts of healthy subjects as well as patients suffering from brain disease. Our method also suggests that the electrophysiological signals underlying the dynamics of fMRI time series consist of all-or-none temporally localized events, analogous to the avalanches of neural activity observed in recordings of local field potentials (LFP), an observation of potentially high neurobiological relevance.

Keywords: functional connectome, functional connectivity, dimensionality reduction, point process, resting state fMRI

INTRODUCTION

The human brain comprises an interconnected network of cortical and sub-cortical regions globally linked by anatomical long-range tracts of connections. The mapping of the corresponding functional connections at a particular spatial scale (termed connectome in contemporary neuroscience; Sporns et al., 2004; Sporns, 2011) is an important ingredient in the process of understanding how the human brain can perform diverse cognitive functions. Furthermore, many neurological and psychiatric diseases can be understood in terms of deviations from a healthy connectome (Fox and Greicius, 2010; Kelly et al., 2012).

Advances in neuroimaging methods, such as Diffusion Tensor Imaging (DTI) and Diffusion Spectrum Imaging (DSI) allow the *in vivo* mapping of the human structural connectome at a large-scale (Hagmann et al., 2008). Blood oxygenation level-dependent (BOLD) functional Magnetic Resonance Imaging (fMRI) allows for a functional counterpart of the anatomical connectome, a notion first introduced about a decade ago (Sporns et al., 2004; Eguiluz et al., 2005; Salvador et al., 2005) by computing the statistical covariance between all pairs of BOLD signals. This functional connectome contains information on how all pairs of regions (at a certain spatial scale) relate dynamically and collectively with each other.

These two approaches are being applied by international coordinated efforts to systematically map connectomes in very large populations of subjects and at the highest temporal and spatial resolution currently available (see for instance Biswal et al., 2010; Smith et al., 2013; Van Essen et al., 2013). These efforts will eventually lead to the availability of large-scale databases useful to account for potential inter-subject variability caused by different demographical variables, as well as to reduce the harmful effect of noise and artifacts through massive averaging.

These collaborative efforts need to be paralleled by methodological developments facilitating efficient extraction of relevant information from the data. Common strategies are based on averaging BOLD signals over brain parcellations comprising extended regions, thus reducing the dimensionality of the problem as well as the number of required computations. However, there are several problems inherent to this approach. First, all detail of the functional connectome inside each region of the parcellation is lost. Second, partitions are usually arbitrary and therefore might sub-divide a functionally coherent region into many regions. Different studies have addressed how the properties of parcellation-based networks can change depending on region selection (Wang et al., 2009; Zalesky et al., 2010). Third, efforts to increase the spatial resolution of fMRI sequences are pointless if data will be down-sampled after acquisition by averaging BOLD signals inside a small number of regions in a parcellation.

The objective of this paper is to show how a very sparse representation of brain activity, namely a discrete spatio-temporal point-process, is able to estimate the whole brain voxel-wise functional connectome. This point-process is derived from the times at which the BOLD signals reach some maximum

level of activity, either by detecting crossings of an arbitrary threshold, or by the identification of local peaks, i.e., the point-process comprises large amplitude events in the data. At its core, our proposed method is based on identifying a basis of discrete contributions to resting state fMRI signals, in analogy to other neural recording modalities (such as spikes in intra- and extra-cellular recordings). Following this analogy, once the relevant events are identified, much of the signal (i.e., the stereotypical response associated with a discrete event) can be disregarded without reducing their information content, facilitating data storage, manipulation and interpretation (this analogy is limited, however, since neural recordings provide more sampling points than fMRI recordings and hence a larger number of discrete events). The main merit of this method is to reduce the continuous representation of BOLD signals into a series of timings associated with events of interest, thus (1) drastically reducing the dimensionality of the data, (2) abstracting the relevant information from sources of noise.

It has been shown previously that this method suffices to reproduce large-scale patterns of coordinated activity (Tagliazucchi et al., 2011, 2012a) termed Resting State Networks (RSN; Beckmann et al., 2005) and is essentially identical to the de-convolution of the signals as a series of discrete impulse functions (Petridou et al., 2013). Furthermore, de-convolution into a point-process can lessen the impact of hemodynamic lags for the estimation of causality between BOLD signals (Wu et al., 2013). Here we contribute a systematic evaluation of the capacity of this method to reproduce *all* bivariate relationships between signals (i.e., whole-brain correlation matrices). This validation is obtained, for the first time, from a large database of subjects $n = 1147$ scanned with different parameters at different locations, thus supporting its universal validity.

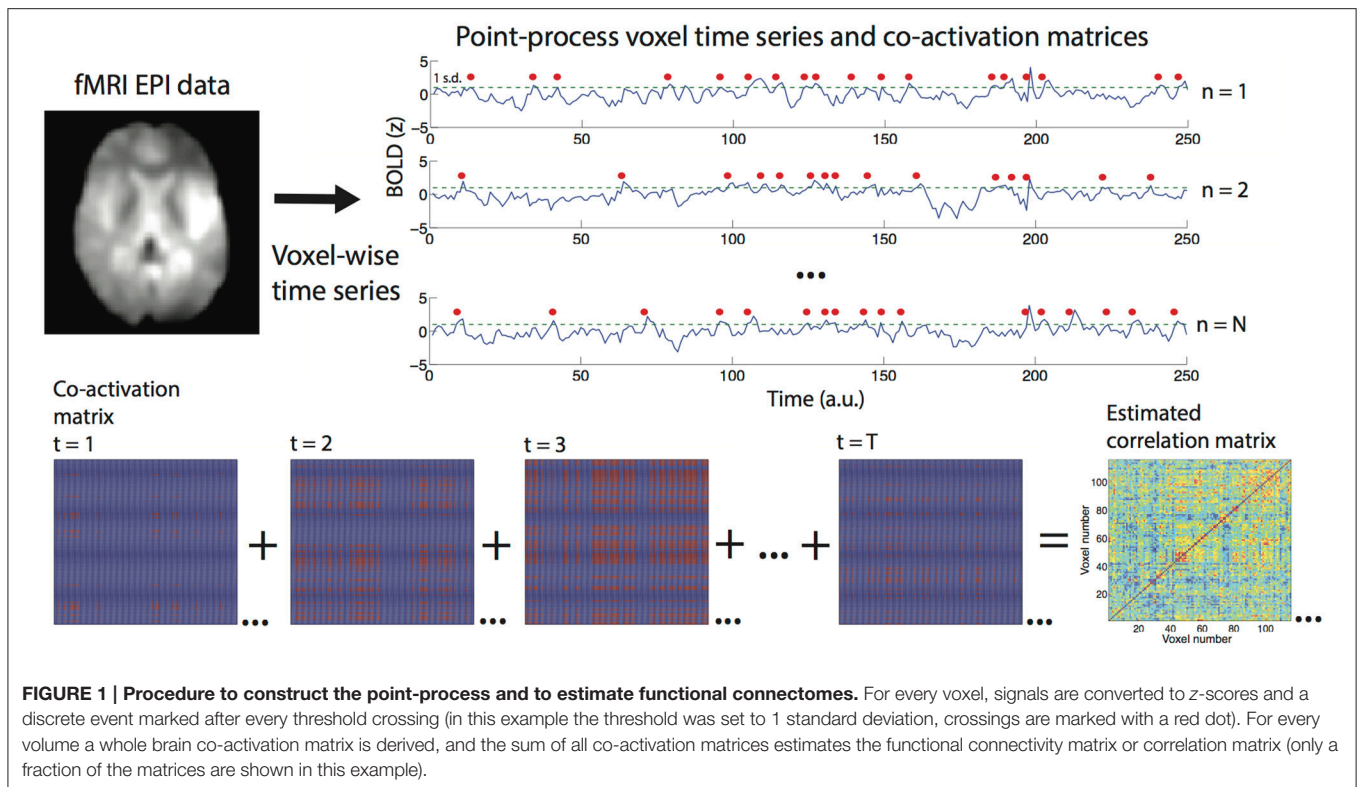
We also investigated whether abstracting the signal into a point-process could yield benefits from the perspective of reducing confounds and noise in the data. For this we adopted a practical, classification-based approach, investigating how accurately connectivity matrices derived from the point-process and from linear correlations could distinguish two groups of subjects (younger and older subjects from the $n = 1147$ database). We hypothesized that keeping high-amplitude events in the data could disregard low-amplitude noise and result in a better classification accuracy than the one obtained using full BOLD signals.

MATERIALS AND METHODS

We will first describe all steps of the proposed method and then introduce different datasets used for validation as well as to show possible applications. The general procedure followed to estimate correlation networks via the point-process analysis is graphically outlined in **Figure 1**.

Voxel-Wise Correlation Matrix

Consider an fMRI measure consisting of N voxels and T volumes, represented as $F_n(t)$, with $0 \leq n \leq N$ and $0 \leq t \leq T$. Thus, $F_n(t)$ represents the BOLD signal at voxel n and time t . The common



definition of voxel-wise correlation matrix (Eguiluz et al., 2005) is as follows,

$$R_{ij} = \frac{< (F_i - < F_i >)(F_j - < F_j >) >}{\sigma(F_i)\sigma(F_j)} \quad (1)$$

where $< F_i >$ and $\sigma(F_i)$ represent the mean value and the standard deviation of the BOLD signal at the voxel i , respectively. Note that according to this definition, for the computation of R_{ij} , Equation (1) must be evaluated $N(N - 1)/2$ times (although not serially in efficient implementations). Often these calculations are used to define functional connectivity networks which in turn allow for further analysis of the resulting graphs.

Constructing the Point-Process

The approach here proposed starts with converting the BOLD signal at every voxel into its z-score, $\tilde{F}_i = \frac{F_i - < F_i >}{\sigma(F_i)}$. This is done under the assumption that, according to our formalism, the absolute amplitude of the BOLD signal carries less information than its temporal evolution (for the biological underpinnings of this assumption please see the Discussion section). To define the point-process, the *a priori* arbitrary threshold γ is selected and the spatio-temporal process $PP_i(t)$ is defined as follows:

$$PP_i(t) = \begin{cases} 1 & \text{if } \tilde{F}_i(t) < \gamma \text{ and } \tilde{F}_i(t+1) > \gamma \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

This point-process was introduced in a previous publication (Tagliazucchi et al., 2012a) where we showed that it suffices to

replicate the topographical features of the major canonical RSN, even though for most values of t and i , $PP_i(t)$ will be zero (indeed, taking $\gamma = 1$, for a signal of $T = 240$ on average the point-process is non-zero for 15 ± 3 time points, or approximately 6% of the data—see Tagliazucchi et al., 2012a). Note that once the point-process is constructed much of the data can be discarded. From a signal comprising 240 values, only a series of (on average) 15 numbers needs to be retained, namely, the timings of the events in the point-process. Clearly, this results in a considerable compression of the fMRI data.

Alternatively, $PP_i(t)$ can be defined by the (high amplitude) local peaks of the BOLD signal. For this, BOLD signals are also converted to z-scores and all sufficiently large peaks (for instance, those above an arbitrary threshold) are the points represented in $PP_i(t)$. The formal definition is as follows,

$$PP_i(t) = \begin{cases} 1 & \text{if } \tilde{F}_i(t) > \tilde{F}_i(t-1) \text{ and } \tilde{F}_i(t) > \tilde{F}_i(t+1) \text{ and} \\ & \tilde{F}_i(t) > \gamma \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Although formally both methods are justified, it will be shown later that either definition of the point-process leads to similar results.

Estimating Correlations from the Point-Process

After converting F_i into $PP_i(t)$ we introduce the following framework to generalize the methods introduced in Tagliazucchi et al. (2012a), from the estimation of seed based correlations

to the efficient computation of all pairs of correlations between voxels. We first define the co-activation matrices $A_{ij}(t)$ as follows:

$$A_{ij}(t) = PP_i(t) PP_j(t) \quad (4)$$

Note that according to this definition, $A_{ij}(t)$ only has two possible values: $A_{ij}(t) = 1$ if at time t the point-process is non-zero both at voxels i and j , and $A_{ij}(t) = 0$ otherwise.

The co-activation matrices defined in Equation (4) can be used to estimate the functional connectivity between all pairs of voxels in the brain by performing a simple matrix addition. Two highly synchronized signals will cross the threshold together most of the time, thus a measure of coupling between the signals can be obtained by counting the number of times the signals crossed the threshold together. This is formalized simply by,

$$C_{ij} = \sum_{t=1}^T A_{ij}(t) = \sum_{t=1}^T PP_i(t) PP_j(t) \quad (5)$$

In matrix notation, this can be succinctly summarized as $C = PP PP^T$, considering PP as a matrix with voxels as rows and time as columns and containing the point-process. The matrix C_{ij} contains in its i, j entry the number of shared co-activations between BOLD signals at voxels i and j . Note that since all A_{ij} are symmetrical matrices, then C_{ij} is also symmetrical. Note also that the matrices $A_{ij}(t)$ contain valuable information about instantaneous co-activations between voxels and as such their analysis might be important to understand the temporal evolution of large-scale synchronization between brain regions (Tagliazucchi et al., 2012b; Hutchison et al., 2013).

The main issue with this matrix as a measure of functional connectivity is that it is not normalized, therefore there is no way to directly decide (for instance) if a perfect synchronization between signals has been reached. An appropriate normalization for this matrix would be as follows,

$$\tilde{C}_{ij} = \frac{C_{ij}}{\max\left(\sum_{t=0}^T PP_i, \sum_{t=0}^T PP_j\right)} = \frac{C_{ij}}{\max(C_{ii}, C_{jj})} \quad (6)$$

This definition of \tilde{C}_{ij} is reasonable since C_{ij} achieves its highest possible value if all threshold crossings are also shared between both voxels. However, one voxel could have all its threshold crossings in common with the other, whereas the opposite might not be true (since the other voxel could have a larger number of crossings in total (this can be the case only if $C_{ii} \neq C_{jj}$), thus normalizing using the maximum between the number of crossings at both voxels is required. Also, \tilde{C}_{ij} is symmetrical with this normalization.

The normalization presented in Equation (6) requires the maximum value between the numbers of threshold crossings at all pairs of voxels. If normalization is needed, then a more efficient approximate solution is to divide by the number of threshold crossings without taking the maximum value, for instance, across rows or columns of the matrix, and then

symmetrizing (if needed) the result by averaging with the transpose:

$$\tilde{C}_{ij} = \frac{1}{2} \left[\frac{C_{ij}}{C_{ii}} + \frac{C_{ji}}{C_{ii}} \right] \quad (7)$$

Note that $\frac{C_{ij}}{C_{ii}}$ deviates from a symmetrical matrix only in the case of different numbers of threshold crossing between voxels ($C_{ii} \neq C_{jj}$). Note also that normalization might not be necessary if comparing fixed-length recordings between two populations, under the assumption that the rate of events in the point-process is not different between groups.

For the computation of \tilde{C}_{ij} all steps can be performed efficiently in vectorized form in any language with matrix manipulation capabilities (for instance, MATLAB or Python with NumPy), in particular, after constructing the point-process in Equation (2), the operations involved consist of a single matrix multiplication (Equations 4 and 5), multiplication by scalars and matrix symmetrization (Equation 7). In this work, all computations were performed using a 8 core CPU running at 2400 MHz with a total of 128 GB built-in memory.

After introducing the core methods, we now discuss the methodology for the validation of our results.

Measures Derived Whole Brain Voxel-Wise Correlations Used for Method Validation

The number of connections derived in a voxel-wise analysis complicates easy visualization of networks and their changes across conditions. Thus, in the many applications of functional connectomes found in the literature, rarely whole-brain voxel-wise networks are directly visualized. Instead, lower-dimensional metrics are to be derived, which are easy to visualize as 3D maps overlaid on brain anatomy. One possible choice is to assess measures of network centrality, this is, how important nodes are in the network, thus collapsing all connections attached to a node into a single number. A straightforward definition in a weighted network is the strength (Barthelemy et al., 2005), defined as:

$$S_i = \sum_{j=1}^N R_{ij} \quad (8)$$

In the present case, using the point-process to estimate correlations, R_{ij} is replaced by \tilde{C}_{ij} . Nodes with the highest strength values are termed hubs and their reorganization has been repeatedly linked to different brain pathologies (Crossley et al., 2014), such as coma (Achard et al., 2012) or Alzheimer's disease (Buckner et al., 2009).

Note that the evaluation of Equation (8) requires the whole brain correlation network. In the case of a voxel-wise network, centrality of nodes (i.e., voxels) can be easily visualized as a 3D map overlaid on an anatomical image.

Another measure employed for validation of our method is the interhemispheric or homotopic connectivity. This is defined as the correlation between the BOLD signal of every voxel and the contralateral voxel. Interhemispheric connectivity is in particular useful to quantify re-organization of functional connectomes for

which left-right asymmetries are expected (as in the case of aging, see Dolcos et al., 2002).

Datasets

To demonstrate the validity of our proposal two different datasets from previously published studies will be used. The first dataset comprises BOLD fMRI recordings from the 1000 Functional Connectomes database, and the second dataset comprises recordings from a recently published study in which combined EEG, EMG, BOLD-fMRI, and physiological data were obtained from 71 subjects.

The Connectome dataset was downloaded from the 1000 Functional Connectome Project online database (http://fcon_1000.projects.nitrc.org). Demographics, scanning parameters, and experimental conditions are described in the database website as well as in Tagliazucchi and Laufs (2014). Only epochs of wakefulness were employed in the present analysis. For more information on sleep vs. wakefulness classification in this dataset (see Tagliazucchi et al., 2012c; Tagliazucchi and Laufs, 2014). Since individual data presents variable length in this data set, normalization (Equation 7) was always required.

Data from a previously published study (Tagliazucchi and Laufs, 2014) was used for the sleep dataset. A total of 71 subjects were selected from a larger dataset on the basis of successful multimodal polysomnographic data recording and quality (written informed consent, approval by the local ethics committee). All subjects were scanned during the evening and instructed to close their eyes and lie still and relaxed. A group of 55 subjects was formed out of the original dataset of 71 subjects by excluding subjects who did not fall asleep. Hypnograms obtained via expert sleep staging based on AASM rules (American Academy of Sleep Medicine, 2007) were scanned for contiguous epochs of wakefulness, N1, N2, and N3 sleep lasting 250 volumes (~ 2 min), resulting in 84 epochs of wakefulness, 16 epochs of N1 sleep, 19 epochs of N2 sleep, and 20 epochs of N3 sleep. Sleep epochs are present (by construction) fixed length in this data set (250 volumes), therefore normalization (Equation 7) was not required under the assumption that sleep does not modify the rate of points in the data.

EEG was recorded via a cap (modified BrainCapMR, EasyCap, Herrsching, Germany) during fMRI acquisition (1505 volumes of T2*-weighted echo planar images, TR/TE = 2080/30 ms, matrix 64×64 , voxel size $3 \times 3 \times 2$ mm³, distance factor 50%; FOV 192 mm²) at 3 T (Siemens Trio, Erlangen, Germany) with an optimized polysomnographic setting [chin and tibial EMG, ECG, EOG recorded bipolarly (sampling rate 5 kHz, low pass filter 1 kHz), 30 EEG channels recorded with FCz as the reference (sampling rate 5 kHz, low pass filter 250 Hz), and pulse oxymetry, respiration recorded via sensors from the Trio (sampling rate 50 Hz)] and MR scanner compatible devices (BrainAmp MR+, BrainAmp ExG; Brain Products, Gilching, Germany).

MRI and pulse artifact correction were performed based on the average artifact subtraction (AAS) method (Allen et al., 1998) as implemented in Vision Analyzer2 (Brain Products, Germany) followed by objective (CBC parameters, Vision Analyzer) ICA-based rejection of residual artifact-laden components after AAS resulting in EEG with a sampling rate of 250 Hz. Good quality

EEG was obtained, which allowed sleep staging by an expert according to the AASM criteria (American Academy of Sleep Medicine, 2007).

fMRI Preprocessing

Using Statistical Parametric Mapping (SPM8) EPI data were realigned, normalized (MNI space) and spatially smoothed (Gaussian kernel, 8 mm³ full width at half maximum). The data were band-pass filtered in the range 0.01–0.1 Hz using a sixth order Butterworth filter. The same procedure was applied to the sleep dataset and to the 1000 Functional Connectomes dataset.

Multivariate Classification

We compared the accuracy of a Random Forest classifier with 100 estimators (implemented in scikit-learn, <http://scikit-learn.org/stable/>) to distinguish younger (<20 years) and older (>40 years) subjects from the 1000 Functional Connectomes dataset. This was based both on strength and interhemispheric connectivity maps obtained via normalized co-activation matrices (derived from the point-process) and standard linear correlation matrices. We applied a 5-fold cross validation procedure combined with feature selection (*F*-test to retain the top 10, 25, 50, 75% features), as well as with all features. Accuracy was reported as the area under the receiver operator characteristic (ROC) curve (AUC).

RESULTS

Correlations between \tilde{C}_{ij} and R_{ij}

We obtained the point-process for both datasets following the procedure illustrated in **Figure 1** and in the methods section. In the case of the 1000 Functional Connectomes dataset we repeated calculations both for voxel-wise networks and for networks based on time series extracted from the AAL template. Using this data, we first evaluated the similitude in the estimation of the connectivity matrix by both methods (point-process analysis with normalization and linear correlations) as a function of the threshold γ used to define the point-process (see Equation 2). Results are shown in **Figure 2** (left) for the average correlation between connectivity networks estimated by both methods as a function of γ . Correlations peaked at 0.6 and were highest for ≈ 0.7 . The histogram of all 1147 correlations obtained using $\gamma = 1$ (**Figure 2**, center) revealed a sharp peak around the mean value. The plot of the entries of the estimated correlation (values of \tilde{C}_{ij}) and the linear correlation (entries of R_{ij}) is shown in **Figure 2** (right). A monotonously increasing relationship was present between both quantities, even though the functional dependency between them was not linear. For low linear correlation values, the point-process co-activation increased slowly and did so more quickly for larger linear correlation values.

We compared the performance of computing voxel-wise functional connectivity matrices using the proposed point-process based method vs. standard linear correlations. In **Figure 2B**, left, the percentage of the time required using linear correlations (*corrcoef.m* MATLAB function, average time 131.48 s on a reference system) was plotted as a function of the threshold. At every threshold value a total of 100 iterations were performed for a single subject and results were then averaged. For

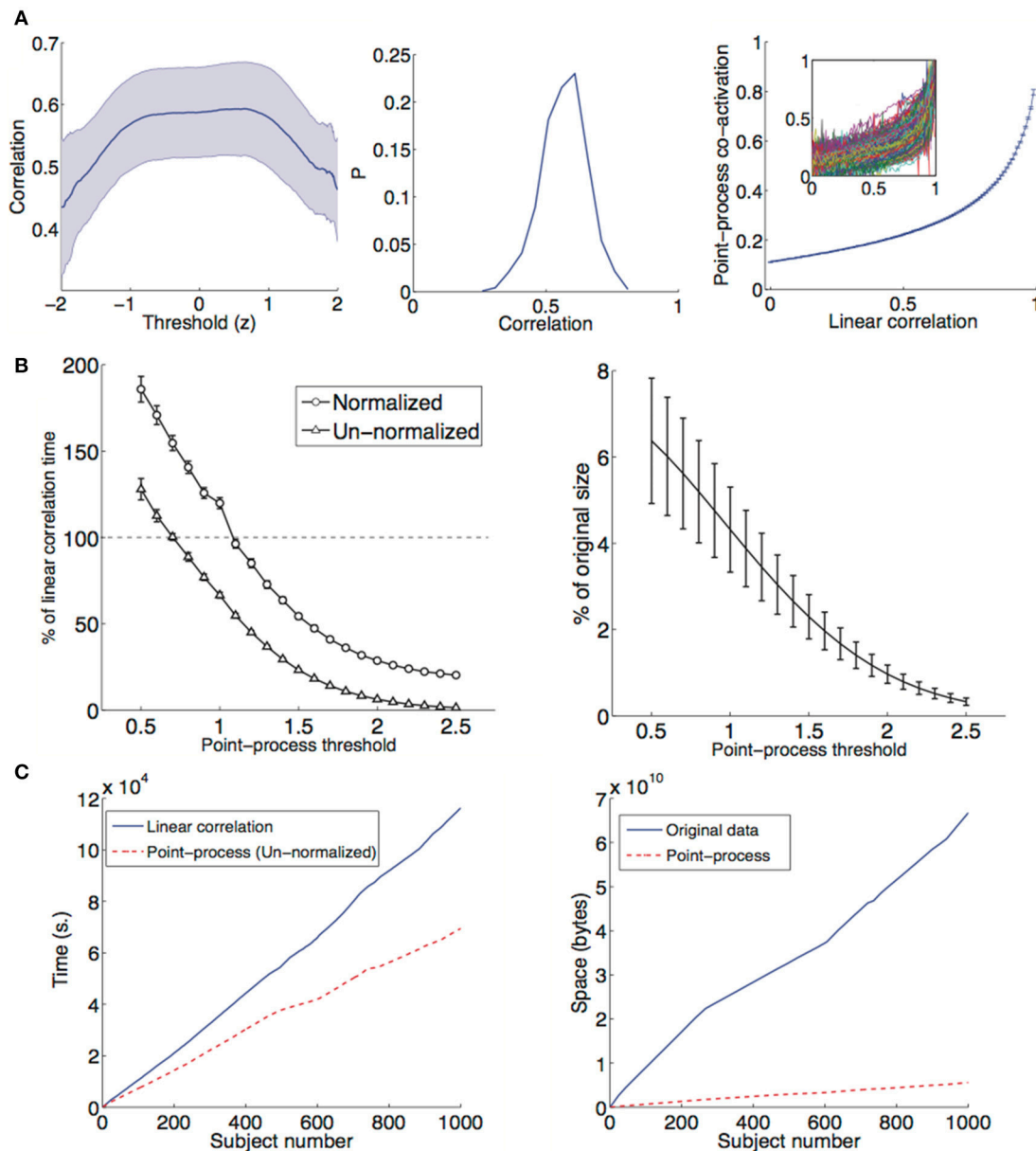


FIGURE 2 | (A) Left: Correlation between R_{ij} and \tilde{C}_{ij} as a function of the threshold [γ in Equation (2); mean \pm SEM]. Connectivity networks were derived from 116 time series extracted from the AAL template in all subjects from the 1000 Functional Connectomes dataset ($n = 1147$). Center: Histogram of all correlation values at $\gamma = 1$, $P =$ probability. Right: Average (mean \pm SEM) plot of the linear correlation coefficient between brain regions (entries of R_{ij}) and the estimate from the point-process analysis (entries of \tilde{C}_{ij}). The inset shows the plot for each one of the 1147 subjects. **(B)** Left: Performance of the point-process based estimation of functional connectivity as a function of the threshold γ (mean \pm SD). Elapsed computation times were obtained for a single subject across 100 repetitions and compared with the performance using linear correlations. Right: Percentage of the original number of data points retained after converting the data to a sparse point-process with $\gamma = 1$, plotted as a function of the threshold (for all subjects in the 1000 Functional Connectomes dataset). **(C)** Left: Cumulative computation time required to compute whole-brain voxel-wise connectivity matrices from 1000 subjects extracted from the Functional Connectomes dataset. An un-normalized point-process with $\gamma = 1$ was used. Right: Cumulative space required to store 1000 subjects from the Functional Connectomes dataset, both for the full data and for a sparse representation based on a point-process with $\gamma = 1$.

thresholds larger than approximately 1 standard deviation, the point-process based method slightly outperformed the standard computation, with performance becoming increasingly better as the threshold was increased and less points were included

in the analysis. However, more evidence needs to be gathered to confirm that the method outperforms the standard linear correlation approach, considering that the routines have not been properly optimized. In **Figure 2B** (right) we plot the percentage

of data points retained after conversion to the point-process. Even for the smallest threshold values, only about 6% of the data was retained. Thus, this very sparse representation of fMRI data contained sufficient information to capture all the differences during deep sleep and in the 1000 Functional Connectomes dataset (see below), requiring but a small fraction of the original time series. Specifically, the required information consists of the (discrete) timings of the events in the point-process (i.e., at which volumes the “points” appear).

To gauge the usefulness of our approach in a real setting, we computed the cumulative time and space required to process (i.e., obtain whole-brain voxel-wise connectivity matrices) and store 1000 subjects extracted from the Functional Connectomes dataset. Results are shown in **Figure 2C**. An un-normalized point-process with a threshold of $\gamma = 1$ resulted in a reduction of computation time (reference system) from a total of ≈ 30 h to ≈ 19 h. However, we note again that more careful experiments need to be performed to compare the time performance of both methods.

We also investigated the sparseness (defined as the percentage of zeros) in the point-process time series and in the associated normalized connectivity matrices (derived via point-process

co-activations). The results are shown in **Figure 3A**. Not only the time series are very sparse ($\approx 95\%$ zeros for a threshold of 1 S.D.) but also the connectivity matrices ($\approx 50\%$ zeros for the same threshold). This results in dramatically smaller file sizes when both the time series and the connectivity matrices are stored (**Figure 3B**).

Strength Maps in Wakefulness vs. Deep Sleep

To compare results obtained by both methods, we applied them to derive the strength maps (Equation 8) from the estimated whole brain voxel-wise correlations in the sleep dataset and to reveal changes between wakefulness and deep sleep. A total of 20 2-min epochs of deep sleep and 84 epochs of wakefulness could be extracted. After deriving the correlation networks, Equation (8) was applied to obtain the voxel-wise spatial distribution of strengths. Results for the contrast wakefulness > deep sleep are shown in **Figure 4A**, both for normalized and un-normalized co-activation matrices, as well as for the point-process derived from BOLD signal peaks instead of threshold crossings. Spatial patterns of decreased strength in deep sleep (comprising frontal, cingulate, primary visual, motor, and auditory cortices) were

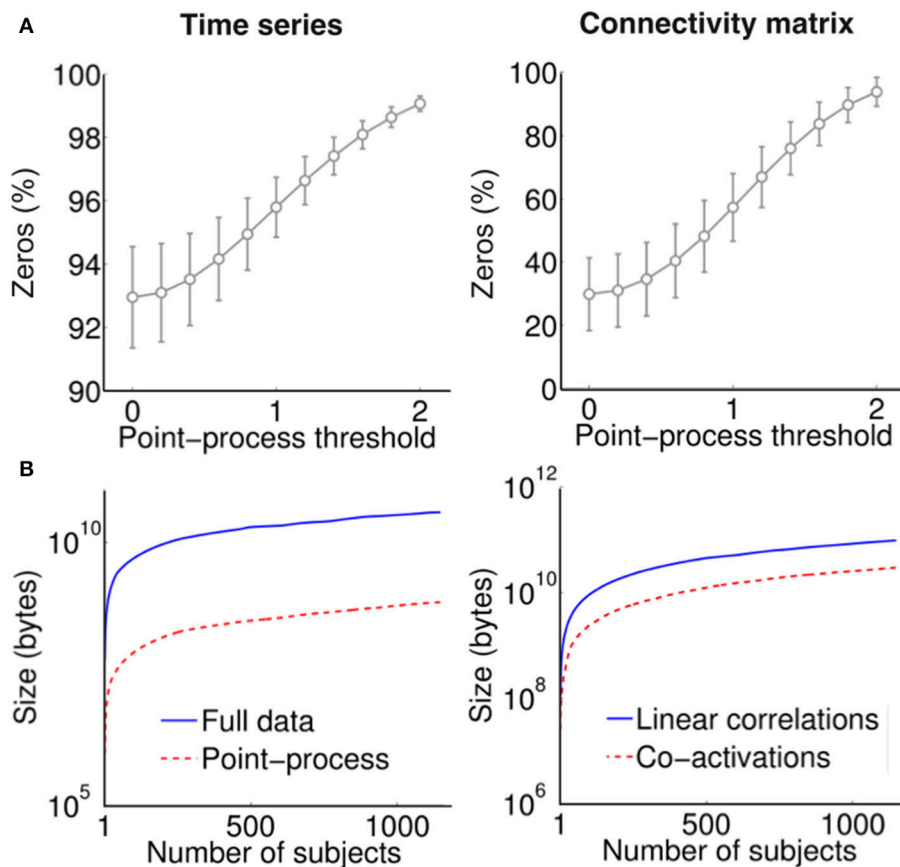
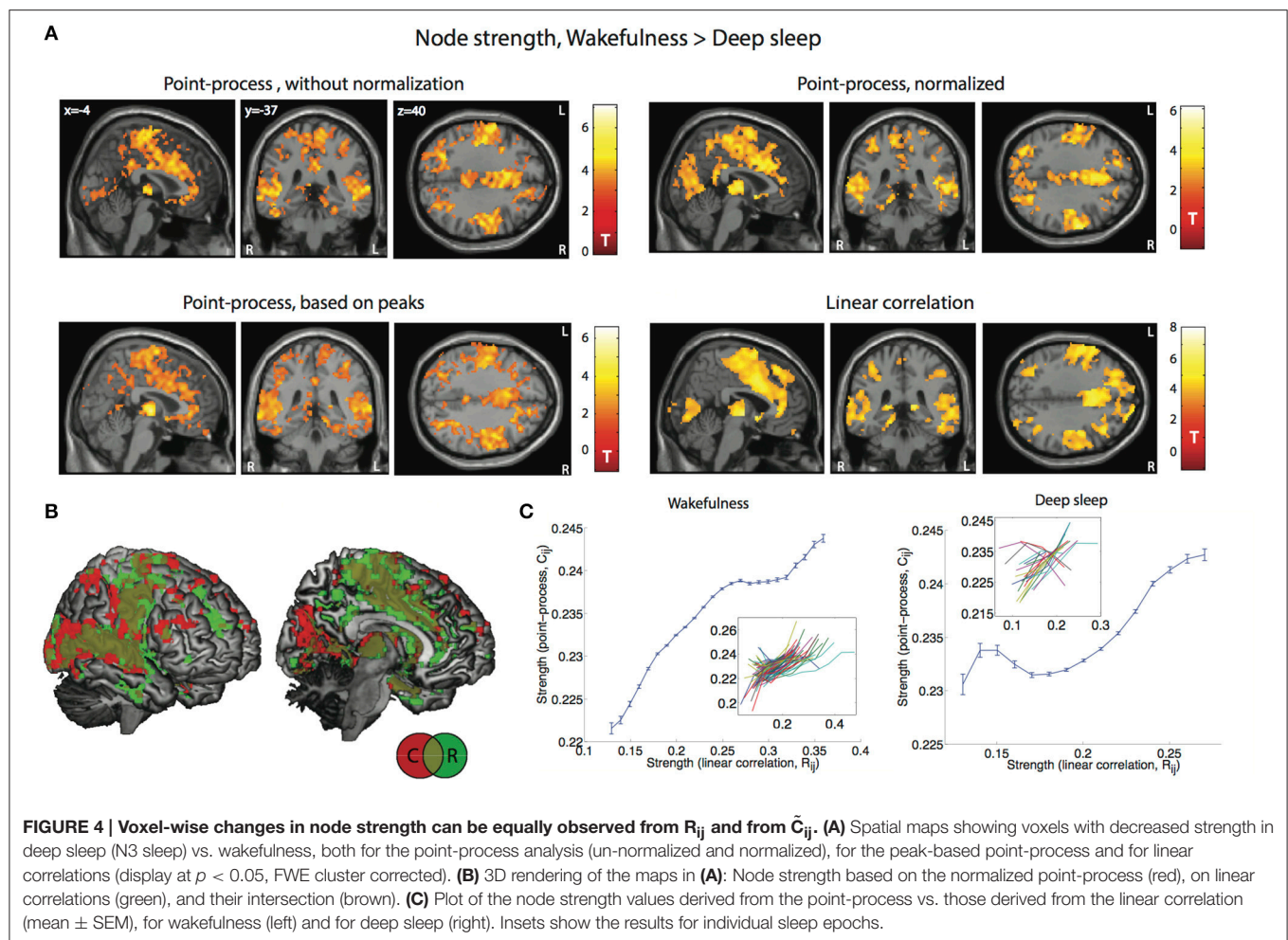


FIGURE 3 | (A) Sparseness (% of zero entries) in the time series (left) and connectivity matrices (right) derived using the point-process for a range of thresholds. **(B)** Cumulative file size (in bytes) of fMRI time series (left) and pair-wise connectivity matrices (right) derived using linear correlations (from the full data) and co-activations (from the point-process with threshold equal to 1).



captured equally well by both methods, as well as by the peak-based point-process. In particular, since fixed epoch lengths were used (250 volumes) results were reproduced with and without normalization of connectivity matrices as derived from the point-process. This similitude can also be seen in **Figure 4B**, in which a joint 3D rendering of both maps shows their spatial agreement. The main plots in **Figure 4C** show node strength values at all voxels computed using the point-process method (entries of \tilde{C}_{ij}) vs. those computed using linear correlations (entries of R_{ij}). The functional dependency was clearly monotonously increasing on average, both for wakefulness and sleep, although two individual epochs of sleep displayed an opposite trend.

Strength Maps in Young vs. Older Subjects

We then studied changes in node strength in the 1000 Functional Connectomes dataset, in particular, we compared a group of subjects younger than 20 years with an older group of subjects older than 40 years. Results can be found in **Figure 5A**. For both methods an increase of (normalized) functional connectivity strength in the older group was observed, comprising a network of regions that included the right parietal cortex, inferior frontal cortex, insula, and the precentral and postcentral gyrus.

Driven by the asymmetry observed in the strength differences between age groups, and by the proposal that the right hemisphere shows accelerated functional decline with aging (Dolcos et al., 2002), we applied linear correlations and the point-process analysis to quantify interhemispheric or homotopic connectivity between groups and compared the respective values. Results are shown in **Figure 5B**. Increased interhemispheric connectivity was observed for the older group of subjects by both methods, comprising areas in the parietal and temporal cortex, as well as in the precentral gyrus.

Finally, an additional calculation was performed to allow for further evaluation of our method. We regressed subject age vs. strength values in two regions of interest extracted from the analysis of young vs. older subjects (right Inferior Parietal Cortex—IPC, right and insular cortex). Strength values were obtained both from connectivity matrices obtained with linear correlations and with the point-process. Results are shown in **Figure 5C**. The plots show a moderate increase in strength with age, which suddenly increased for more mature subjects (age > 40 years approximately). Spearman's rank correlation coefficients were higher for the strength values computed using the point-process.

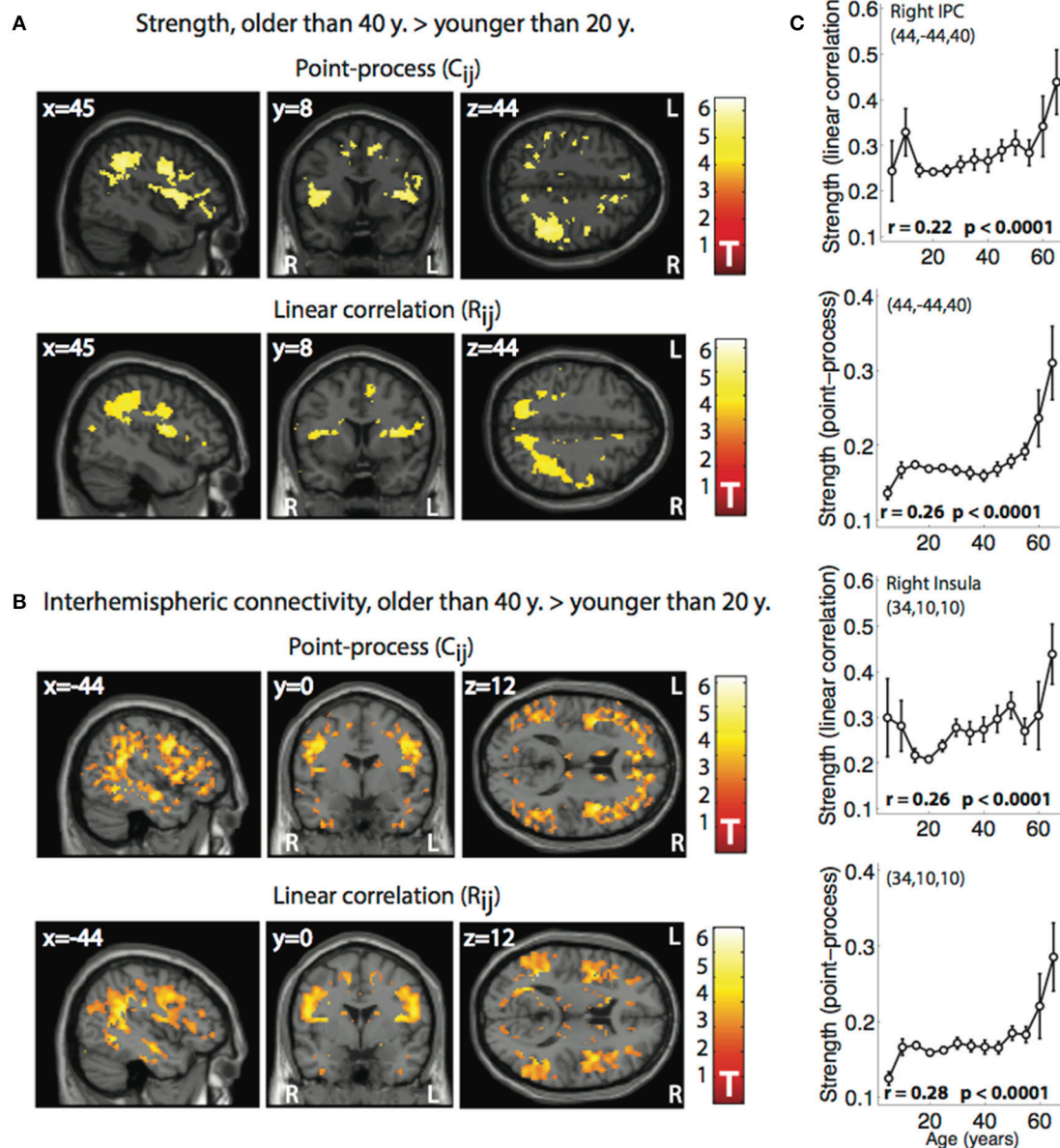


FIGURE 5 | Voxel-wise changes in node strength and interhemispheric connectivity between two age groups, <20 years ($n = 140$) and >40 years ($n = 46$) observed from R_{ij} and from \tilde{C}_{ij} . (A) Spatial maps showing voxels with increased strength in the older group when compared to the younger group, both for the normalized point-process (\tilde{C}_{ij} , top) and for linear correlation (R_{ij} , bottom). Only voxels passing a threshold of $p < 0.05$ (FWE corrected) are shown. **(B)** Spatial maps showing voxels with increased interhemispheric connectivity in the older group when compared with the younger group, both for results obtained from the normalized point-process (\tilde{C}_{ij} , top) and for linear correlation (R_{ij} , bottom). Only clusters passing a threshold of $p < 0.05$ (FWE corrected) are shown. **(C)** Plots subject of age (in years) vs. strength values (derived from linear correlations and the normalized point-process) extracted from two regions of interest (right Inferior Parietal Cortex—IPC, and right insular cortex; mean \pm SEM). An almost monotonous (but clearly non-linear) relationship between age and network centrality is observed.

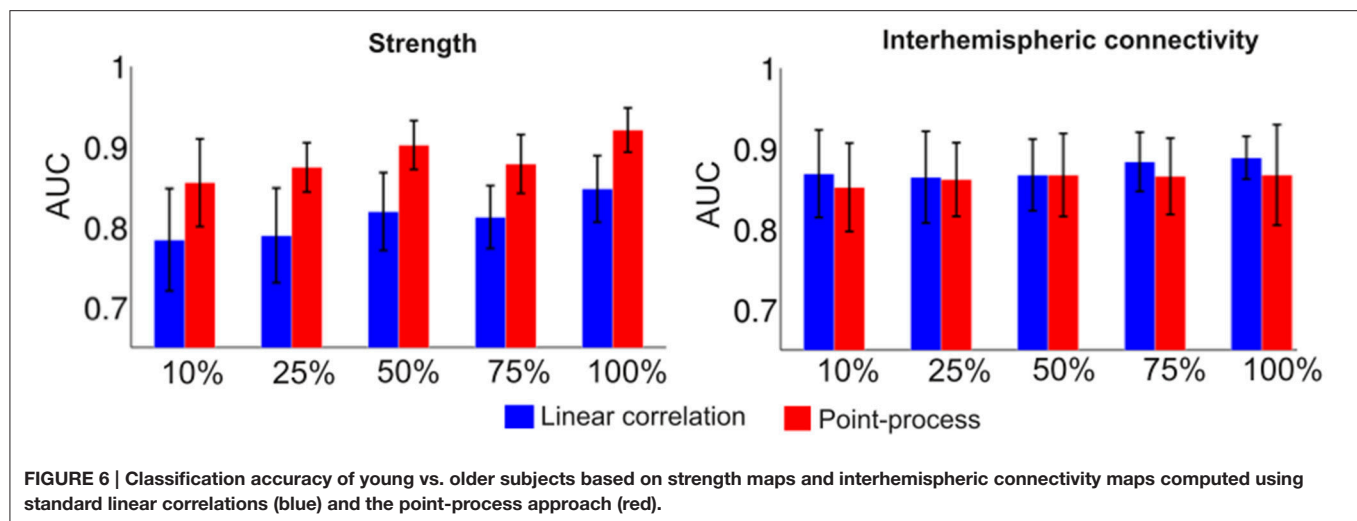
Classification of Young vs. Older Subjects

We implemented the classifier described in the methods to investigate how accurately subjects could be classified by age using strength and interhemispheric connectivity maps, computed with both linear correlations and normalized point-process co-activations. Results are presented in **Figure 6**. We observed similar classification accuracy for the computation based on inter-hemispheric connectivity, and

higher classification accuracy for point-process co-activations vs. linear correlations for the computation based on strength maps.

DISCUSSION

We are witnessing in recent times how neuroscience, and in particular neuroimaging, is moving at a fast pace toward the accumulation and analysis of very large volumes of data. A



number of international collaborations is aiming to break new ground in the scale and speed of data collection, including the 1000 Functional Connectomes Project (Biswal et al., 2010), the NIH BRAIN Initiative (Insel et al., 2013), as well as the Human Connectome Project (Van Essen et al., 2013). These studies span hundreds of subjects scanned at high temporal resolution, resulting in very large datasets. Exploratory analyses of this data may thus benefit from biologically principled dimensionality reduction.

While it is obvious that having large volumes of data reduces the negative effect of noise, artifacts and the relative importance of the mathematical models employed to analyze it [a position eloquently defended by Halevy et al. (2009) in their seminal article “The Unreasonable Effectiveness of Data”], it is also true that the handling of redundant data might be inefficient, both from a computational perspective and in terms of distinguishing the real contributors to the signal from sources of noise. In this line of thought, we have shown that the introduction of a sparse representation of fMRI datasets can reproduce findings obtained from full time series while keeping on the order of 1% of the original data. With respect to vulnerability to noise, sudden head movements can induce spurious points in the process, however, these can be identified from the realignment parameters and erased, following the strategy of scan censoring (Siegel et al., 2014) but eliminating single points (instead of continuous segments of data) from the analyses (see Tagliazucchi et al., 2014 for an application). A consequence of defining the point-process based on high amplitude excursions of the signal is that the impact of physiological noise sources affecting low amplitude fluctuations (Cordes et al., 2002) will be lessened.

Sleep Validation Dataset: Loss of Connectivity in the Thalamus, Frontal, Midline, and Auditory Cortices

We validated our method by first computing correlation between connectivity matrices as obtained by both methods over > 1000 subjects in the Functional Connectomes dataset, as well

as by comparing voxel-wise network strength (a measure of centrality computed from the voxel-wise network of functional connections) between wakefulness and deep sleep and between two age groups extracted from the 1000 Functional Connectomes dataset. In this latter dataset we also obtained the distribution of voxel-wise inter-hemispheric connectivity. The maps of altered network strength in deep sleep and the age-dependent effect observed in the 1000 Functional Connectomes dataset are of biological relevance themselves, as we are not aware of prior reports of these results. Deep sleep resulted in a loss of connectivity across all voxels located in frontal and cingulate cortices, as well as in the primary auditory cortex (Heschl’s gyrus) and the thalamus. These are plausible correlates of reduced awareness (frontal and cingulate cortex) and loss of sensory engagement with the environment (primary auditory cortex and thalamus) resulting in increased arousal thresholds (Tagliazucchi et al., 2013).

Age Groups Validation Dataset: Increased Connectivity with Age in Inferior Parietal and (Pre-)Frontal Cortices

With respect to the two different age groups extracted from the 1000 Functional Connectomes database, regions central to working memory processes (inferior parietal and frontal cortices, prefrontal cortex) showed “over-connectivity” in the older group of subjects. The meaning of this result is less clear, especially in the light of reports showing an inverse relationship between seed-based functional connectivity and age (Sambataro et al., 2010). However, voxel-based strength maps do not require any *a priori* anatomical hypotheses (i.e., seed selection) and thus might be capable of capturing more global changes in connectivity as opposed to the aforementioned approach. Interestingly, changes in the node strength values were mostly located in the right hemisphere. It has been noted by Dolcos et al. (2002) that the right hemisphere shows a more marked decline with aging, a fact supported so far by evidence from working memory neuroimaging experiments. The changes observed

by the authors were hypothesized to be of compensatory origin, which is compatible with the outcome of our analyses (increased overall connectivity in the right hemisphere of older subjects). Prompted by this observation, we also found differences in interhemispheric connectivity located in a set of regions overlapping with those involved with changes in node strength.

Why Few Points Are Sufficient to Reproduce Functional Connectomes

It is worthwhile discussing the reasons underlying the effectiveness of our approach, since it might be surprising that a small fraction of the data suffices to capture all bivariate relationships between BOLD signals (functional connectome) without sacrificing (and even enhancing) classification accuracy.

From a signal processing perspective the answer is relatively straightforward: keeping large amplitude events can increase the signal-to-noise ratio, since it discards low-amplitude activity containing a larger noise component. This non-linear filtering selectively amplifies the importance of those time points at which the signal amplitude becomes relatively large and therefore the signal-to-noise ratio increases. Physiological artifacts have been shown to affect BOLD signals at low frequencies and low amplitudes (Cordes et al., 2002) and signals measured in white matter and cerebrospinal fluid (which do not reflect activity of neural origin and are commonly employed as proxies for physiological confound time series) present smaller amplitude fluctuations compared to those in gray matter (see for instance Tagliazucchi et al., 2013). This situation can result in selective down-weighting of physiological noise when only large-amplitude excursions of the signals are considered.

From a biological point of view, the challenge is to understand why the fMRI time series can be effectively represented as a train of discrete impulses, a view of BOLD time series also supported by studies performing blind de-convolution of spontaneous activity (Petridou et al., 2013). Electrophysiological experiments reveal that Local Field Potentials (LFP) are spatio-temporally distributed as power law avalanches (Beggs and Plenz, 2003): most frequently, spontaneous LFP increases span a limited spatial area, however, at certain (discrete) points in time, LFP might extend up to the size of the tissue under study (an event termed avalanche). If LFP avalanches are, indeed, distributed following a scale-free power law, then macroscopic events (i.e., in the centimeter scale) should be observed, which would be sufficient to elicit a measurable hemodynamic response (considering the correlation observed between LFP and BOLD signals, see Logothetis et al., 2001). Indeed, spatio-temporal avalanches of activity can also be observed with fMRI, following the same statistical laws as the electrophysiological avalanches (Tagliazucchi et al., 2012a). Large amplitude macroscopic LFP increases were reported in the monkey cortex (Thiagarajan et al., 2010) and termed coherence potentials. These large-scale events are also stereotypical (in the words of the authors, much like action potentials at the single-cell level) and thus fulfill all the theoretical requirements for the electrophysiological

underpinnings of the events in the spatio-temporal fMRI point-process.

Contributors to the Resting State fMRI Signal

One of the main limitations of fMRI compared to other non-invasive neuroimaging techniques (EEG, MEG) is its limited temporal resolution. This limitation not only stems from the relatively slow acquisition of whole-brain volumes (i.e., long TRs, in the order of seconds) but also from the coupling between neural activity and the signal measured by fMRI. This coupling blurs temporally localized activity into a temporally extended response (given by the HRF). Therefore, improvement in fMRI sampling rates will only result in a better-sampled HRF, with no gain in the measurement of underlying neural activity, unless the distortion caused by the HRF can be inverted.

Our results suggest that the fMRI resting state signal comprises a temporal succession of well-localized events. The identification of these events has been shown to match a formal de-convolution of fMRI time series (Tagliazucchi et al., 2012a; Petridou et al., 2013). This inversion of the HRF blurring can allow to capitalize on improvements in fMRI acquisition rates. While the contributors to the task-evoked fMRI signal have been thoroughly investigated, this remains to be done in the context of spontaneous brain activity; the possibility of reducing resting state fMRI signals to a few high-amplitude events and still estimate all pair-wise interactions represents an important first step in this direction, and suggests a focus for future studies on the electrophysiological basis of spontaneous fMRI fluctuations.

Caveats and Limitations

Generally, this procedure should yield equivalent results for any dataset in which high amplitude events do not arise spuriously as artifacts and represent important information in the data. From a neurophysiological perspective, the fulfillment of these conditions has been already demonstrated for BOLD time series by means of inverting the Hemodynamic Response Function (HRF) convolution of neuronal sources (de-convolution). As discussed in the previous sections, LFP giving rise to metabolic changes reflected in the BOLD signal are temporally cluttered into avalanches of activity (Beggs and Plenz, 2003; Tagliazucchi et al., 2012a; Shriki et al., 2013), presumably underlying the high information content of BOLD signal high amplitude events.

The main drawbacks of the proposed method are: (1) the non-linear relationship between linear correlation and its estimated value using the point-process (i.e., point-process co-activation, **Figure 2C**) and (2) the slowing down of the computation time when following the normalization given by Equation (6), unless properly optimized. With respect to the first concern, while not linear, the relationship is clearly monotonic and by extracting its functional form, connectivity estimated using the point-process can be properly normalized to have a linear co-variation with standard functional connectivity. This non-linear shape can be explained by the dismissal of low amplitude events in the point-process and their associated contributions to linear correlations. Therefore, correlations can increase faster than point-process co-activations, giving rise to the convex shape seen in **Figure 2A**,

right panel. The second concern (normalization) does not affect the results unless performing comparisons between time series of different length, thus having a different number of points. Normalizing by the length of the time series offers a solution to this issue.

Related Findings

Given the relative novelty of the present approach, caution should be exercised concerning the interpretation of the results to avoid making exaggerated claims. Nevertheless, it is encouraging and reassuring to see a body of publications consistent with the main idea of the present paper. Indeed, since the first observation (Tagliazucchi et al., 2012a) that the timing of high-activity events in BOLD signals allows the reconstruction of major RSN, different research groups have reproduced and built on this result (Davis et al., 2013; Liu and Duyn, 2013; Liu et al., 2013; Amico et al., 2014; Jiang et al., 2014; Li et al., 2014). The analysis of spontaneous voxel co-activation is a natural continuation of functional connectivity studies: instead of asking whether two voxels are engaged in synchronized fluctuations over a relatively long period of time, the question is shifted to whether two voxels become jointly activated (i.e., present high activity above their baseline levels) and what are the timings and properties of these co-activations. Interestingly, it has been shown that co-activation patterns contain additional information not available to standard functional connectivity analyses (Liu et al., 2013) and has also been used to characterize the dynamics of different brain states (Amico et al., 2014; Chen et al., 2015). In the present report we show that the spatio-temporal point-process extracted from whole-brain BOLD signals suffices to estimate all pairs of functional connections (i.e., the functional connectomes) with reasonable accuracy (as demonstrated by its usefulness to capture differences in connectivity between brain states/groups of subjects) with a very small fraction of the data

(on the order of 1%), and thus can be taken as an equivalent (but sparser) representation of the data. We believe these results should prompt an in-depth exploration of high amplitude events in BOLD time series, in particular, their neural correlates and potential relationship to LFP neural avalanches, a signature of self-organized criticality in the human brain (Chialvo, 2010).

In conclusion, as fMRI datasets grow larger, tools to rapidly store, process, and explore them become increasingly valuable. The present report validates a strategy defining a sparse representation of these complex four-dimensional datasets, which keeps only the timing of large BOLD events and thus allows for reasonable fMRI compression. This technique both empowers neuroimaging collaborative projects aimed at gathering and understanding vast amounts of data, and suggests a temporally intermittent organization for brain hemodynamic activity, likely reflecting discrete electrophysiological events spreading throughout the cerebral cortex. Vice versa, if we assume that the sub-threshold BOLD activity is not mere noise nor redundant, this reminds us that with functional connectivity analyses we take but a peek through a keyhole onto the wealth of brain function.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct, and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

Work supported by CONICET (Argentina) and LOEWE Neuronale Koordination Forschungsschwerpunkt Frankfurt—NeFF (Germany). We thank Astrid Morzelewski for data acquisition and sleep scoring together with Kolja Jahnke and Sandra Anti, Ralf Deichmann, and Steffen Volz for extensive MRI support and our volunteers for participation in the study.

REFERENCES

- Achard, S., Delon-Martin, C., Vértes, P. E., Renard, F., Schenck, M., Schneider, F., et al. (2012). Hubs of brain functional networks are radically reorganized in comatose patients. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20608–20613. doi: 10.1073/pnas.1208933109
- Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R., and Lemieux, L. (1998). Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. *Neuroimage* 8, 229–239. doi: 10.1006/nimg.1998.0361
- American Academy of Sleep Medicine (2007). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*, Darien, CT.
- Amico, E., Gomez, F., Di Perri, C., Vanhaudenhuyse, A., Lesenfants, D., Boveroux, P., et al. (2014). Posterior cingulate cortex-related co-activation patterns: a resting state fMRI study in propofol-induced loss of consciousness. *PLoS ONE* 9:e100012. doi: 10.1371/journal.pone.0100012
- Barthelemy, M., Barrat, A., Pastor-Satorras, R., and Vespignani, A. (2005). Characterization and modeling of weighted networks. *Phys. A* 346, 34–43. doi: 10.1016/j.physa.2004.08.047
- Beckmann, C. F., DeLuca, M., Devlin, J. T., and Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. Roy. Soc. B Biol. Sci.* 360, 1001–1013. doi: 10.1098/rstb.2005.1634
- Beggs, J. M., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *J. Neurosci.* 23, 11167–11177.
- Biswal, B. B., Mennes, M., Zuo, X. N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734–4739. doi: 10.1073/pnas.0911855107
- Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., et al. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J. Neurosci.* 29, 1860–1873. doi: 10.1523/JNEUROSCI.5062-08.2009
- Chen, J. E., Chang, C., Greicius, M. D., and Glover, G. H. (2015). Introducing co-activation pattern metrics to quantify spontaneous brain network dynamics. *NeuroImage* 111, 476–488. doi: 10.1016/j.neuroimage.2015.01.057
- Chialvo, D. R. (2010). Emergent complex neural dynamics. *Nat. Phys.* 6, 744–750. doi: 10.1038/nphys1803
- Cordes, D., Haughton, V., Carew, J. D., Arfanakis, K., and Maravilla, K. (2002). Hierarchical clustering to measure connectivity in fMRI resting-state data. *Magn. Reson. Imaging* 20, 305–317. doi: 10.1016/S0730-725X(02)00503-9
- Crossley, N. A., Mechelli, A., Scott, J., Carletti, F., Fox, P. T., McGuire, P., et al. (2014). The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* 137, 2382–2395. doi: 10.1093/brain/awu132

- Davis, B., Jovicich, J., Iacovella, V., and Hasson, U. (2013). Functional and developmental significance of amplitude variance asymmetry in the BOLD resting-state signal. *Cereb. Cortex* 24, 1332–1350. doi: 10.1093/cercor/bhs416
- Dolcos, F., Rice, H. J., and Cabeza, R. (2002). Hemispheric asymmetry and aging: right hemisphere decline or asymmetry reduction. *Neurosci. Biobehav. Rev.* 26, 819–825. doi: 10.1016/S0149-7634(02)00068-4
- Eguíluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M., and Apkarian, A. V. (2005). Scale-free brain functional networks. *Phys. Rev. Lett.* 94:018102. doi: 10.1103/physrevlett.94.018102
- Fox, M. D., and Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Front. Sys. Neurosci.* 4:19. doi: 10.3389/fnsys.2010.00019
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., et al. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biol.* 6:e159. doi: 10.1371/journal.pbio.0060159
- Halevy, A., Norvig, P., and Pereira, F. (2009). The unreasonable effectiveness of data. *Intell. Syst. IEEE* 24, 8–12. doi: 10.1109/MIS.2009.36
- Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., and Menon, R. S. (2013). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Hum. Brain. Mapp.* 34, 2154–2177. doi: 10.1002/hbm.22058
- Insel, T. R., Landis, S. C., and Collins, F. S. (2013). The NIH brain initiative. *Science* 340, 687–688. doi: 10.1126/science.1239276
- Jiang, X., Lv, J., Zhu, D., Zhang, T., Hu, X., Guo, L., et al. (2014). “Integrating group-wise functional brain activities via point processes,” in *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on* (Beijing: IEEE), 669–672.
- Kelly, C., Biswal, B. B., Craddock, R. C., Castellanos, F. X., and Milham, M. P. (2012). Characterizing variation in the functional connectome: promise and pitfalls. *Trends. Cogn. Sci.* 16, 181–188. doi: 10.1016/j.tics.2012.02.001
- Li, W., Li, Y., Hu, C., Chen, X., and Dai, H. (2014). Point process analysis in brain Networks of Patients with diabetes. *Neurocomputing* 145, 182–189. doi: 10.1016/j.neucom.2014.05.045
- Liu, X., Chang, C., and Duyn, J. H. (2013). Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns. *Front. Sys. Neurosci.* 7:101. doi: 10.3389/fnsys.2013.00101
- Liu, X., and Duyn, J. H. (2013). Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4392–4397. doi: 10.1073/pnas.1216856110
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157. doi: 10.1038/35084005
- Petridou, N., Gaudes, C. C., Dryden, I. L., Francis, S. T., and Gowland, P. A. (2013). Periods of rest in fMRI contain individual spontaneous events which are related to slowly fluctuating spontaneous activity. *Hum. Brain. Mapp.* 34, 1319–1329. doi: 10.1002/hbm.21513
- Salvador, R., Suckling, J., Coleman, M. R., Pickard, J. D., Menon, D., and Bullmore, E. (2005). Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb. Cortex* 15, 1332–1342. doi: 10.1093/cercor/bhi016
- Sambataro, F., Murty, V. P., Callicott, J. H., Tan, H. Y., Das, S., Weinberger, D. R., et al. (2010). Age-related alterations in default mode network: impact on working memory performance. *Neurobiol. Aging* 31, 839–852. doi: 10.1016/j.neurobiolaging.2008.05.022
- Shriki, O., Alstott, J., Carver, F., Holroyd, T., Henson, R. N., Smith, M. L., et al. (2013). Neuronal avalanches in the resting MEG of the human brain. *J. Neurosci.* 33, 7079–7090. doi: 10.1523/JNEUROSCI.4286-12.2013
- Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., et al. (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Hum. Brain Mapp.* 35, 1981–1996. doi: 10.1002/hbm.22307
- Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., et al. (2013). Resting-state fMRI in the human connectome project. *Neuroimage* 80, 144–168. doi: 10.1016/j.neuroimage.2013.05.039
- Sporns, O. (2011). The human connectome: a complex network. *Ann. N.Y. Acad. Sci.* 1224, 109–125. doi: 10.1111/j.1749-6632.2010.05888.x
- Sporns, O., Chialvo, D. R., Kaiser, M., and Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends. Cogn. Sci.* 8, 418–425. doi: 10.1016/j.tics.2004.07.008
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., and Chialvo, D. R. (2012a). Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Front. Physiol.* 3:15. doi: 10.3389/fphys.2012.00015
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., Montoya, P., and Chialvo, D. R. (2011). Spontaneous BOLD event triggered averages for estimating functional connectivity at resting state. *Neurosci. Lett.* 488, 158–163. doi: 10.1016/j.neulet.2010.11.020
- Tagliazucchi, E., Carhart-Harris, R., Leech, R., Nutt, D., and Chialvo, D. R. (2014). Enhanced repertoire of brain dynamical states during the psychedelic experience. *Hum. Brain Mapp.* 35, 5442–5456. doi: 10.1002/hbm.22562
- Tagliazucchi, E., and Laufs, H. (2014). Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep. *Neuron* 82, 695–708. doi: 10.1016/j.neuron.2014.03.020
- Tagliazucchi, E., von Wegner, F., Morzelewski, A., Borisov, S., Jahnke, K., and Laufs, H. (2012c). Automatic sleep staging using fMRI functional connectivity data. *Neuroimage* 63, 63–72. doi: 10.1016/j.neuroimage.2012.06.036
- Tagliazucchi, E., von Wegner, F., Morzelewski, A., Brodbeck, V., Jahnke, K., and Laufs, H. (2013). Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15419–15424. doi: 10.1073/pnas.1312848110
- Tagliazucchi, E., Von Wegner, F., Morzelewski, A., Brodbeck, V., and Laufs, H. (2012b). Dynamic BOLD functional connectivity in humans and its electrophysiological correlates. *Front. Hum. Neurosci.* 6:339. doi: 10.3389/fnhum.2012.00339
- Thiagarajan, T. C., Lebedev, M. A., Nicolelis, M. A., and Plenz, D. (2010). Coherence potentials: loss-less, all-or-none network events in the cortex. *PLoS Biol.* 8:e1000278. doi: 10.1371/journal.pbio.1000278
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., and Ugurbil, K. (2013). The WU-Minn human connectome project: an overview. *Neuroimage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041
- Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., et al. (2009). Parcellation-dependent small-world brain functional networks: a resting-state fMRI study. *Hum. Brain Mapp.* 30, 1511–1523. doi: 10.1002/hbm.20623
- Wu, G. R., Liao, W., Stramaglia, S., Ding, J. R., Chen, H., Marinazzo, D. (2013). A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. *Med. Image Anal.* 17, 365–374. doi: 10.1016/j.media.2013.01.003
- Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., et al. (2010). Whole-brain anatomical networks: does the choice of nodes matter?. *Neuroimage* 50, 970–983. doi: 10.1016/j.neuroimage.2009.12.027

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer JV and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Tagliazucchi, Siniatchkin, Laufs and Chialvo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Mixed Effects Models for Resampled Network Statistics Improves Statistical Power to Find Differences in Multi-Subject Functional Connectivity

Manjari Narayan^{1*} and Genevera I. Allen^{1,2,3}

¹ Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA, ² Department of Statistics, Rice University, Houston, TX, USA, ³ Jan and Dan Duncan Neurological Research Institute and Department of Pediatrics-Neurology at Baylor College of Medicine, Houston, TX, USA

OPEN ACCESS

Edited by:

Bertrand Thirion,
Institut National de Recherche en
Informatique et Automatique, France

Reviewed by:

Joshua T. Vogelstein,
Johns Hopkins University, USA
Felix Carbonell,
Biospective Inc., Canada

*Correspondence:

Manjari Narayan
manjari.narayan@gmail.com

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 23 September 2015

Accepted: 07 March 2016

Published: 12 April 2016

Citation:

Narayan M and Allen GI (2016) Mixed
Effects Models for Resampled
Network Statistics Improves Statistical
Power to Find Differences in
Multi-Subject Functional Connectivity.
Front. Neurosci. 10:108.
doi: 10.3389/fnins.2016.00108

Many complex brain disorders, such as autism spectrum disorders, exhibit a wide range of symptoms and disability. To understand how brain communication is impaired in such conditions, functional connectivity studies seek to understand individual differences in brain network structure in terms of covariates that measure symptom severity. In practice, however, functional connectivity is not observed but estimated from complex and noisy neural activity measurements. Imperfect subject network estimates can compromise subsequent efforts to detect covariate effects on network structure. We address this problem in the case of Gaussian graphical models of functional connectivity, by proposing novel two-level models that treat both subject level networks and population level covariate effects as unknown parameters. To account for imperfectly estimated subject level networks when fitting these models, we propose two related approaches— R^2 based on resampling and random effects test statistics, and R^3 that additionally employs random adaptive penalization. Simulation studies using realistic graph structures reveal that R^2 and R^3 have superior statistical power to detect covariate effects compared to existing approaches, particularly when the number of within subject observations is comparable to the size of subject networks. Using our novel models and methods to study parts of the ABIDE dataset, we find evidence of hypoconnectivity associated with symptom severity in autism spectrum disorders, in frontoparietal and limbic systems as well as in anterior and posterior cingulate cortices.

Keywords: functional connectivity, Gaussian graphical models, Markov networks, covariates, mixed effects models, resampling methods, lasso, network statistics

1. INTRODUCTION

One of the goals of neuroimaging studies of intrinsic or “resting state” brain activity, is to discover specific and stable imaging based biomarkers or phenotypes of neuropsychiatric and neurological disorders. Typically, resting state studies seek to infer *functional connectivity* or functional relationships between distinct brain regions from observed neurophysiological activity. Advances in resting state studies using fMRI (Menon, 2011; Bullmore, 2012; Craddock et al., 2013;

Smith et al., 2013) suggest that functional connectivity could yield neuroimaging biomarkers for diagnosis and personalized treatment for a wide range of disorders.

For instance, many studies have found differences either in individual functional connections or in overall patterns of connectivity in autism spectrum disorders (Di Martino et al., 2014a; Uddin, 2014), Alzheimer's (Buckner et al., 2009; Tam et al., 2014), Depression (Tao et al., 2013; Lui et al., 2014; Kaiser et al., 2015) and others (Meda et al., 2012; van den Heuvel et al., 2013; Palaniyappan et al., 2013). However, simple group level differences between two distinct samples are challenging to interpret in many disorders. Autism, for example, is a diagnostic label that masks many diverse clinical symptoms (Lenroot and Yeung, 2013; Insel, 2014). Thus, the biological relevance of group level differences in network structure between Autism and healthy populations is unclear for individual subjects. One solution to find more meaningful differences in network structure is to study whether behavioral and affective symptoms measured by cognitive scores are associated with variations in individual functional networks. This paper offers a novel and rigorous statistical framework to find and test such covariate effects on functional connectivity metrics, when functional connectivity is defined using Gaussian graphical models.

Functional connectivity refers to latent relationships that cannot be directly observed via any modality of functional neuroimaging. Instead, it must be estimated from observations of neurophysiological activity. In fMRI studies, we first observe changes in the BOLD response over time either across thousands of voxels or over hundreds of brain regions, defined anatomically or functionally. Then depending on the specific statistical definition for functional connectivity, we estimate a functional connectivity network per subject using within-subject BOLD observations. For example, in a pairwise correlation model of functional connectivity, if the mean time-series of two brain regions are correlated then they are functionally connected. Thus, one popular approach to estimate functional connectivity is to compute sample correlations between every pair of brain regions. An increasingly popular alternative is to use Gaussian graphical models (GGMs) based on partial correlations to define functional connectivity. Here, if two brain regions are partially correlated, that is if the mean time-series of two brain regions remain correlated after regressing out the time-series of other brain regions, then they are functionally connected. For multivariate normal data, a zero partial correlation between two brain regions is equivalent to independence between the activity of two brain regions conditional on the activity of all other intermediate brain regions. Thus, GGMs eliminate indirect connections between regions provided by pairwise correlations and are increasingly popular in neuroimaging (Marrelec et al., 2006; Smith et al., 2011; Varoquaux et al., 2012; Craddock et al., 2013). Consequently, employing GGMs for functional connectivity enables us to discover network differences that implicate nodes and edges directly involved in producing clinical symptoms and provide stronger insights into network structures truly involved in the disease mechanism. For the rest of this paper, we define functional connectivity in terms of GGMs and

discuss approaches to conducting inference on network metrics for such network models.

The functional connectivity of a single experimental unit or subject is rarely the final object of interest. Rather, most neuroimaging studies (Bullmore, 2012; Bullmore and Sporns, 2012; Zuo et al., 2012) are interested in identifying network biomarkers, or broader patterns of functional connectivity shared across individuals who belong to some distinct population or display some clinical phenotype. A popular approach (Bullmore and Sporns, 2009) to find such network biomarkers is through topological properties of network structure. Common properties or metrics either measure specialization of network components into functionally homogenous modules, or measure how influential brain regions integrate information across distinct network components. However, recall that functional connectivity in individual subjects is unknown and unobserved. Consequently, many multi-subject fMRI studies first estimate functional connectivity for every subject and then assuming these subject networks are fixed and known, compute topological metrics of these networks using the Brain Connectivity Toolbox (Rubinov and Sporns, 2010). Finally, they compare and contrast these estimated networks or estimated network metrics to infer group level network characteristics. Typical neuroimaging studies that seek to detect covariate effects on network structure (Warren et al., 2014; Hahamy et al., 2015) conduct a single level regression with network metrics as the response and cognitive scores as the covariate, and subsequently use standard *F*-tests for covariate testing. New methods to conduct such network inference either emphasize novel topological metrics (van den Heuvel and Sporns, 2011; Alexander-Bloch et al., 2012) or novel approaches to study covariate effects for known networks for complex experimental designs with longitudinal observations or multiple experimental conditions (Simpson et al., 2013; Ginestet et al., 2014; Kim et al., 2014). However, these existing approaches assume estimated functional networks are perfectly known quantities. In contrast, we seek to explicitly investigate the consequences of using estimated, and often imperfectly estimated, functional networks and their corresponding network metrics on subsequent inference for covariate effects.

Before considering the consequences of using estimated networks, one might ask why individual network estimates might be unreliable to begin with. Statistical theory informs us that estimated networks can be unreliable in two possible ways. First, high dimensional networks with a large number of nodes estimated from a limited number of fMRI observations in a session possess substantial sampling variability (Bickel and Levina, 2008; Rothman et al., 2008; Ravikumar et al., 2011; Narayan et al., 2015). Second, when assuming sparsity in the network structure in the form of thresholded or penalized network estimates to overcome high dimensionality, we often obtain biased network estimates in the form of false positive or false negative edges (Ravikumar et al., 2011). Such errors in estimating networks are particularly exacerbated (Narayan et al., 2015) when networks are well connected with modest degrees, as is the case in neuroimaging. Additionally, empirical evidence from neuroimaging studies also suggest that sample correlation based estimates of individual resting state networks

are unreliable. For instance test re-test studies (Shehzad et al., 2009; Van Dijk et al., 2010; Braun et al., 2012) that measure inter-session agreement of estimated functional networks within the same subject find that sample intra-class correlations vary between 0.3 and 0.7, indicating non-negligible within subject variability. While we expect many sources of variation contribute to such inter-session variability within a subject including natural variations due to differences in internal cognitive states, recent work by Birn et al. (2013); Hacker et al. (2013); Laumann et al. (2015) suggests that sampling variability due to limited fMRI measurements play a significant role. These studies find that increasing the length of typical fMRI sessions from 5–10 min to 25 min substantially improves inter-session agreement of functional networks. Given the accumulating theoretical and empirical evidence of these methodological limitations, we assume that obtaining perfect estimates of individual networks is unlikely in typical fMRI studies. Instead, we seek to highlight the importance of accounting for imperfect estimates of functional networks in subsequent inferential analyses.

Failure to account for errors in estimating statistical networks reduces both generalizability and reproducibility of functional connectivity studies. Statistical tests that compare functional networks but do not account for potentially unreliable network estimates lack either statistical power or type I error control or both. For instance, Narayan and Allen (2013); Narayan et al. (2015) investigate the impact of using estimated networks when testing for two-sample differences in edge presence or absence between groups. When individual subject graphical models cannot be estimated perfectly, Narayan et al. (2015) show that standard two-sample test statistics are both biased and overoptimistic, resulting in poor statistical power and type I error control. Though this paper is similar in spirit to previous work (Narayan et al., 2015) in emphasizing the adverse effects of using estimated networks to study differences in functional connectivity, the unique contribution of this work are as follows: (1) Whereas previous work considered simple two-sample tests, we consider general covariate effects (that include both binary and continuous covariates) to link symptom severity to individual variations in functional connectivity. (2) We propose methods relevant to network metrics beyond the edge level. Finally, we provide empirical results such as statistical power analyses that offer greater practical guidance on choosing sample size and planning data analysis for future studies.

The paper is organized as follows. In Section 2 we provide new statistical models that explicitly link subject level neurophysiological data to population level covariate effects for network metrics of interest and provide new statistical algorithms and test statistics using resampling and random penalization for testing covariate effects. While the models and methods we propose can detect covariate effects on many well behaved network metrics (Balachandran et al., 2013) at the edge level (Tomson et al., 2013), node level (Buckner et al., 2009; Zuo et al., 2012) and community level (Alexander-Bloch et al., 2012; Tomson et al., 2013), we investigate the benefits of our methods to discover covariate effects on connection density. Using realistic simulations of graph structure for GGMs in Section 3, we demonstrate our proposed resampling framework substantially

improves statistical power over existing approaches, particularly for typical sample size regimes in fMRI studies. Finally, in Section 4 we demonstrate that our proposed methods can detect biologically relevant signals in a resting state fMRI dataset for autism spectrum disorders.

2. MODELS AND METHODS

We seek new methods to detect covariate effects when populations of functional networks are unknown. To achieve this, we first need statistical models that describe how each measurement of brain activity denoted by $y_j^{(i)}$ arises from unknown functional brain network with p nodes in the i^{th} subject and how individual variations in a population of brain networks are related to some population level mean. Thus, for any network model and any network metric under investigation, we propose the following general two-level models to investigate covariate effects in functional connectivity. In subsequent sections, we provide specific instances of these models investigated in this paper.

Subject Level: $y_j^{(i)} \stackrel{iid}{\sim} \mathcal{N}_p(0, \Sigma^{(i)})$ and

Population Level: $u(\text{Network}_{(i)}) \stackrel{iid}{\sim} \mathbb{P}_{\mu^{(i)}, v^2}$ (1)

where $\Sigma^{(i)}$ is the covariance, Network_i is an adjacency matrix derived from either the covariance, the inverse covariance $\Theta = (\Sigma^{(i)})^{-1}$ or their correlational counterparts and $u(\cdot)$ denotes some network metric over the brain network. In this paper, we assume the individual measurements of brain activity at the subject level follow a multivariate normal distribution. At the population level, we assume that the effect of covariates on the network metrics follows a generalized linear model (Searle et al., 2009) where the mean and variance of the relevant continuous or discrete probability distribution, \mathbb{P} , for the network metric of interest is given by $\mu^{(i)}$ and v^2 .

Suppose that we denote any network metric in the i^{th} subject as $u^{(i)}$ and the vector of network metrics as $\mathbf{u} = (u^{(1)}, \dots, u^{(n)})$, then the population mean is given by $\mu = \mathbb{E}(\mathbf{u})$ and population variance is given by $\text{Var}(u^{(i)}) = v^2$. Then the generalized linear model for the population mean is given by

$$g(\mu) = \mathbf{X}\beta + \mathbf{Z}\gamma \quad (2)$$

Here $g(\mu)$ is a link function either reduces to $g(\mu) = \mu$ in linear models, or takes other forms such as the logit function for non-linear models; \mathbf{X} is the $n \times (q + 1)$ matrix of the intercept and q covariates of interest with corresponding coefficients $\beta = (\beta_0, \beta_1, \dots, \beta_q)$ while \mathbf{Z} is the $n \times r$ matrix of nuisance covariates and corresponding regression coefficients γ . X_i and Z_i denote the q dimensional explanatory covariate and r dimensional nuisance covariate for the i^{th} subject, respectively.

In this paper, we seek to test the hypothesis that explanatory covariates have a statistically significant covariate effect on network metrics. Here $\beta_{\setminus 0}$ denotes the coefficients for

explanatory covariates. Thus, the null \mathcal{H}_0 and alternative hypothesis \mathcal{H}_1 are

$$\mathcal{H}_0 : \beta_{\setminus 0} = 0, \quad \mathcal{H}_1 : \beta_{\setminus 0} \neq 0 \quad (3)$$

This section is organized as follows—In Section 2.1, we specifically employ Gaussian graphical model of functional connectivity at the subject level and investigate covariate effects using linear models for density based network metrics for the population level. Standard statistical analyses in neuroimaging studies estimate each level of these two level models separately. Thus, such approaches first estimate functional connectivity networks by fitting subject level models. However, they assume individual subject networks and their metrics are known when they fit the population level model and conduct inference on covariate effects. In Section 2.2 we discuss how such statistical procedures that assume functional connectivity networks are known lose statistical power to detect covariate effects. To address this problem, we introduce two related methods that utilize resampling, random adaptive penalization, and random effects that we call, R^2 and R^3 in Section 2.3. These methods ameliorate potential biases and sampling variability in estimated network metrics, thus improving statistical power to detect covariate effects.

2.1. Two Level Models for Covariate Effects

We begin by studying the earlier subject level network model in Equation (1) specifically for networks given by Gaussian graphical models. Recall that the p -variate random vector $y_j^{(i)}$ denotes BOLD observations or average BOLD observations within p regions of interest, at the j^{th} time point for the i^{th} subject. We assume $y_j^{(i)}$ has a multivariate normal distribution,

$$y_j^{(i)} \stackrel{iid}{\sim} \mathcal{N}_p(0, (\Theta^{(i)})^{-1}), \quad (4)$$

where the network model of interest is derived from the inverse covariance or precision matrix $\Theta^{(i)}$, $j = 1, \dots, t$, and $i = 1, \dots, n$. In subsequent sections, we denote the $t \times p$ data matrix of observations by $\mathbf{Y}^i = (y_1^{(i)}, \dots, y_t^{(i)})$ and the random variable associated with each brain region as Y_k . Although fMRI observations are autocorrelated across time and thus dependent (Woolrich et al., 2001; Worsley et al., 2002), we assume that these observations can be made approximately independent via appropriate whitening procedures discussed in our case study in Section 4.

Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ denote a Gaussian graphical model that consists of vertices $\mathcal{V} = \{1, 2, \dots, p\}$ and edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. Here, the presence of an edge $(k, l) \in \mathcal{E}$ implies that the random variables Y_k and Y_l at nodes/vertices k and l are statistically dependent conditional on all the other vertices $\mathcal{V} \setminus \{k, l\}$. For multivariate normal distributions, a non-zero value in the (k, l) entry of the inverse covariance matrix $\Theta^{(i)}$ is equivalent to the conditional independence relationships, $Y_k \perp Y_l | Y_{\mathcal{V} \setminus \{k, l\}}$. Thus, we define functional connectivity networks where edges indicate *direct* relationships between two brain regions using the non-zero

entries of $\Theta^{(i)}$. For a more thorough introduction to graphical models, we refer the reader to Lauritzen (1996).

Following the neuroimaging literature (Bullmore and Sporns, 2009), we consider network metrics to be functions of a binary adjacency matrix. The adjacency matrix of each individual subject network in our model (Equation 4) is given by the support of the inverse covariance matrix $\mathbb{I}\{\Theta^{(i)} \neq 0\}$. Network metrics that measure topological structure of networks are widely used in neuroimaging (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010). While any of these network metrics can be incorporated into our two level models, we have found that many metrics originally proposed when studying a deterministic network are not suitable for covariate testing in the presence of individual variations in a population of networks. Recently, Balachandran et al. (2013) suggests that several discontinuous network metrics which include betweenness centrality, clustering coefficients defined at the node level and potentially many others are not suitable for inference. Thus, this paper focuses on well behaved topological metrics, namely density based metrics. Formally, the density or number of connections in any binary adjacency matrix A is given by $\sum_{k=1}^p \sum_{l=1}^p A_{kl}$. However, rather than defining density over the whole graph, the density can be restricted to a subnetwork (subnetwork density) or over a single node (node density or degree) or simply at the edge level (edge presence). At the node level, density is a simple measure of influence or centrality of a single brain region of interest (Rubinov and Sporns, 2010; Power et al., 2013). At the subnetwork level, density is popularly used (Honey et al., 2007; Bullmore and Sporns, 2009) to measure an excess or deficit of long range connections either within or between groups of brain regions with a distinct functional purpose. While we investigate node and subnetwork density in this paper, alternative network metrics amenable to inference include binary metrics such as edge presence (Meda et al., 2012; Narayan et al., 2015) or co-modularity relationships between nodes (Bassett et al., 2013; Tomson et al., 2013).

2.1.1. Population Model for Network Metrics

As described earlier, given the subject level model and a network metric of interest, we use a generalized linear model in Equation (2) to describe the deterministic relationship between the population mean for the network metrics and various covariates of interest. Depending on whether a network metric is continuous or binary valued, this general linear model takes the form of linear or logistic-linear models.

However, we also require a probability model to describe how a random sample of individual network metrics deviate from the population mean. When the network metric $u^{(i)}$ is continuous valued, the link function in Equation (2) reduces to the identity $g(\mu) = \mu$. For network metrics $u^{(i)}$ such as global, subnetwork or node density, we use the following linear model with normal errors,

$$u^{(i)} \stackrel{iid}{\sim} \mathcal{N}(X_i \beta + Z_i \gamma, \sigma^2) \quad (5)$$

For metrics such as edge presence and co-modularity that take discrete binary values $\{0, 1\}$, a widely used link function (Williams, 1982; Agresti, 2002) for the generalized linear model

(Equation 2) is the logit function. The resulting logistic-linear model takes the following form

$$\mathbb{E}(u^{(i)}) = [1 + \exp(X_i\beta + Z_i\gamma)]^{-1} \quad (6)$$

For the remainder of this paper, we consider normal models for node and subnetwork density.

2.2. Motivation for New Test Statistics

To understand why new statistical methods are necessary to fit our two-level models, consider the our covariate testing problem (Equation 3) for node and subnetwork density. Suppose the subject level networks in Equation (4) and corresponding metrics are known precisely for each subject. In this case, we employ standard least squares estimation with corresponding F -tests for linear regression to test our null hypothesis for covariate effects (Equation 3).

In practice however, not only is the covariate effect β unknown, the underlying graphical model $\Theta^{(i)}$ and the network metric $u^{(i)}$ is also unknown and are all estimated from data. In **Figure 1** we contrast the ideal scenario where the population of networks and corresponding network metrics are exactly known with the practical scenario where these network metrics are estimated from data. (See Section 3.1 for details on how we simulate data.) Applying a standard linear regression to known network metrics reveals an oracle estimate of the covariate effect (blue line). In contrast, when the standard approach described is applied to estimated network metrics (orange line), the size of the covariate effect is substantially reduced. However, by employing the R^3 approach (green line) that we introduce in the next section, we account for errors in estimating networks, thereby improving statistical power.

Two issues arise when we estimate network metrics from data. First, instead of true network metrics, $u^{(i)}$, our estimated network metrics, $\tilde{u}^{(i)}$, are a function of observations $\mathbf{Y}^{(i)}$. Thus, each estimate, $\tilde{u}^{(i)}$, possesses additional sampling variability. However, since we only acquire one network estimate per subject, standard least squares estimation cannot account for this additional variability. Additionally graph selection errors in network estimation potentially bias network metric estimates. Previously, Meinshausen and Bühlmann (2006); Ravikumar et al. (2011); Narayan et al. (2015) show that in finite sample settings where the number of independent observations t within a subject is comparable to the number of nodes p , we expect false positive and false negative edges in network estimates. Such graph selection errors increase with the complexity of the network structure, governed by factors such as the level of sparsity, maximum node degree as well as the location of edges in the network. Since functional connectivity networks are moderately dense and well connected with small world structure (Achard et al., 2006), edges in these networks might be selected incorrectly. Observe that in **Figure 1**, we obtain larger estimates of node and subnetwork density for individual networks where true node or subnetwork densities are small and the reverse for truly large node or subnetwork densities. As a result, individual variation in estimated metrics no longer reflects the true effect of the covariate, resulting in loss of statistical power. For a detailed overview of how selection errors in estimating network structure propagate to group level inferences, we refer the reader to Section 2 of Narayan et al. (2015).

To overcome these obstacles, we use resampling to empirically obtain the sampling variability of estimated network metrics, $\tilde{u}^{(i)}$, and propagate this uncertainty using mixed effects test statistics for the covariate effect $\hat{\beta}$. Moreover, by aggregating

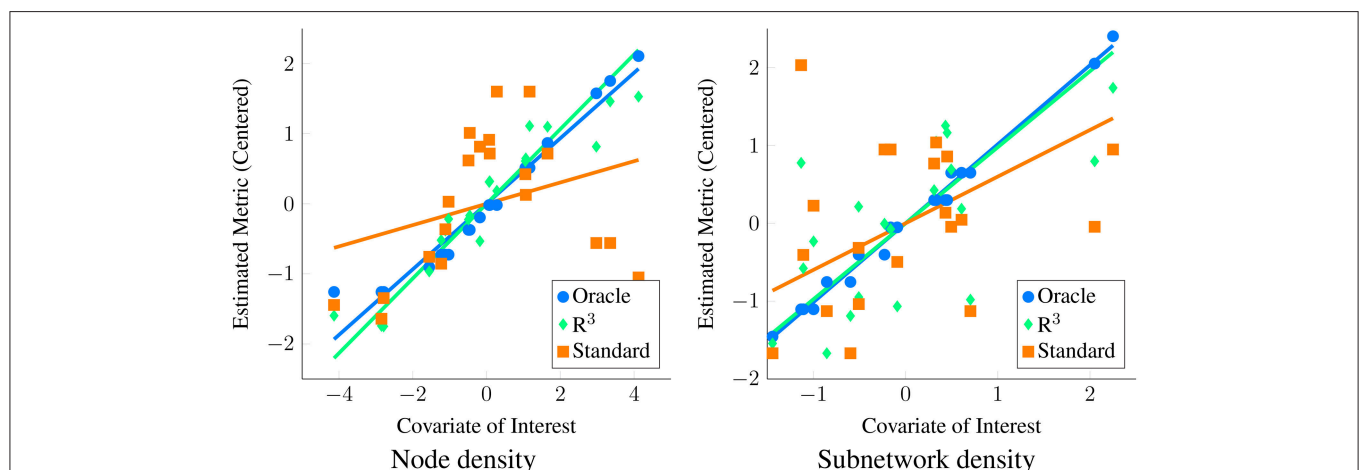


FIGURE 1 | Motivation for new statistical framework R^3 . Here, we simulate covariate effects on the metric of interest, namely the degree centrality or node density (left) and subnetwork density (right) with ($p = 50, n = 20, t = 200$). We illustrate covariate effects in the ideal scenario where network metrics are known perfectly in blue. Unfortunately, in functional connectivity networks, statistical errors in estimating graphical models are inevitable and these propagate to estimates of network metrics. As a result, when we estimate node and subnetwork density for each subject and conduct tests for covariate effects using standard F -tests, we fail to see a clear relationship between metrics and covariate of interest (orange) using linear regression. This loss of statistical power occurs when standard test statistics assume that estimates of density are correct. In contrast, when we account for errors in graph estimation and selection using R^3 test statistics (green), we have greater statistical power to detect covariate effects on density metrics. Algorithmic details of the R^3 approaches can be found in Section 2.

network statistics across resamples and optionally incorporating adaptive penalization techniques, we sufficiently improve network estimates and corresponding network metrics to obtain more accurate estimates of the covariate effects.

2.3. Procedure for Testing Covariate Effects

In order to improve statistical power, we propose a resampling framework that integrates network estimation with inference for fixed covariate effects at the population level. We provide two related procedures to test covariate effects— R^2 that employs resampling (RS) and random effects test statistics (RE), and R^3 that employs resampling (RS), random adaptive penalization (RAP) and random effect test statistics (RE). Intuitively, our algorithm consists of first obtaining initial estimates of the sparsity levels in individual subject networks. Then, to estimate the sampling variability of each subject network empirically, we resample within subject observations and re-estimate the networks of each subject. Additionally, in the case of R^3 we simultaneously apply random adaptive penalties when re-estimating the networks. Network metrics are computed on each of the resampled networks, giving us multiple pseudo-replicates of network metrics per subject. Finally, we model these resampled network statistics using simple mixed effects models to derive test statistics for population level covariate effects. After performing our procedure, one can use well known parametric or non-parametric approaches to obtain p -values and correct for multiplicity of test statistics when necessary. Thus, our resampling framework consists of three components, graph estimation and selection, resampling and optionally RAP, and covariate testing via mixed effects models. We discuss each of these ingredients separately before putting them together in **Algorithm 1**.

2.3.1. Graphical Model Estimation

Many approaches such as sparse regularized regression (Meinshausen and Bühlmann, 2006), sparse penalized maximum likelihood (ML) or the graphical lasso (Yuan and Lin, 2007; Friedman et al., 2008) and others (Cai et al., 2011; Zhou et al., 2011) can be used to estimate $\Theta^{(i)}$ in our subject level model (Equation 4). We use the QuIC solver (Hsieh et al., 2011, 2013) to fit a weighted graphical lasso to obtain estimates of $\Theta^{(i)}$.

$$\hat{\Theta}_{\Lambda^{(i)}}^{(i)}(\mathbf{Y}^{(i)}) = \arg \min_{\Theta > 0} \text{Tr}(\hat{\Sigma}^{(i)} \Theta) - \log \det(\Theta) + \|\Lambda^{(i)} \circ \Theta\|_{1,\text{off}} \quad (7)$$

where $\hat{\Sigma}^{(i)}$ is the empirical sample covariance, $\hat{\Sigma}^{(i)} = \frac{1}{t}(\mathbf{Y}^{(i)\top} \mathbf{Y}^{(i)})$, and \circ denotes the Hadamard dot product. The term $\|\Theta\|_{1,\text{off}} = \sum_{k < l} |\theta_{k,l}|$ is the ℓ_1 penalty on the off-diagonals entries. Since the sample correlation rather than covariance is commonly used in neuroimaging, we employ sample correlation matrix, $\tilde{\Sigma}^{(i)}$. The two are equivalent when $\mathbf{Y}^{(i)}$ has been centered and scaled. Given any estimate of the inverse covariance matrix $\hat{\Theta}^{(i)}$, the estimated adjacency matrix for each subject is thus given by $\mathbb{I}(\hat{\Theta}^{(i)} \neq 0)$ and network statistics can be computed

accordingly. For our R^3 procedure, we employ a symmetric weight matrix of penalties $\Lambda^{(i)}$ obtained by randomly perturbing an initial penalty parameter $\lambda^{(i)}$. For our R^2 this weight matrix $\Lambda^{(i)}$ reduces to a scalar value $\lambda^{(i)}$ for all off-diagonal entries, giving us the standard graphical lasso. In order to estimate these initial penalty parameters $\lambda^{(i)}$, we employ StARS (Liu et al., 2010), a model selection criterion that is asymptotically guaranteed to contain the true network, and works well with neuroimaging data. The β parameter of StARS is set to 0.1 in our work.

2.3.2. Resampling and Random Adaptive Penalization

Since network estimates depend on the underlying observations $\mathbf{Y}^{(i)}$, we employ resampling techniques to estimate the sampling variability of $\tilde{u}^{(i)}$. Recall that estimates of a network metric, $\tilde{u}^{(i)}$, are a function of estimated networks $\mathbb{I}(\hat{\Theta}^{(i)}(\mathbf{Y}^{(i)}) \neq 0)$. Unfortunately, closed form finite sample distributions for sparse penalized estimates of $\hat{\Theta}^{(i)}$ (Berk et al., 2013) as well as sampling distributions of network metrics (Balachandran et al., 2013) are still an emerging area of research. Thus, our problem differs from standard univariate GLM analyses employed in both voxel-wise activation studies and seed-based correlational analysis (Penny et al., 2003; Fox et al., 2006) where closed form asymptotic formulas for sample variance at the subject level are incorporated into the group level analyses. To tackle the issue of unknown sampling variability we build an empirical distribution of network statistics, where we perturb the data by sampling m out of t observations with replacement (bootstrap) (Efron and Tibshirani, 1993) or without replacement (subsampling) (Politis et al., 1999) and re-estimate the network metrics per resample. By aggregating network statistics across resamples within each subject (Breiman, 1996a), we gain the additional benefit of variance reduction (Bühlmann and Yu, 2002) for individual subject metrics. Many variations of resampling techniques exist to handle dependencies (Lahiri, 2013) in spatio-temporal data. Since we assume approximately independent observations, from here on our resampling consists of sampling t out of t observations with replacement.

Recall that our method R^2 is a variant of R^3 , that only involves resampling without random adaptive penalties. Here we obtain a bootstrapped network estimate $\hat{\Theta}^{*(i,b)}$, and a corresponding network metric $\tilde{u}^{*(i,b)}$ in Step 1 of our **Algorithm 1** for each of $B = 100$ resamples. For our alternative procedure, R^3 , however, we not only use resampling, but simultaneously perturb the initial regularization parameters $\lambda^{(i)}$ for every resample. This amounts to solving a weighted graphical lasso to re-estimate the network, where the weights are given by random adaptive penalties. Our motivation to use R^3 is based on previous work in the context of two-sample tests for edge differences. Narayan et al. (2015) show that random penalization significantly improved power over pure resampling to detect differential edges when the networks were moderately dense. Given this result, we sought to investigate the benefits of random penalization for more general network metrics. Intuitively, we anticipate that density based metrics beyond the edge level are immune to some graph selection errors. For instance, when false negatives are compensated by an equal number of false positive edges within the same

node or subnetwork, node or subnetwork density values remain unchanged. However, graph selection errors that do not cancel each other out result in a net increase or decrease in density, thus contributing to loss of power. In these scenarios, we expect R^3 to offer additional statistical power to test covariate effects.

Whereas general network metrics, require global properties of the network structure be preserved, the standard randomized graphical lasso (Meinshausen and Bühlmann, 2010) penalizes every edge randomly such that topological properties of the network could be easily destroyed within each resample. Thus, we seek to randomly perturb selected models in a manner less destructive to network structure. To achieve this, we adaptively penalize (Zhou et al., 2011) entries of $\Theta^{(i)}$. Strongly present edges are more likely to be true edges and should thus be penalized less, whereas weak edges are more likely to be false and should be penalized more. As long as we have a good initial estimate of where the true edges in the network are, we can improve network estimates by adaptively re-estimating the network, while simultaneously using random penalties to account for potential biases in the initial estimates. In order to obtain a reliable initial estimate of network structure, we take advantage of the notion of stability as a measure of confidence popularized by Breiman (1996b); Meinshausen and Bühlmann (2010). Here the stability of an edge within a network across many resamples measures how strongly an edge is present in the network. When an edge belongs to the true network with high stability we randomly decrease the associated penalty by a constant κ . Conversely, we randomly increase the penalty by κ for an edge with low stability. Similar to Narayan et al. (2015), we fix the constant κ to $0.25\lambda_{max}^{(i)}$. Here $\lambda_{max}^{(i)}$ is the regularization parameter that results in the all zero graph for a subject. We call this approach random adaptive penalization (RAP) as it builds on the previous random penalization approach of Narayan et al. (2015) but adaptively perturbs the regularization parameters using initial stability scores along the lines of the random lasso (Wang et al., 2011).

Since, random adaptive penalization depends on an initial estimate of the stability of every edge in the network, we take advantage of the basic resampling step in **Algorithm 1** to obtain a stability score matrix $\hat{\Pi}^{(i)}$ for each subject. The entries of this matrix provide a proportion that takes values in the interval (0, 1). Once we have the stability scores, we consider an additional set of $B = 100$ resamples to implement RAP. Thus, in step 2 of **Algorithm 1**, we form a matrix of random penalties $\Lambda_{RAP}^{(i,b)}$ per resample b . For each edge (k, l) the corresponding adaptive penalty is determined by perturbing initial $\hat{\lambda}^{(i)}$ by an amount κ using a Bernoulli random variable. The probability of success of each Bernoulli r.v is determined by the corresponding stability score for that edge.

$$\Lambda_{RAP}^{(i,b)} = \begin{cases} \hat{\lambda}^{(i)} + \kappa & \text{Ber}(1 - \Pi_{kl}^{(i)}) \\ \hat{\lambda}^{(i)} - \kappa & \text{Ber}(\Pi_{kl}^{(i)}) \end{cases} \quad (8)$$

Putting these components together, R^3 consists of first running Step 1 of **Algorithm 1** to obtain stability scores and then using an additional B resamples based on random adaptive penalization,

summarized in Step 2 of **Algorithm 1** to obtain nB resampled network metrics $\tilde{u}^{(i,b)}$. Note that in subsequent steps we omit the superscripts in $\Lambda_{RAP}^{(i,b)}$ for notational convenience.

Algorithm 1 : R^2 and R^3 Procedures for Testing Covariates Effects on Network Metrics

Step 0: Initial Parameters

Input: $\mathbf{Y}^{(i)}$, **Output:** $\hat{\lambda}^{(i)}$

Estimate $\hat{\lambda}^{(i)}$ using graphical model estimation and selection (StARS) for each subject i .

Step 1: Subject Level Resampling

Input: $(\mathbf{Y}^{(i)}, \hat{\lambda}^{(i)}, B = 100)$, **Output:** Either $\tilde{u}^{*(i,b)}$ or $\hat{\Pi}^{(i)}$

(a) FOR $b = 1, \dots, B$ in the i^{th} subject

- (i) Bootstrap the data $\mathbf{Y}^{(i)}$ to get $\mathbf{Y}^{*(i,b)}$ and sample correlation matrix $\tilde{\Sigma}^{*(i,b)}$
- (ii) Perform a standard graphical lasso $\hat{\Theta}_{\hat{\lambda}^{(i)}}^{*(i,b)}(\tilde{\Sigma}^{*(i,b)})$ in Equation (7)
- (iii) **If R^2 :** Compute network statistic $\tilde{u}^{*(i,b)}$ defined in Section 2.1

END

(b) **If R^3 :** Estimate stability scores $\hat{\Pi}^{(i)} = \frac{1}{B} \sum_b \mathbb{I}(\hat{\Theta}_{\hat{\lambda}^{(i)}}^{*(i,b)}(\tilde{\Sigma}^{*(i,b)}) \neq 0)$

Step 2: Subject Level Resampling & Random Adaptive Penalization (R^3 only)

Input: $(\mathbf{Y}^{(i)}, \hat{\Pi}^{(i)}, \hat{\lambda}^{(i)}, B = 100)$, **Output:** $\tilde{u}^{*(i,b)}$

(a) FOR $b = 1, \dots, B$ in the i^{th} subject

- (i) Bootstrap the data $\mathbf{Y}^{(i)}$ to get $\mathbf{Y}^{*(i,b)}$ and sample correlation matrix $\tilde{\Sigma}^{*(i,b)}$
- (ii) Using stability scores from Step 1(b), compute random adaptive penalties $\Lambda_{RAP}^{(i,b)}$ in Equation (8)
- (iii) Using a weighted graphical lasso, estimate $\hat{\Theta}_{\Lambda_{RAP}}^{*(i,b)}(\tilde{\Sigma}^{*(i,b)})$ in Equation (7)
- (iv) Compute network statistic $\tilde{u}^{*(i,b)}$ defined in Section 2.1

END

Step 3: Population Level Inference for $\hat{\beta}$ using Random Effects

Input: $\{\{\tilde{u}^{*(i,b)}\}_{b=1}^B\}_{i=1}^n$, **Output:** $\hat{\beta}$ and p -values

- (a) Estimate fixed covariate effects $\hat{\beta}$ using mixed effects models. (Section B.1)
 - (b) Compute mixed effects test statistic and p -values in Equation (B.1)
-

2.3.3. Test Statistics for Network Metrics

Both R^2 and R^3 yield a total of nB resampled network statistics that possess two levels of variability. If we applied single level regression techniques to test the covariate effect in Equation (3), we would in effect assume that all the nB resampled

statistics were independent. Test statistics that assume nB independent observations, despite the availability of only n independent clusters of size B are known to be overoptimistic (Laird and Ware, 1982; Liang and Zeger, 1993). To address this overoptimism, a more reasonable assumption is that resampled statistics between any two subjects are independent, whereas within subject resampling statistics are positively correlated. Just as we commonly employ mixed effects models to account for two levels of variation in repeated measures data, we employ similar two-level models to derive test statistics for resampled network metrics.

Let \mathbf{U}_i^* denote the vector $B \times 1$ vector of resampled statistics per subject $\{\tilde{u}^{*(i,b)}\}$. In the case of real valued density metrics, we use a linear mixed effects (LME) model for repeated measures (Laird and Ware, 1982) to account for the two levels of variability in resampled statistics.

$$\mathbf{U}_i^* = \beta_0 + \underbrace{X_i \beta_{\setminus 0}}_{\text{Between Subject}} + \underbrace{Z_i \gamma + R_i a_i}_{\text{Within Subject}} + e_i^* \quad (9)$$

$$\text{Var}(\mathbf{U}_i^*) = V_i = \phi^2 \mathbf{I}_B + R_i v^2 R_i^\top \quad (10)$$

Here a_i are *i.i.d* subject level random intercepts with variance $\text{Var}(a_i) = v^2$, $R_i = \mathbf{1}_{B \times 1}$ is the random effect design matrix, and e_i^* is independent of a_i and captures within subject sampling variability with variance $\text{Var}(e_i) = \phi^2 \mathbf{I}_B$ where \mathbf{I} denotes the identity. From hereon, we ignore the intercept β_0 , and assume that β denotes the $(q \times 1)$ vector of explanatory fixed effects.

Estimation and inference for linear mixed effect models are well covered in the neuroimaging literature in the context of functional activation studies and longitudinal designs (Beckmann et al., 2003; Bernal-Rusiel et al., 2013). We employ standard estimators and test statistics for linear mixed effects models including generalized least squares estimators for $\hat{\beta}$ and corresponding restricted maximum likelihood (ReML) estimators of variance to obtain F -test statistics to test the null hypothesis regarding β , the covariate effects. A thorough review of mixed effects models can be found in Agresti (2015) and we also spell these out in more detail for our methods in Supplementary Materials.

3. SIMULATION STUDY

In this section, we seek to evaluate our framework for testing covariate effects by conducting a rigorous power analysis using realistic fMRI network structures. We obtain realistic network structures for fMRI functional connectivity by using networks estimated from real data as the basis of our simulated networks. First, we synthetically create multivariate data according to our two-level models using realistic graph structures in Section 3.1. Since we know the true structure of graphical models and their network metrics we empirically measure statistical power and type-I error for all methods. Then, in Section 3.2 we offer two key results. First, by employing simulations using two-level models of variability in Equation (4) that reflect how functional networks are analyzed in practice, we provide a more realistic assessment

of when we lose statistical power due to sample sizes (t, n) and covariate signal-to-noise (SNR) controlled by population variance v^2 . Second, we show that both R^2 and R^3 mitigate the challenges discussed in Section 2.2 and improve statistical power over standard test statistics under various sample sizes and covariate SNR regimes.

3.1. Simulation Setup for Node and Subnetwork Density

We simulate multivariate data according to our two level models in Section 2.1. We know from previous work that the graph structure or location of non-zeros in the inverse covariance (Narayan et al., 2015) influences the difficulty of estimating individual subject networks accurately. Using a group level empirical inverse correlation matrix obtained from 90 healthy subjects in the Michigan sample of the ABIDE dataset, preprocessed in Section 4, we threshold entries smaller than $\tau = |.25|$ to create a baseline network A_0 that contributes to the intercept term β_0 of our model (Equation 4). Illustrations of this baseline network can be found in Figure A.0 in the Supplementary Materials. Then we create individual adjacency matrices and network metrics $u^{(i)}$ according to the linear model (Equation 5). We create inverse correlation matrices $\Theta^{(i)}$ using the graph structure provided by A_0 and ensure $\Theta^{(i)}$ is positive definite.

Our main focus in the simulation study is to conduct a rigorous power analysis to detect covariate effects on node density and subnetwork density under a range of sample sizes and population variability and demonstrate the benefits of using R^3 and R^2 over standard approaches. Recall from Section 2.1 that node density is the degree of a node, while the subnetwork density is the number of connections between sets of nodes that make up a submatrix or subnetwork of the inverse covariance matrix. We obtain empirical estimates of statistical power by measuring the proportion of times we successfully reject $\hat{\beta}_{\setminus 0} = 0$ at level $\alpha = .05$, in the presence of a true covariate effect $\beta_{\setminus 0} \neq 0$, across 150 monte-carlo trials for a simulation scenario. Similarly, we obtain an empirical estimate of type I error by measuring the proportion of times we reject $\hat{\beta}_{\setminus 0} = 0$ at level $\alpha = .05$ in the presence of a null covariate effect of $\beta_{\setminus 0} = 0$.

Although one could choose to vary a large number of parameters for these simulations, we focus on the parameters most important for a power analysis, sample sizes and population variance, (t, n, v^2) , while fixing other parameters such as number of covariates to $q = 1$, $r = 0$ and number of nodes to $p = 50$. We present a 3×3 panel of 9 power analyses of node density in **Figure 2** where we vary $t = \{p, 2p, 4p\}$ along the y-axis and $v^2 = \{0.1, 0.25, 0.5\}$ along the x-axis. Then within each sub-panel, we evaluate statistical power at subject sample sizes of $n = \{5, 10, \dots, 95\}$. For the entire 3×3 panel we hold the intercept and covariate effect fixed at $\beta_0 = 2$, $\beta_1 = 1$. Thus, each sub-panel illustrates statistical power as a function of subject sample size n for a fixed value of (t, v^2) . Similarly, in **Figure 3** we present power analyses for subnetwork density where we hold the intercept and covariate fixed at $\beta_0 = 5$, $\beta_1 = 2$ and use subnetworks of size $0.1p = 10$ nodes. We use

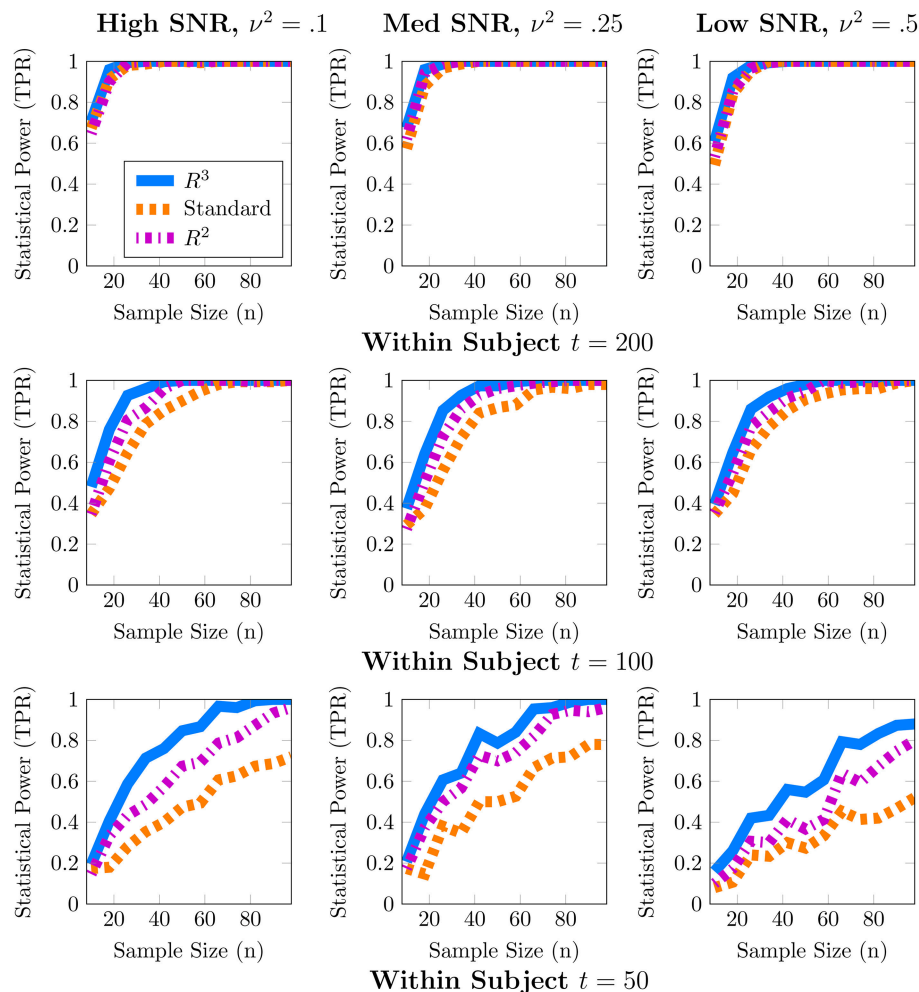


FIGURE 2 | Statistical power analysis for node density. When node density varies with an explanatory covariate ($q = 1$), statistical power to detect this covariate effect improves with subject sample size n but crucially depends on the number of independent fMRI samples t from a single session and relative size of the covariate effect, $\beta_1 = 1$, to population variance ν^2 (covariate SNR). When $t \approx p$, estimates of node density are both highly variable and potentially biased. By accounting for these issues, R^3 and R^2 improve estimates of network metrics, thus exceeding 80% power, whereas the standard F -test is substantially less powerful. Note that R^3 and R^2 are more powerful at smaller sample sizes compared to the standard approach. However, when fMRI samples become sufficiently large at $t \approx 4p$, all methods become similarly powerful for detecting covariate effects of node density. Empirical statistical power is defined as $\frac{\# \text{ of times reject } \mathcal{H}_0}{\# \text{ of Monte Carlo Trials}}$ when the alternative is true in Equation (3).

larger values for covariate effects to ensure that the number of edges in a subnetwork are realistically large for a subnetwork with 10 nodes. While the sample sizes (t, n) are identical to those in node density, we increase $\nu^2 = \{0.4, 1, 2\}$ to match β . This ensures that covariate signal to noise ratio $\frac{\|X\beta_1\|_2^2}{\nu^2}$ is similar for both metrics. Note that the intercept values β_0 in both power analyses were based on the average node degree in A_0 or average subnetwork density for subnetworks of size 10 in A_0 . For each power analysis, we have a corresponding simulation of type-I error, obtained by setting $\beta_1 = 0$ while keeping all other parameters equivalent. The full set of type-I error control results are presented in Supplementary Materials, and one representative simulation for each metric is presented in Figure 4.

3.2. Simulation Results

In these simulations, our methods, R^3 and R^2 , empirically outperform standard methods in terms of statistical power, particularly when within subject observations are comparable to the dimension of the network, and subject networks are harder to estimate correctly. Recall from Section 2.2 that we expect to lose statistical power when individual subject networks are difficult to estimate correctly, due to additional sampling variability and bias in network metrics. As expected, power analyses for both metrics in Figures 2, 3 reveal that statistical power deteriorates as observations t available for subject network estimation reduces. Moreover, this loss of statistical power cannot always be compensated by larger subject sample sizes n . For example, the best achievable statistical power at large subject

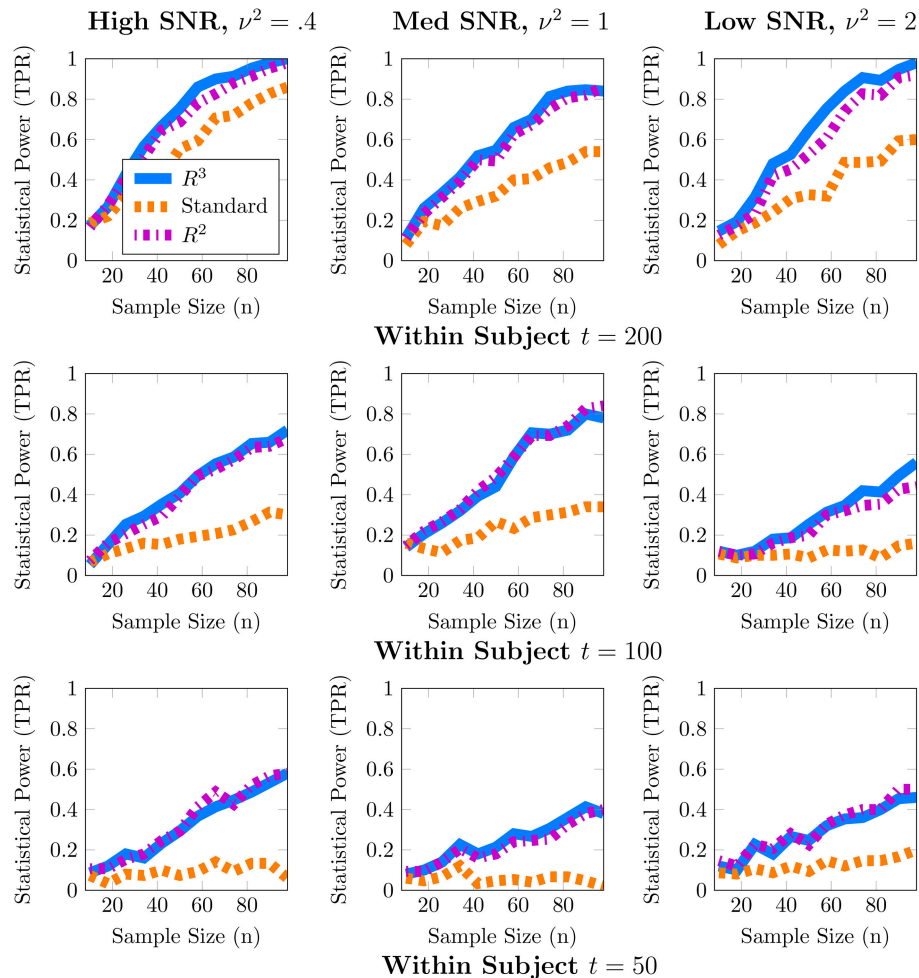


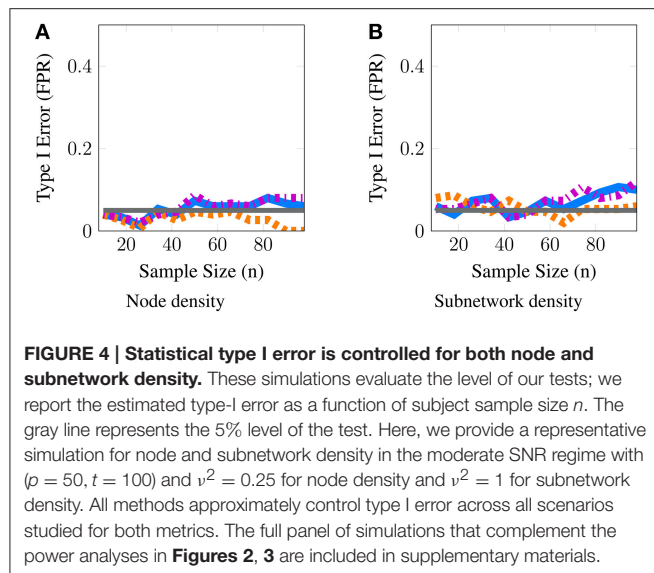
FIGURE 3 | Statistical power for subnetwork density. When subnetwork density varies with an explanatory covariate ($q = 1$), statistical power to detect this effect improves with subject sample size n but crucially depends on the number of independent fMRI samples t from a single session and the relative size of the covariate effect, $\beta_1 = 2$, to the population variance ν^2 (covariate SNR). For many values of (t, p) estimates of subnetwork density are both highly variable and potentially biased. By accounting for these issues, both R^3 and R^2 test statistics substantially improve statistical power across all regimes at smaller subject sample sizes, whereas the standard F -test is substantially less powerful. We note that covariate effects on subnetwork metrics are particularly hard to detect when $t \approx p$, with statistical power often below 60%. Empirical statistical power is defined as $\frac{\# \text{ of times reject } \mathcal{H}_0}{\# \text{ of Monte Carlo Trials}}$ when the alternative is true in Equation (3).

samples of $n \approx 100$ begins to deteriorate when $t = p$. While, the best achievable statistical power often exceeds 90% for node density when $t > p$, it drops as low as 80% for R^3 and R^2 . The standard approach in contrast drops below 60% node density. In the case of subnetwork density, statistical power for R^3 and R^2 exceed 80% when $t = 4p$, this drops as low as 60% at more modest sample sizes of $t = 2p$ and further down to 40% at $t = p$. The standard approach falls to below 40% more quickly at $t = 2p$ and below 20% when $t = p$.

Just as with subject sample size, when individual network estimation is easy in our simulations with larger within subject observations of $t = 4p$, the covariate signal to noise ratio or SNR has an almost negligible impact on statistical power. However, as t decreases, network estimation becomes harder and consequently, all methods become much more sensitive to SNR. For example, in

regimes where $t = 2p$, network estimation is moderately hard but detecting covariate effects is achievable at high SNR. However, we observe that all methods lose power as covariate SNR decreases. We also observe that loss of statistical power due to SNR is more pronounced at smaller subject sample sizes of $n < 60$. Such a result is expected since sampling variability of covariate effect β_1 is proportional to population variance ν^2 and decreases with larger subject sample sizes n .

We noted earlier in Section 2.3 that we expect the benefits of R^3 over R^2 to be the greatest for finest scale metrics at the edge level which are most sensitive to graph selection errors and decrease as metrics measure density at more global levels. Whereas, random penalization improves statistical power relative to R^2 for two-sample differences at the edge level Narayan et al. (2015), they share similar statistical power for



node and subnetwork density in most simulations presented here, with some marginal benefits for node density. R^3 offers greater benefits over R^2 at small sample sizes t for networks that are more sparse and where the stability of true edges over false edges can be improved via random penalties. All methods, including R^3 and R^2 are unable to detect covariate effects when estimation of individual networks becomes unreliable under high density regimes. We provide additional simulations that vary the sparsity of baseline networks in Figure A.3 in the Supplementary Materials.

Finally, in **Figure 4**, we provide evidence that type-I error is controlled by all methods for both node and subnetwork density. The full panel of simulations that complement **Figures 2, 3** are included in Supplementary Materials.

From these simulations we conclude that resampling based approaches are more efficient, i.e., they have higher statistical power for both node and subnetwork density at smaller subject sample sizes n , particularly for smaller t and lower covariate SNR. Another insight from these simulations is that given a fixed budget of fMRI session time, it is preferable to increase the number of within session observations t per subject for fewer number of subjects n in order to maximize statistical power. For studies where each fMRI session consists of observations comparable to the size of networks ($t, p \in [100, 200]$), as well as for studies that cannot recruit a large number of subjects, our methods, R^3 and R^2 , make better use of available data and improve statistical power compared to standard approaches to network analysis.

4. CASE STUDY

A number of recent studies on autism spectrum disorders (ASD) have found differences in functional connectivity that were correlated with symptom severity as measured by Autism Diagnostic Interview (ADI) or Autism Diagnostic Observation Schedule (ADOS). However, the majority of these studies that

link symptom severity to functional connectivity derive networks using pairwise correlations (Supekar et al., 2013; Uddin et al., 2013b). An important shortcoming of studying differences in pairwise correlation networks is that edges in a true correlational network might be present due to the effect of “common causes” elsewhere in the brain and do not necessarily represent a direct flow of information. Thus, while correlational networks can provide network biomarkers for autism (Supekar et al., 2013), it is more problematic to infer network mechanisms of behavioral deficits in ASD exclusively using correlational networks. However, by studying previously implicated regions and subnetworks using Gaussian graphical models (GGMs), we strengthen the interpretation of variations in network structure linked to autism severity. Thus, by employing our two level models (Equation 1) based on GGMs to detect covariate effects, we enable scientists to infer that any network differences linked with behavioral deficits implicate nodes and edges directly involved in the disease mechanism. Guided by the successes of our simulation study, we employ R^3 to investigate the relationship between cognitive scores on node and subnetwork densities in autism spectrum disorders. In particular, we conduct tests for covariate effects on two density metrics, the node density and subnetwork density. Node density counts the number of connections between a single region of interest to all other regions whereas subnetwork density counts the number of connections between sets of regions or subnetworks. We investigate nodes and subnetworks hypothesized in the literature (Uddin, 2014) to be involved in regulating attention to salient events and explanatory for behavioral deficits in ASD.

4.1. ABIDE Data Collection and Preprocessing

We use resting state fMRI data collected from the Autism Brain Imaging Data Exchange (ABIDE) project (Di Martino et al., 2014b) and preprocessed by the Preprocessed Connectomes Project (PCP) (Craddock and Bellec, 2015) using the configurable-pipeline for analysis of connectomes or (C-PAC) toolbox (Craddock, 2014; Giavasis, 2015). In order to properly account for site effects, we choose to focus on two major sites with relatively large samples, UCLA and Michigan, resulting in 98 and 140 subjects per site. While both ADOS and ADI-R cognitive scores are available for these sites, we focus on ADOS scores obtained using the Gotham algorithm (Gotham et al., 2009), which is known to be comparable across different age groups.

The ABIDE data was acquired (Di Martino et al., 2014b) using T2 weighted functional MRI images with scan parameters $TR = 2$ at the Michigan site and $TR = 3$ at the UCLA site. Subsequently, this data was minimally preprocessed using the C-PAC utility (Craddock and Bellec, 2015; Giavasis, 2015), including slice timing correction, motion realignment and motion correction using 24 motion parameters, and normalization of images to Montreal Neurological Institute (MNI) 152 stereotactic space at $3 \times 3 \times 3 \text{ mm}^3$ isotropic resolution. The pipeline was also configured to regress out nuisance signals from the fMRI time-series. The nuisance variables included were physiological confounds such as heart beat and respiration, tissue signals

and low frequency drifts in the time-series. We did not regress out the global signal as this operation is known to introduce artifacts in the spatial covariance structure (Murphy et al., 2009). Additionally, we did not apply band pass filtering as this would interfere with subsequent temporal whitening that we describe later in this section. Preprocessed data without bandpass filtering and global signal regression is available using the *noglobalnofilt* option in the PCP project. Finally, the spatial time-series was parcellated into times-series \times regions of interest using the Harvard-Oxford atlas distributed with FSL (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>). Here we included $p = 110$ regions of interest including 96 cortical regions and 14 subcortical regions. Regions corresponding to white matter, brain stem and cerebellum were excluded. The resulting time-series \times regions data matrix for each individual subject is ($t = 116, p = 110$) for UCLA subjects and ($t = 300, p = 110$) for Michigan subjects. This preprocessed dataset has been archived in a public repository (<http://dx.doi.org/10.6084/m9.figshare.1533313>).

4.2. Previously Implicated Subnetworks and Regions

Distinct lines of evidence suggest the involvement of limbic, fronto-parietal, default mode and ventral attention regions in ASD. Uddin (2014) summarize the evidence in favor of a salience-network model to explain behavioral dysfunction in responding to external stimuli. According to this model, the salience network regions that span traditional limbic and ventral attention systems play a vital role in coordinating information between the default mode regions involved in attending to internal stimuli and the fronto-parietal regions involved in regulating attention to external stimuli. Together, these interactions enable appropriate behavioral responses to “salient” or important events in the external environment. Uddin et al. (2013a) conducted a network-based prediction study and found that connectivity features of the anterior cingulate cortex, and the anterior insula, predict an increase ADOS repetitive behavior scores. Similarly, another study by Di Martino et al. (2009) also implicates connectivity of anterior insula and anterior cingulate cortex to deficits in social responsiveness in Autism. Cherkassky et al. (2006); Monk et al. (2009) implicate posterior cingulate connectivity within the default mode network in ASD. Alaerts et al. (2013) show that deficits in emotion recognition were correlated with network features in the right posterior superior temporal sulcus, a result also supported in the wider literature (Uddin et al., 2013b).

Additionally, we also major findings from previous analyses of the ABIDE dataset that include the UCLA or Michigan subject samples. Whole brain voxelwise analysis by Di Martino et al. (2014b) revealed covariate effects associated with the mid insula, posterior insula, posterior cingulate cortex and thalamus. Group level two-sample tests of functional segregation and integration in seed based functional connectivity (Rudie et al., 2012a,b) reveal differences in the amygdala, IFG right pars opercularis.

Based on our review of existing literature, we seek to detect covariate effects with respect to 23 hypotheses regarding the density of connections. Of these 23 hypotheses, 13 correspond

to density of connections of nodes or brain regions with respect to the whole brain, and 10 correspond to the density within and between 4 large scale functional subnetworks. These regions are defined using the Harvard-Oxford atlas with large scale subnetworks provided by Yeo et al. (2011). **Figure 5** illustrates the volumes associated with the 13 regions of interest. **Figure 6** illustrates the four large scale functional brain networks we consider, namely, the default mode, the frontoparietal, the limbic and the ventral attention networks as defined by Yeo et al. (2011). By explicitly testing the density of long-range connections in brain regions and networks previously linked with ASD, we aim to identify network structures at the node and subnetwork level that are directly involved in behavioral deficits.

4.2.1. Testing for Covariate Effects via R^3

We employ the linear model from Equation (5) for node and subnetwork density to test the null hypothesis that ADOS covariates have no effect on density. For this analysis, we jointly consider two related explanatory covariates, the ADOS Social Affect (SA) and the ADOS Restricted, Repetitive Behavior (RRB) scores ($q = 2$), while accounting for differences in clinical evaluation across sites, by incorporating site as a nuisance covariate ($r = 1$). We eliminate subjects without ADOS cognitive scores, leaving us with $n = 100$ autism subjects. Thus, the final data tensor for covariate tests contains either $t = 116$ (UCLA) or $t = 300$ (Michigan) time-points for $p = 110$ brain regions in $n = 100$ subjects.

Before applying the R^3 procedure from Section 2.3 to the preprocessed ABIDE dataset, we need to ensure fMRI observations are approximately independent. By whitening temporal observations, we ensure that estimating individual subject networks is more efficient. We achieve this by first estimating the temporal precision matrix $\hat{\Omega} = \sum_{i=1}^n \mathbf{Y}^{(i)}(\mathbf{Y}^{(i)})^\top$ using the banded regularization procedure of Bickel and Levina (2008) for autoregressive data and whitening the fMRI time-series of each subject $\tilde{\mathbf{Y}}^{(i)} = \hat{\Omega}^{1/2} \mathbf{Y}^{(i)}$. To choose the number of lags, we conduct model selection via cross-validation (Bickel and Levina, 2008). Given these whitened observations, we apply the R^3 procedure outlined in **Algorithm 1**. We initialize regularization parameters using StARS and subsequently perturb these parameters according to RAP as described in Section 2.3. Since we have a total of 23 node density and subnetwork density hypotheses, we control the false discovery rate at the 5% level using the Benjamini-Yekutieli procedure (Benjamini and Yekutieli, 2001).

4.3. ABIDE Data Analysis: Results

Tables 1, 2 show statistically significant covariate effects for 3 subnetwork hypotheses and 5 regions of interest. Notable findings amongst subnetwork hypotheses in **Table 1** are that an increase in behavioral deficits indicated by restricted and repetitive behavior scores (RRB) and social affect (SA) is associated with a decrease in connection densities in frontoparietal-based subnetworks. The 3 prominent findings involve connection densities between the frontoparietal to limbic subnetworks, between the frontoparietal to ventral attention subnetworks and between the default mode and limbic

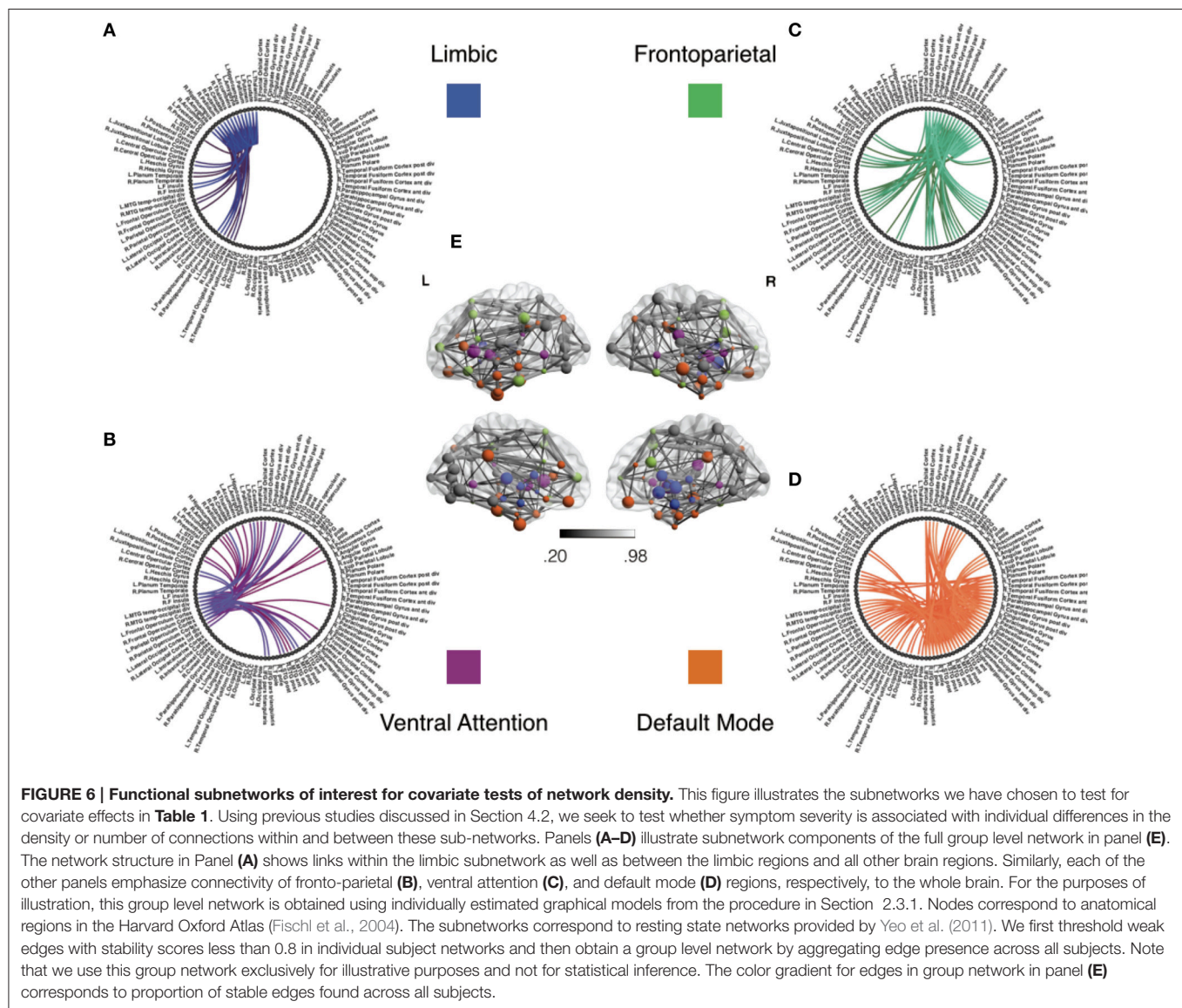
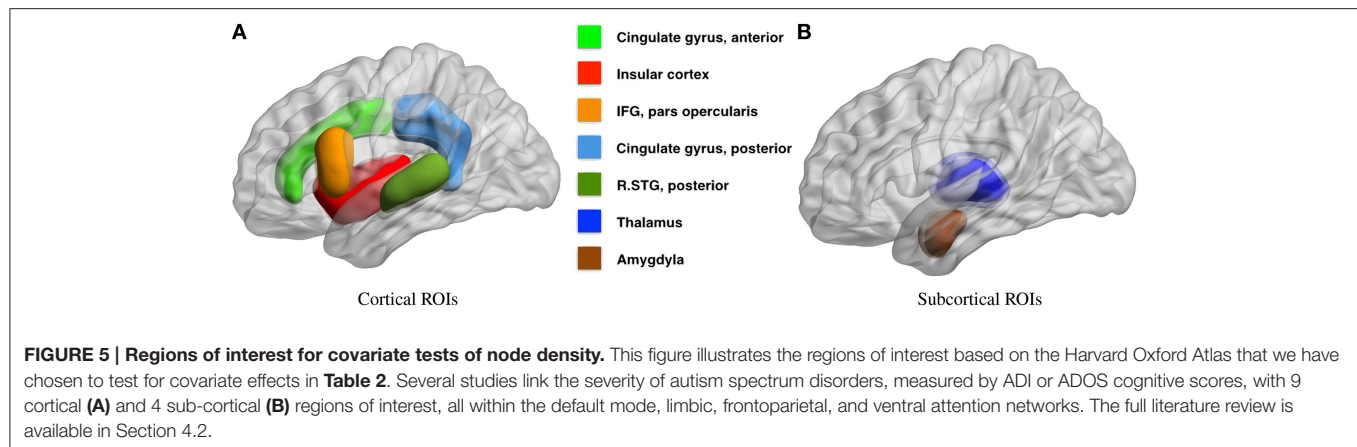


TABLE 1 | Joint ADOS covariate effects on subnetwork density.

Subnetwork 1	Subnetwork 2	pval (RRB + SA)	RRB	CI (L)	CI (U)	SA	CI (L)	CI (U)	SITE	CI (L)	CI (U)
Default	Default	0.061200	-2.66	-7.10	1.78	-0.66	-2.92	1.59	0.22	-5.85	6.30
Default	Frontoparietal	0.010000	-2.34	-4.97	0.29	-0.18	-1.52	1.15	-0.52	-4.12	3.08
Default	Limbic	0.004530*	-1.51	-3.06	0.04	-0.10	-0.89	0.69	-0.30	-2.42	1.82
Default	Ventral attention	0.038000	-0.83	-1.76	0.10	0.08	-0.40	0.55	0.37	-0.91	1.64
Frontoparietal	Frontoparietal	0.007030	-1.36	-3.23	0.52	-0.51	-1.47	0.44	-0.47	-3.04	2.10
Frontoparietal	Limbic	0.000088*	-1.15	-1.98	-0.31	0.00	-0.43	0.43	0.23	-0.92	1.38
Frontoparietal	Ventral attention	0.003793*	-0.61	-1.16	-0.06	0.03	-0.25	0.31	0.75	0.00	1.50
Limbic	Limbic	0.530000	-0.19	-1.70	1.32	-0.35	-1.11	0.42	-0.77	-2.83	1.29
Limbic	Ventral attention	0.955000	0.01	-0.45	0.46	-0.05	-0.28	0.18	-0.69	-1.31	-0.06
Ventral attention	Ventral attention	0.196000	-0.05	-0.50	0.40	-0.21	-0.44	0.02	-0.24	-0.86	0.37

We jointly test the effects of two ADOS covariates on subnetwork density while accounting for site effects as a nuisance covariate. Here, the most prominent findings suggest that a decrease in the number of direct connections between frontoparietal to limbic, between frontoparietal to ventral attention subnetworks and between default to limbic subnetworks is linked with increased ADOS symptom severity. This result is consistent with the hypothesis that abnormalities within the salience network, comprising anterior cingulate cortex (a region within our frontoparietal network) and insula (a region within our ventral attention network), results in a failure to regulate between attention to external stimuli vs. attention to internal thoughts. A total of three subnetworks, denoted by *, survive corrections for multiplicity, using false discovery control over all 23 hypotheses tested at the 5% level using Benjamini-Yekutieli. Although estimates of site effects were non-zero, individual confidence intervals for most site effects are close to or include zero and were thus not statistically significant after corrections for multiplicity. Results are discussed further in Section 4.3

TABLE 2 | Joint ADOS covariate effects on node density.

Subnetwork	Region	pval (RRB + SA)	RRB	CI (L)	CI (U)	SA	CI (L)	CI (U)	SITE	CI (L)	CI (U)
Default	L. Cingulate post.	0.004600*	-0.68	-1.35	-0.02	-0.01	-0.34	0.33	-0.05	-0.96	0.86
Default	R. Cingulate post.	0.009000	-0.49	-0.96	-0.01	0.03	-0.21	0.27	0.39	-0.26	1.03
Default	R. pSTG	0.010900	-0.41	-0.85	0.04	-0.02	-0.24	0.21	0.08	-0.53	0.69
Frontoparietal	R. Cingulate ant.	0.002100*	-0.30	-0.65	0.04	-0.08	-0.26	0.10	0.62	0.14	1.09
Frontoparietal	R. IFG pars oper	0.004100*	-0.69	-1.30	-0.09	0.06	-0.25	0.36	0.05	-0.78	0.88
Frontoparietal	L. Cingulate ant.	0.005400*	-0.55	-1.14	0.05	-0.06	-0.36	0.24	-0.05	-0.86	0.76
Frontoparietal	L. IFG pars oper	0.058000	-0.29	-0.70	0.11	-0.01	-0.22	0.19	0.07	-0.48	0.62
Limbic	R. Thalamus	0.004200*	-0.46	-1.02	0.10	-0.12	-0.41	0.16	-0.65	-1.41	0.12
Limbic	L. Thalamus	0.037700	-0.49	-1.20	0.21	-0.08	-0.43	0.28	-0.57	-1.53	0.39
Limbic	R. Amygdyla	0.092500	-0.29	-0.23	0.21	-0.01	-0.72	0.14	-0.26	-0.85	0.33
Limbic	L. Amygdyla	0.175100	-0.14	-0.49	0.21	-0.07	-0.25	0.11	0.30	-0.19	0.78
Ventral attention	L. Insula	0.223300	-0.11	-0.46	0.24	-0.08	-0.26	0.09	-0.24	-0.72	0.23
Ventral attention	R. Insula	0.306800	-0.13	-0.60	0.34	-0.10	-0.34	0.14	-0.76	-1.40	-0.12

We jointly test the effects of two ADOS covariates on node density while accounting for site effects as a nuisance covariate. Notably, we find that a decrease in the number of direct connections between left posterior cingulate cortex (PCC) and anterior cingulate cortex (ACC) with all other regions is linked with an increase in ADOS symptom severity. This result corroborates previous findings that ACC (a component of the salience network) and PCC connectivity might be directly involved behavioral deficits ASD. A total of five regions, denoted by *, survive corrections for multiplicity, using false discovery control over all 23 hypotheses tested at the 5% level using Benjamini-Yekutieli. Although estimates of site effects were non-zero, individual confidence intervals for most site effects are close to or include zero and were thus not statistically significant after corrections for multiplicity. Results are discussed further in Section 4.3.

subnetworks. Individual regression coefficients and confidence intervals for RRB and SA suggest that of the two covariates, RRB scores particularly dominate the decrease in subnetwork density for two of these results, particularly the frontoparietal-limbic subnetwork. The most prominent results amongst region of interest hypotheses in **Table 2** suggest that ADOS symptom severity is again associated with hypoconnectivity or a decrease in the number of connections between each of the following regions with the rest of the network—bilateral pairs of anterior cingulate cortex (ACC); left posterior cingulate cortex (PCC); the right inferior frontal gyrus (IFG); and the thalamus. Note that we use a conservative Benjamini-Yekutieli procedure (Benjamini

and Yekutieli, 2001) to control for FDR at the 5% level under arbitrary dependence amongst the 23 hypotheses tested. Under a less conservative procedure, Benjamini-Hochberg (Benjamini and Hochberg, 1995), four additional hypotheses including the within-frontoparietal subnetwork and the right PCC are statistically significant at 5% FDR control. While the regression coefficients for site effects are non-zero in both analyses, most confidence intervals either contain zero or are very close to zero and not statistically significant. The one exception amongst our prominent findings, the right ACC, shows statistically significant site effects. We also find site effects for two hypotheses where we did not detect ADOS effects, namely, the limbic to ventral

attention subnetwork and right insula. However, these site effects are not statistically significant after correcting for multiplicity.

Our analysis strongly implicates the frontoparietal-limbic subnetwork, and frontoparietal-ventral attention subnetworks, as well as posterior/anterior cingulate cortical connections with the rest of the brain, in behavioral deficits of ASD. Since we identify these regions and subnetworks using partial correlation measures of functional connectivity, our results provide strong evidence that these network components are directly involved in ASD. In particular, since the salience network (Buckner et al., 2013; Uddin et al., 2013a) is thought to comprise the ACC, which falls within our frontoparietal network, and insular regions that overlap limbic and ventral attention networks in our analysis, our subnetwork findings are consistent with the salience network explanation for behavioral deficits in autism. Additionally, our findings strongly implicate frontoparietal-limbic relationships. While our region of interest analysis found abnormalities in thalamar connectivity, a component of the limbic network, other limbic regions could also be directly involved in ASD and thus warrant further study.

We contrast our findings on the 23 a-priori hypotheses in Section 4.2 with previous analyses that were obtained by conducting network analyses on correlational networks, including previous analyses of the same ABIDE dataset. Our analysis detects only a subset of previous covariate effects on ASD network structure when using GGM based networks via R^3 . Correlational network analysis using the UCLA and UM samples of ABIDE (Rudie et al., 2012b; Di Martino et al., 2014b) as well as those from alternative sites (Uddin et al., 2013b) link insular, amygdalar connectivity with autism symptoms, whereas we do not detect strong effects for these regions for density metrics. The absence of strong covariate effects using GGMs suggests that the insular and amygdalar connections might be associated with behavioral deficits in autism only due to indirect correlations with other regions of interest. Similarly, although we find abnormalities in the PCC, a region within the default mode network, and between the default-mode and the limbic regions, we failed to find abnormalities linking the default mode with frontoparietal or ventral attention networks. This suggests that previous findings involving the default mode network could have been the result of indirect pairwise correlations, possibly driven by PCC and limbic regions. Although we use novel functional connectivity models and methods to analyze the ABIDE dataset, some of our choices of a-priori hypotheses for this analysis, notably, the inclusion of IFG pars opercularis and the amygdala for node density, were guided by alternative analyses of the ABIDE dataset (Rudie et al., 2012b; Di Martino et al., 2014b). Thus, we need further validation of the purported effects of ADOS on IFG pars opercularis density.

5. DISCUSSION

This paper investigates an understudied issue in neuroimaging—the impact of (often imperfectly) estimated functional networks on subsequent population level inference to find differences across functional networks. Using an important class of network

models for functional connectivity, Gaussian graphical models, we demonstrate that neglecting errors in estimated functional networks reduces statistical power to detect covariate effects for network metrics. While lack of statistical power due to small subject sizes is well documented in neuroimaging (Button et al., 2013), recent test re-test studies (Birn et al., 2013; Laumann et al., 2015) suggest that typical fMRI studies of 5–10 min are highly susceptible to lack of statistical power. This paper provides additional evidence that within subject sample size, t , is important for well powered studies. For typical studies where t is comparable to the number of nodes p , errors in estimating functional networks can be substantial and not accounted for by standard test statistics. We show that our methods to mitigate this problem, R^2 and R^3 , are always at least as powerful if not substantially more powerful than standard test statistics under a variety of sample sizes and covariate signal-to-noise regimes. Additionally, regardless of the methods employed, our power analyses suggest that in many scenarios, particularly when subject level networks are large, a more efficient use of a fixed experimental budget would be to collect more within subject measurements and fewer subject samples in order to maximize statistical power to detect covariate effects. While we demonstrate this result on the joint importance of within and between subject sample sizes using density based network metrics, we expect such results to hold more generally whenever population level functional connectivity analyses are conducted in a two-step manner where subject level networks are estimated initially and population level metrics then explicitly depend on the quality of subject level network estimates. In practice, we additionally need to incorporate other considerations beyond statistical power in choosing within subject scan length such as increase in movement or the discomfort to participants particularly in patient populations. These issues related to statistical power warrant further investigation in future work.

This paper also highlights the scientific merits of employing explicit density based metrics in graphical models of functional connectivity to gain insights into disease mechanisms at a macroscopic level using the ABIDE dataset (Di Martino et al., 2014b). In Section 4, we sought to detect covariate effects on the density of direct, long range functional connections in Autism Spectrum Disorders (ASD). Notably, our results in Section 4.3, at both the subnetwork and node level favor the hypoconnectivity hypothesis for behavioral deficits in ASD. Specifically, we find that a reduction in directly involved long-range functional connections between parcellated regions of interest increases ADOS symptom severity. Assuming that the salience network model of autism dysfunction is correct (Uddin, 2014), our results suggest that reduced interactions between the executive control network and the salience network, as well as default mode to the salience network might be responsible for ASD symptoms. Since we employ GGM based models, a plausible interpretation of such hypoconnectivity is that regions in ventral attention and limbic systems fail to adequately communicate with frontoparietal regions that participate in executive control and default mode regions that participate in internal attention. A previous study found evidence of hyperconnectivity when counting the number of local voxelwise connections in Keown et al. (2013). Our results

do not contradict this finding since a network architecture of ASD could involve both reduced long range connections as well as increased density of local connections (Rudie and Dapretto, 2013). Other results on hyperconnectivity (Supekar et al., 2013; Uddin et al., 2013a) do not explicitly employ degree or density of connections to measure hyper or hypo-connectivity but measure the strength of the mean pairwise correlation within and between regions and subnetworks. While the effect in Supekar et al. (2013) appears to be a large and robust finding, the correlational model of connectivity employed in their analysis could be misleading since it includes both direct and indirect functional connections and does not explicitly measure the density of connections. While further studies are needed to resolve the questions raised by Rudie and Dapretto (2013) on this matter, we emphasize that since graphical models of functional connectivity capture direct functional connections, such models enable stronger scientific conclusions regarding functional network mechanisms compared to purely correlational models where edges do not necessarily reflect direct communication between regions.

As we discuss in the simulation results in Section 3.2, our ability to detect covariate effects in populations of graphical models deteriorates in highly dense regimes of network structure where the density or number of edges in the network increases substantially while the number of within subject observations remains limited, or when the individual networks contain a large number of hub-like structures (Ravikumar et al., 2011; Zhou et al., 2011). Since our resampling based methods are a framework that employ existing graph estimation algorithms (Section 2.3), they inherit the strengths and limitations of the specific graph estimation algorithm in such high density regimes. By incorporating new and improved estimators (Yang et al., 2014) for graphical models at the level of individual subjects, we expect corresponding variants of our resampling framework to detect covariate effects under a wider range of network density regimes.

While this paper specifically considers network models (Equation 1) where neuroimaging data is distributed according to a multivariate normal, alternative distributions can be employed for the subject level model in Equation (1), including matrix variate distributions (Allen and Tibshirani, 2012; Zhou et al., 2014) that can account for the serial correlation in temporal observations, and non-parametric graphical models (Lafferty et al., 2012) that relax assumptions of normality. Furthermore, while we focus on resting state functional connectivity in fMRI in this work, our concern regarding errors in estimating large functional networks is applicable to other imaging modalities including EEG/MEG studies. In fact, our two level models (Equation 1) and R^3 framework can be easily extended to functional network analyses based on partial coherence (Sato

et al., 2009) networks or vector autoregressive models (Koenig et al., 2005; Schelter et al., 2006) that are popular in EEG/MEG studies. Additionally, our results are highly relevant to dynamic functional connectivity (Chang and Glover, 2010) analyses where studies estimate separate time-varying functional networks per subject using short sliding-windows of 30–60 s rather than 5–10 min. In such a high dimensional setting where $t \ll p$, our power analyses in **Figures 2, 3** suggest that such dynamic network analyses will be highly underpowered and could benefit from our methods. Thus, extensions of the R^3 framework for dynamic connectivity analyses as well as other multivariate network models is a promising avenue of research. Other areas of investigation include inference for partial correlation strength and corresponding weighted network analysis, as well as including high dimensional covariates in our general linear model (Equation 2). Overall, this work reveals that accounting for imperfectly estimated functional networks dramatically improves statistical power to detect population level covariate effects, thus highlighting an important new direction for future research.

6. DATA SHARING

The preprocessed ABIDE dataset used in this paper is available at <http://dx.doi.org/10.6084/m9.figshare.1533313>. Software for reproducing our analysis is provided at <https://bitbucket.org/gastats/monet/downloads>.

AUTHOR CONTRIBUTIONS

MN and GA conceived and designed the research. MN conducted data analysis. MN and GA wrote and revised the paper.

FUNDING

MN and GA are supported by NSF DMS 1264058. MN is supported an AWS (Amazon Web Services) research grant for computational resources.

ACKNOWLEDGMENTS

The authors thank Steffie Tomson for helpful discussions and advice on preprocessing the ABIDE dataset.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnins.2016.00108>

REFERENCES

- Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* 26, 63–72. doi: 10.1523/JNEUROSCI.3874-05.2006
- Agresti, A. (2002). *Categorical Data Analysis*, Vol. 359. Hoboken, NJ: John Wiley & Sons.
- Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. Hoboken, NJ: John Wiley & Sons.
- Alaerts, K., Woolley, D. G., Steyaert, J., Di Martino, A., Swinnen, S. P., and Wenderoth, N. (2013). Underconnectivity of the superior temporal sulcus

- predicts emotion recognition deficits in autism. *Soc. Cogn. Affect. Neurosci.* 9, 1589–1600. doi: 10.1093/scan/nst156
- Alexander-Bloch, A., Lambiotte, R., Roberts, B., Giedd, J., Gogtay, N., and Bullmore, E. (2012). The discovery of population differences in network community structure: new methods and applications to brain functional networks in schizophrenia. *NeuroImage* 59, 3889–3900. doi: 10.1016/j.neuroimage.2011.11.035
- Allen, G. I., and Tibshirani, R. (2012). Inference with transposable data: modelling the effects of row and column correlations. *J. R. Stat. Soc. Ser. B* 74, 721–743. doi: 10.1111/j.1467-9868.2011.01027.x
- Balachandran, P., Airolidi, E., and Kolaczyk, E. (2013). Inference of network summary statistics through network denoising. arXiv:1310.0423.
- Bassett, D. S., Porter, M. A., Wymbs, N. F., Grafton, S. T., Carlson, J. M., and Mucha, P. J. (2013). Robust detection of dynamic community structure in networks. *Chaos Interdiscipl. J. Nonlin. Sci.* 23, 013142. doi: 10.1063/1.4790830
- Beckmann, C. F., Jenkinson, M., and Smith, S. M. (2003). General multilevel linear modeling for group analysis in fmri. *Neuroimage* 20, 1052–1063. doi: 10.1016/S1053-8119(03)00435-X
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300.
- Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188. doi: 10.1214/aos/1013699998
- Berk, R., Brown, L., Buja, A., Zhang, K., and Zhao, L. (2013). Valid post-selection inference. *Ann. Stat.* 41, 802–837.
- Bernal-Rusiel, J. L., Greve, D. N., Reuter, M., Fischl, B., Sabuncu, M. R., Initiative, A. D. N., et al. (2013). Statistical analysis of longitudinal neuroimage data with linear mixed effects models. *Neuroimage* 66, 249–260. doi: 10.1016/j.neuroimage.2012.10.065
- Bickel, P. J., and Levina, E. (2008). Regularized estimation of large covariance matrices. *Ann. Stat.* 36, 199–227. doi: 10.1214/009053607000000758
- Birn, R. M., Molloy, E. K., Patriat, R., Parker, T., Meier, T. B., Kirk, G. R., et al. (2013). The effect of scan length on the reliability of resting-state fmri connectivity estimates. *Neuroimage* 83, 550–558. doi: 10.1016/j.neuroimage.2013.05.099
- Braun, U., Plichta, M. M., Esslinger, C., Sauer, C., Haddad, L., Grimm, O., et al. (2012). Test–retest reliability of resting-state connectivity network characteristics using fmri and graph theoretical measures. *Neuroimage* 59, 1404–1412. doi: 10.1016/j.neuroimage.2011.08.044
- Breiman, L. (1996a). Bagging predictors. *Mach. Learn.* 24, 123–140. doi: 10.1007/BF00058655
- Breiman, L. (1996b). Heuristics of instability and stabilization in model selection. *Ann. Stat.* 24, 2350–2383. doi: 10.1214/aos/1032181158
- Buckner, R. L., Krienen, F. M., and Yeo, B. T. (2013). Opportunities and limitations of intrinsic functional connectivity mri. *Nat. Neurosci.* 16, 832–837. doi: 10.1038/nn.3423
- Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., et al. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to alzheimer's disease. *J. Neurosci.* 29, 1860–1873. doi: 10.1523/JNEUROSCI.5062-08.2009
- Bühlmann, P., and Yu, B. (2002). Analyzing bagging. *Ann. Stat.* 30, 927–961. doi: 10.1214/aos/1031689014
- Bullmore, E. (2012). Functional network endophenotypes of psychotic disorders. *Biol. Psychiatry* 71, 844. doi: 10.1016/j.biopsych.2012.03.019
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–349. doi: 10.1038/nrn3214
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., et al. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376. doi: 10.1038/nrn3475
- Cai, T., Liu, W., and Luo, X. (2011). A constrained ℓ_1 minimization approach to sparse precision matrix estimation. *JASA* 106, 594–607. doi: 10.1198/jasa.2011.tm10155
- Chang, C., and Glover, G. H. (2010). Time–frequency dynamics of resting-state brain connectivity measured with fmri. *Neuroimage* 50, 81–98. doi: 10.1016/j.neuroimage.2009.12.011
- Cherkassky, V. L., Kana, R. K., Keller, T. A., and Just, M. A. (2006). Functional connectivity in a baseline resting-state network in autism. *Neuroreport* 17, 1687–1690. doi: 10.1097/01.wnr.0000239956.45448.4c
- Craddock, C. (2014). “Preprocessed connectomes project: open sharing of preprocessed neuroimaging data and derivatives,” in *61st Annual Meeting (AACAP)*, (San Diego, CA: AACAP).
- Craddock, R. C., and Bellec, P. (2015). *Preprocessed Connectomes Project: Abide*. Availbel online at: <http://preprocessed-connectomes-project.github.io/abide/>
- Craddock, R. C., Jbabdi, S., Yan, C.-G., Vogelstein, J. T., Castellanos, F. X., Di Martino, A., et al. (2013). Imaging human connectomes at the macroscale. *Nat. Methods* 10, 524–539. doi: 10.1038/nmeth.2482
- Di Martino, A., Fair, D. A., Kelly, C., Satterthwaite, T. D., Castellanos, F. X., Thomason, M. E., et al. (2014a). Unraveling the miswired connectome: a developmental perspective. *Neuron* 83, 1335–1353. doi: 10.1016/j.neuron.2014.08.050
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014b). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659–667. doi: 10.1038/mp.2013.78
- Di Martino, A., Shehzad, Z., Kelly, C., Roy, A. K., Gee, D. G., Uddin, L. Q., et al. (2009). Relationship between cingulo-insular functional connectivity and autistic traits in neurotypical adults. *Am. J. Psychiatry* 166, 891–899. doi: 10.1176/appi.ajp.2009.08121894
- Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. London, UK: Chapman & Hall.
- Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., et al. (2004). Automatically parcellating the human cerebral cortex. *Cereb. Cortex* 14, 11–22. doi: 10.1093/cercor/bhg087
- Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., and Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10046–10051. doi: 10.1073/pnas.0604187103
- Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441. doi: 10.1093/biostatistics/kxm045
- Giavasis, S. (2015). *Configurable-Pipeline for the Analysis of Connectomes Version 0.3.8*. Available online at: <http://dx.doi.org/10.5281/zenodo.14298>
- Ginestet, C. E., Fournel, A. P., and Simmons, A. (2014). Statistical network analysis for functional mri: summary networks and group comparisons. *Front. Comput. Neurosci.* 8:51. doi: 10.3389/fncom.2014.00051
- Gotham, K., Pickles, A., and Lord, C. (2009). Standardizing ados scores for a measure of severity in autism spectrum disorders. *J. Autism Dev. Disord.* 39, 693–705. doi: 10.1007/s10803-008-0674-3
- Hacker, C. D., Laumann, T. O., Szirmai, N. P., Baldassarre, A., Snyder, A. Z., Leuthardt, E. C., et al. (2013). Resting state network estimation in individual subjects. *Neuroimage* 82, 616–633. doi: 10.1016/j.neuroimage.2013.05.108
- Hahamy, A., Behrmann, M., and Malach, R. (2015). The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nat. Neurosci.* 18, 302–309. doi: 10.1038/nn.3919
- Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10240–10245. doi: 10.1073/pnas.0701519104
- Hsieh, C.-J., Sustik, M. A., Dhillon, I. S., and Ravikumar, P. (2011). Sparse inverse covariance matrix estimation using quadratic approximation. *Adv. Neural Inf. Proces. Syst.* 24, 2330–2338. arXiv:1306.3212
- Hsieh, C.-J., Sustik, M. A., Dhillon, I. S., Ravikumar, P. K., and Poldrack, R. (2013). Big & quic: Sparse inverse covariance estimation for a million variables. *Adv. Neural Inf. Proces. Syst.* 26, 3165–3173.
- Insel, T. R. (2014). The nimh research domain criteria (rdoc) project: precision medicine for psychiatry. *Am. J. Psychiatry* 171, 395–397. doi: 10.1176/appi.ajp.2014.14020138
- Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D., and Pizzagalli, D. A. (2015). Large-scale network dysfunction in major depressive disorder a meta-analysis of resting-state functional connectivity. *JAMA Psychiatry* 72, 603–611. doi: 10.1001/jamapsychiatry.2015.0071

- Keown, C. L., Shih, P., Nair, A., Peterson, N., Mulvey, M. E., and Müller, R.-A. (2013). Local functional overconnectivity in posterior brain regions is associated with symptom severity in autism spectrum disorders. *Cell Rep.* 5, 567–572. doi: 10.1016/j.celrep.2013.10.003
- Kim, J., Wozniak, J. R., Mueller, B. A., Shen, X., and Pan, W. (2014). Comparison of statistical tests for group differences in brain functional networks. *NeuroImage* 101, 681–694. doi: 10.1016/j.neuroimage.2014.07.031
- Koenig, T., Studer, D., Hubl, D., Melie, L., and Strik, W. (2005). Brain connectivity at different time-scales measured with eeg. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1015–1024. doi: 10.1098/rstb.2005.1649
- Lafferty, J., Liu, H., and Wasserman, L. (2012). Sparse nonparametric graphical models. *Statist. Sci.* 27, 519–537. doi: 10.1214/12-STS391
- Lahiri, S. N. (2013). *Resampling Methods for Dependent Data*. New York, NY: Springer Science & Business Media
- Laird, N. M., and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* 38, 963–974. doi: 10.2307/2529876
- Laumann, T. O., Gordon, E. M., Adeyemo, B., Snyder, A. Z., Joo, S. J., Chen, M.-Y., et al. (2015). Functional system and areal organization of a highly sampled individual human brain. *Neuron* 87, 657–670. doi: 10.1016/j.neuron.2015.06.037
- Lauritzen, S. L. (1996). *Graphical Models*, Vol. 17, New York, NY: Oxford University Press.
- Lenroot, R. K., and Yeung, P. K. (2013). Heterogeneity within autism spectrum disorders: what have we learned from neuroimaging studies? *Front. Hum. Neurosci.* 7:733. doi: 10.3389/fnhum.2013.00733
- Liang, K.-Y., and Zeger, S. L. (1993). Regression analysis for correlated data. *Annu. Rev. Public Health* 14, 43–68. doi: 10.1146/annurev.pu.14.050193.000355
- Liu, H., Roeder, K., and Wasserman, L. (2010). Stability approach to regularization selection (stars) for high dimensional graphical models. *Adv. Neural Inf. Process. Syst.* 24, 1432. arXiv:1006.3316.
- Lui, S., Wu, Q., Qiu, L., Yang, X., Kuang, W., Chan, R. C., et al. (2014). Resting-state functional connectivity in treatment-resistant depression. *Am. J. Psychiatry* 168, 642–648. doi: 10.1176/appi.ajp.2010.10101419
- Marrelec, G., Krainik, A., Duffau, H., Pélégriani-Issac, M., Lehericy, S., Doyon, J., et al. (2006). Partial correlation for functional brain interactivity investigation in functional MRI. *Neuroimage* 32, 228–237. doi: 10.1016/j.neuroimage.2005.12.057
- Meda, S. A., Gill, A., Stevens, M. C., Lorenzoni, R. P., Glahn, D. C., Calhoun, V. D., et al. (2012). Differences in resting-state functional magnetic resonance imaging functional network connectivity between schizophrenia and psychotic bipolar probands and their unaffected first-degree relatives. *Biol. Psychiatry* 71, 881–889. doi: 10.1016/j.biopsych.2012.01.025
- Meinshausen, N., and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Stat.* 34, 1436–1462. doi: 10.1214/009053606000000281
- Meinshausen, N., and Bühlmann, P. (2010). Stability selection. *J. R. Stat. Soc. Ser. B* 72, 417–473. doi: 10.1111/j.1467-9868.2010.00740.x
- Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci.* 15, 483–506. doi: 10.1016/j.tics.2011.08.003
- Monk, C. S., Peltier, S. J., Wiggins, J. L., Weng, S.-J., Carrasco, M., Risi, S., et al. (2009). Abnormalities of intrinsic functional connectivity in autism spectrum disorders. *Neuroimage* 47, 764–772. doi: 10.1016/j.neuroimage.2009.04.069
- Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., and Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* 44, 893–905. doi: 10.1016/j.neuroimage.2008.09.036
- Narayan, M., and Allen, G. I. (2013). “Randomized approach to differential inference in multi-subject functional connectivity,” in *Pattern Recognition in Neuroimaging (PRNI), 2013 International Workshop on (IEEE)* (Philadelphia, PA), 78–81. doi: 10.1109/PRNI.2013.29
- Narayan, M., Allen, G. I., and Tomson, S. (2015). Two sample inference for populations of graphical models with applications to functional connectivity. arXiv:1502.03853.
- Palaniyappan, L., Simmonite, M., White, T. P., Liddle, E. B., and Liddle, P. F. (2013). Neural primacy of the salience processing system in schizophrenia. *Neuron* 79, 814–828. doi: 10.1016/j.neuron.2013.06.027
- Penny, W. D., Holmes, A., and Friston, K. (2003). Random effects analysis. *Hum. Brain Funct.* 2, 843–850. doi: 10.1016/B978-012264841-0/50044-5
- Politis, D., Romano, J., and Wolf, M. (1999). *Subsampling*. New York, NY: Springer.
- Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N., and Petersen, S. E. (2013). Evidence for hubs in human functional brain networks. *Neuron* 79, 798–813. doi: 10.1016/j.neuron.2013.07.035
- Ravikumar, P., Wainwright, M. J., Raskutti, G., and Yu, B. (2011). High-dimensional covariance estimation by minimizing ℓ_1 -penalized log-determinant divergence. *Electron. J. Statist.* 5, 935–980. doi: 10.1214/11-EJS631
- Rothman, A. J., Bickel, P. J., Levina, E., and Zhu, J. (2008). Sparse permutation invariant covariance estimation. *Electron. J. Statist.* 2, 494–515. doi: 10.1214/08-EJS176
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Rudie, J., Brown, J., Beck-Pancer, D., Hernandez, L., Dennis, E., Thompson, P., et al. (2012a). Altered functional and structural brain network organization in autism. *NeuroImage* 2, 79–94. doi: 10.1016/j.nicl.2012.11.006
- Rudie, J. D., and Dapretto, M. (2013). Convergent evidence of brain overconnectivity in children with autism? *Cell Rep.* 5, 565–566. doi: 10.1016/j.celrep.2013.10.043
- Rudie, J. D., Shehzad, Z., Hernandez, L. M., Colich, N. L., Bookheimer, S. Y., Iacoboni, M., et al. (2012b). Reduced functional integration and segregation of distributed neural systems underlying social and emotional information processing in autism spectrum disorders. *Cereb. Cortex* 22, 1025–1037. doi: 10.1093/cercor/bhr171
- Sato, J. R., Takahashi, D. Y., Arcuri, S. M., Sameshima, K., Moretton, P. A., and Baccalà, L. A. (2009). Frequency domain connectivity identification: an application of partial directed coherence in fmri. *Hum. Brain Mapp.* 30, 452–461. doi: 10.1002/hbm.20513
- Schelter, B., Winterhalder, M., Eichler, M., Peifer, M., Hellwig, B., Guschlbauer, B., et al. (2006). Testing for directed influences among neural signals using partial directed coherence. *J. Neurosci. Methods* 152, 210–219. doi: 10.1016/j.jneumeth.2005.09.001
- Searle, S. R., Casella, G., and McCulloch, C. E. (2009). *Variance Components*, Vol. 391, Hoboken, NJ: Wiley-Interscience.
- Shehzad, Z., Kelly, A. C., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q., et al. (2009). The resting brain: unconstrained yet reliable. *Cereb. Cortex* 19, 2209–2229. doi: 10.1093/cercor/bhn256
- Simpson, S. L., Bowman, F. D., and Laurienti, P. J. (2013). Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Statist. Surv.* 7, 1–36. doi: 10.1214/13-SS103
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for fmri. *NeuroImage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., et al. (2013). Functional connectomics from resting-state fmri. *Trends Cogn. Sci.* 17, 666–682. doi: 10.1016/j.tics.2013.09.016
- Supekar, K., Uddin, L. Q., Khouzam, A., Phillips, J., Gaillard, W. D., Kenworthy, L. E., et al. (2013). Brain hyperconnectivity in children with autism and its links to social deficits. *Cell Rep.* 5, 738–747. doi: 10.1016/j.celrep.2013.10.001
- Tam, A., Orban, P., Dansereau, C., Belleville, S., Benoit, S., Dagher, A., et al. (2014). Connectome-wide analysis of differences between normal aging, mild cognitive impairment, and dementia of the alzheimer’s type using resting-state fmri connectivity. *Alzheimer Dement.* 10, P827–P828. doi: 10.1016/j.jalz.2014.05.089
- Tao, H., Guo, S., Ge, T., Kendrick, K. M., Xue, Z., Liu, Z., et al. (2013). Depression uncouples brain hate circuit. *Mol. Psychiatry* 18, 101–111. doi: 10.1038/mp.2011.127
- Tomson, S. N., Narayan, M., Allen, G. I., and Eagleman, D. M. (2013). Neural networks of colored sequence synesthesia. *J. Neurosci.* 33, 14098–14106. doi: 10.1523/JNEUROSCI.5131-12.2013
- Uddin, L. Q. (2014). Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* doi: 10.1038/nrn3857
- Uddin, L. Q., Supekar, K., Lynch, C. J., Khouzam, A., Phillips, J., Feinstein, C., et al. (2013a). Salience network-based classification and prediction of symptom severity in children with autism. *JAMA Psychiatry* 70, 869–879. doi: 10.1001/jamapsychiatry.2013.104

- Uddin, L. Q., Supekar, K., and Menon, V. (2013b). Reconceptualizing functional brain connectivity in autism from a developmental perspective. *Front. Hum. Neurosci.* 7:458. doi: 10.3389/fnhum.2013.00458
- van den Heuvel, M. P., and Sporns, O. (2011). Rich-club organization of the human connectome. *J. Neurosci.* 31, 15775–15786. doi: 10.1523/JNEUROSCI.3539-11.2011
- van den Heuvel, M. P., Sporns, O., Collin, G., Scheewe, T., Mandl, R. C., Cahn, W., et al. (2013). Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry* 70, 1–10. doi: 10.1001/jamapsychiatry.2013.1328
- Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., and Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J. Neurophysiol.* 103, 297–321. doi: 10.1152/jn.00783.2009
- Varoquaux, G., Gramfort, A., Poline, J., and Thirion, B. (2012). Markov models for fmri correlation structure: is brain functional connectivity small world, or decomposable into networks? *J. Physiol.* 106, 212–221. doi: 10.1016/j.jphysparis.2012.01.001
- Wang, S., Nan, B., Rosset, S., and Zhu, J. (2011). Random lasso. *Ann. Appl. Stat.* 5, 468. doi: 10.1214/10-AOAS377
- Warren, D. E., Power, J. D., Bruss, J., Denburg, N. L., Waldron, E. J., Sun, H., et al. (2014). Network measures predict neuropsychological outcome after brain injury. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14247–14252. doi: 10.1073/pnas.1322173111
- Williams, D. A. (1982). Extra-binomial variation in logistic linear models. *Appl. Statist.* 31, 144–148. doi: 10.2307/2347977
- Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fmri data. *Neuroimage* 14, 1370–1386. doi: 10.1006/nimg.2001.0931
- Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G., Morales, F., et al. (2002). A general statistical analysis for fmri data. *Neuroimage* 15, 1–15. doi: 10.1006/nimg.2001.0933
- Yang, E., Lozano, A. C., and Ravikumar, P. K. (2014). Elementary estimators for graphical models. *Adv. Neural Inf. Proces. Syst.* 2159–2167.
- Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.*, 106, 1125–1165. doi: 10.1152/jn.00338.2011
- Yuan, M., and Lin, Y. (2007). Model selection and estimation in the gaussian graphical model. *Biometrika* 94, 19–35. doi: 10.1093/biomet/asm018
- Zhou, S., Rütimann, P., Xu, M., and Bühlmann, P. (2011). High-dimensional covariance estimation based on gaussian graphical models. *J. Mach. Learn. Res.* 12, 2975–3026. arXiv:1009.0530
- Zhou, S. et al. (2014). Gemini: graph estimation with matrix variate normal instances. *Ann. Stat.* 42, 532–562. doi: 10.1214/13-AOS1187
- Zuo, X.-N., Ehmke, R., Mennes, M., Imperati, D., Castellanos, F. X., Sporns, O., et al. (2012). Network centrality in the human functional connectome. *Cereb. Cortex* 22, 1862–1875. doi: 10.1093/cercor/bhr269

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Narayan and Allen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



An Efficient and Reliable Statistical Method for Estimating Functional Connectivity in Large Scale Brain Networks Using Partial Correlation

Yikai Wang¹, Jian Kang², Phebe B. Kemmer¹ and Ying Guo^{1*}

¹ Department of Biostatistics and Bioinformatics, The Rollins School of Public Health, Emory University, Atlanta, GA, USA,

² Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA

OPEN ACCESS

Edited by:

Brian Caffo,
Johns Hopkins University, USA

Reviewed by:

Baxter P. Rogers,
Vanderbilt University, USA
Xi Luo,
Brown University, USA

*Correspondence:

Ying Guo
yguo2@emory.edu

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 30 November 2015

Accepted: 13 March 2016

Published: 31 March 2016

Citation:

Wang Y, Kang J, Kemmer PB and
Guo Y (2016) An Efficient and Reliable
Statistical Method for Estimating
Functional Connectivity in Large Scale
Brain Networks Using Partial
Correlation. *Front. Neurosci.* 10:123.
doi: 10.3389/fnins.2016.00123

Currently, network-oriented analysis of fMRI data has become an important tool for understanding brain organization and brain networks. Among the range of network modeling methods, partial correlation has shown great promises in accurately detecting true brain network connections. However, the application of partial correlation in investigating brain connectivity, especially in large-scale brain networks, has been limited so far due to the technical challenges in its estimation. In this paper, we propose an efficient and reliable statistical method for estimating partial correlation in large-scale brain network modeling. Our method derives partial correlation based on the precision matrix estimated via Constrained L1-minimization Approach (CLIME), which is a recently developed statistical method that is more efficient and demonstrates better performance than the existing methods. To help select an appropriate tuning parameter for sparsity control in the network estimation, we propose a new *Dens*-based selection method that provides a more informative and flexible tool to allow the users to select the tuning parameter based on the desired sparsity level. Another appealing feature of the *Dens*-based method is that it is much faster than the existing methods, which provides an important advantage in neuroimaging applications. Simulation studies show that the *Dens*-based method demonstrates comparable or better performance with respect to the existing methods in network estimation. We applied the proposed partial correlation method to investigate resting state functional connectivity using rs-fMRI data from the Philadelphia Neurodevelopmental Cohort (PNC) study. Our results show that partial correlation analysis removed considerable between-module marginal connections identified by full correlation analysis, suggesting these connections were likely caused by global effects or common connection to other nodes. Based on partial correlation, we find that the most significant direct connections are between homologous brain locations in the left and right hemisphere. When comparing partial correlation derived under different sparse tuning parameters, an important finding is that the sparse regularization has more shrinkage effects on negative functional connections than on positive connections, which supports previous findings that many of the negative brain connections are due to non-neurophysiological effects. An R package “DensParcorr” can be downloaded from CRAN for implementing the proposed statistical methods.

Keywords: network analysis, functional connectivity, fMRI, partial correlation, precision matrix, CLIME, L1 regularization

INTRODUCTION

In recent years, network-oriented analyses have shown great promise for understanding brain organization and its involvement in mental disorders. With the advancement of neuroimaging technologies, the study of whole-brain functional connectivity analysis using functional magnetic resonance imaging (fMRI) data has stimulated an enormous amount of interest (Biswal et al., 1995; Bullmore and Sporns, 2009; Deco et al., 2011; Satterthwaite et al., 2015; Zhang et al., 2015). In particular, there has been a strong focus on investigating intrinsic brain connectivity using resting-state fMRI (rs-fMRI), which measures the spontaneous low-frequency fluctuations in the blood oxygen level dependent (BOLD) signal in subjects at rest (Ogawa et al., 1990; Dosenbach et al., 2010).

Various methods have been proposed for assessing the brain connectivity between selected network nodes. One of the simplest and most frequently used methods in the neuroimaging community is via pairwise correlations between BOLD time courses from two brain network nodes. These correlations are of great interest to neuroscientists in that they can reflect the functional connectivity between brain regions and help explore the overall network structure of the whole brain (Church et al., 2009; Seeley et al., 2009).

However, there are well-known limitations in the correlation analysis. Pearson correlation, which we will henceforth refer to as “full correlation,” only reflects the marginal association between network nodes and is not an appropriate tool for capturing the true or direct functional connection between them. For example, a large correlation between a pair of nodes can appear due to their common connections to a third-party node, even if the two nodes are not directly connected (Smith et al., 2011). Using full correlation, investigators often identify significant connections between a large number of node pairs in brain networks. It is difficult to distinguish which of these significant correlations reflect true functional connections and which are caused by confounding factors such as global effects or third-party nodes.

A network modeling method that has shown great potential in addressing this major issue is partial correlation (Smith, 2012). Partial correlation measures the direct connectivity between two nodes by estimating their correlation after regressing out effects from all the other nodes in the network, hence avoiding spurious effects in network modeling. A partial correlation value of zero implies an absence of direct connections between two nodes given all the other nodes. Through a set of extensive and realistic simulation studies, Smith et al. (2011) compared the performance of a wide range of network modeling methods for fMRI data and found that partial correlation is among the top methods that performed excellently under various types of scenarios and showed high sensitivity to detect true functional connections.

Although it has been shown to have major advantages in studying brain connectivity, the application of partial correlation in the neuroimaging community has been limited. This is mainly because the estimation of partial correlation is more difficult than full correlation. Direct estimation based on the regression approach is inefficient in terms of computational time and often fails due to the multicollinearity among node time series. A

more efficient way to estimate the full set of partial correlations is via the inverse of the covariance matrix, also known as the precision matrix (Marrelec et al., 2006), where the off-diagonals of a precision matrix have a one-to-one correspondence with partial correlations (Peng et al., 2009).

Estimation of the precision matrix is not a trivial task since it involves the inversion of the covariance matrix, especially for a large dimensional case. Furthermore, a precision matrix needs to satisfy the positive definite condition which further increases difficulty in its estimation. In neuroimaging applications, this task could become even more challenging because there are often a large number of nodes in brain networks and a limited number of observations at each node (e.g., shorter fMRI scans) (Zhang et al., 2015). Under this setting, estimation of the precision matrix requires a huge computational load and may not be stable. A few methods have been developed for this purpose in the neuroimaging community (Schmittmann et al., 2015). Schäfer and Strimmer (2005) developed a shrinkage approach to estimate the covariance matrix. Moore-Penrose inverse of the covariance matrix can also be applied to directly estimate the precision matrix (Ben-Israel and Greville, 2003). Moreover, the most popular approach is to apply the sparse regularization via the L1 penalty (Meinshausen and Bühlmann, 2006; Friedman et al., 2008; Liu and Luo, 2012) in estimating the precision matrix under the normality assumption (Smith et al., 2011). As an extension, several works were proposed to relax the normality assumption for graphical models (Liu et al., 2012; Han et al., 2013). The existing approaches can become quite time consuming when the dimension of the precision matrix becomes high. Furthermore, based on our experiments, when estimating large-scale brain networks, the existing computational tools used in the community often either have computational issues or lack of accuracy in capturing some key features in brain organization. Finally, the sparse regularization estimation usually requires the selection of a tuning parameter to control the sparsity of the estimated precision matrix, and the results vary significantly depending on the choice. Currently, the selection of the tuning parameter is often fairly subjective in applications.

In this paper, we present a more efficient and reliable statistical procedure for estimating partial correlation in brain network modeling under the regularized precision matrix framework. The proposed procedure first estimates the precision matrix via the Constrained L1-minimization Approach (CLIME) (Cai et al., 2011). Compared with other regularization methods such as Lasso, CLIME is shown to have better theoretical properties as well as computational advantages. Theoretically, CLIME precision matrix estimators are shown to converge to the true precision matrix at a faster rate as compared to the traditional L1 regularization methods. Computationally, CLIME can be easily implemented by linear programming and is scalable to a high dimensional precision matrix with a large number of nodes. As with the other regularization methods, CLIME requires the setting of a tuning parameter for controlling the sparsity. The existing selection methods often face challenges in estimating large-scale brain networks in that they either tend to select overly dense networks or are computationally expensive. To address this issue, we propose a method to provide a systematic approach that

allows the users to make a more informed choice of the tuning parameter. Specifically, we propose a *Dens* criterion function that reflects how dense the estimated precision matrix is under various tuning parameters. Then by setting a desired density level one would like to achieve, the users can find the appropriate tuning parameter to use for CLIME. The proposed *Dens*-based selection method is easy to implement, computationally much faster than existing methods, and provides users the flexibility to control the sparsity of the estimated precision matrix. Simulation studies show that our *Dens*-based method demonstrates similar or better accuracy in estimating the precision matrix as compared to the more complicated and computationally expensive selection methods. We also show via a real fMRI data example that the selection of the tuning parameter based on the proposed method is highly consistent across subjects. After estimating the precision matrix using CLIME with the chosen tuning parameter, we provide the formula for deriving the partial correlation matrix from the precision matrix.

We apply the proposed partial correlation estimation procedure to investigate direct brain functional connectivity using resting state fMRI data collected in the Philadelphia Neurodevelopmental Cohort (PNC) study (Satterthwaite et al., 2014). We compare the direct brain connectivity pattern based on partial correlation with the marginal brain connectivity based on full correlation. We examine edges in the brain network that are consistently identified by both the partial correlation and full correlation method vs. edges for which the two methods show inconsistent results. Additionally, we examine how the partial-correlation-based direct connectivity networks change when we impose different levels of sparsity in the estimated network.

METHODS

Partial Correlation: Definition and Derivation

In this section, we first introduce the concept and definition of partial correlation under the brain network modeling framework. To set notation, let $\mathbf{X} = \{X_1, \dots, X_M\}$ denote the fMRI BOLD signal at M nodes ($M \times 1$ vector) in the network in an fMRI scan. Let \mathbf{X}_t , $t = 1, \dots, T$, denote the T realizations of \mathbf{X} in fMRI scans obtained during a scanning session. Partial correlation between nodes i and j is defined as the correlation between X_i and X_j conditioning on all the other nodes, i.e.:

$$\rho_{ij} = \text{corr}(X_i, X_j | X_{-(i,j)}), X_{-(i,j)} = \{X_k : 1 \leq k \leq M, k \neq i, j\}, \\ i, j = 1, \dots, M, i \neq j.$$

In the context of brain networks, partial correlation is the correlation between time series of two nodes, after adjusting for the time series from all other network nodes (Smith et al., 2011). As an example, consider a simple three-node network ($M = 3$). To derive the partial correlation between nodes 1 and 2, we first regress the time series of node 1 against the time series of node 3 and denote the residual as $\mathbf{R}_{1|3}$, then regress the time series of node 2 against the time series of node 3 and denote the residual as $\mathbf{R}_{2|3}$; the partial correlation between node 1 and 2 can then be obtained as the correlation between $\mathbf{R}_{1|3}$ and $\mathbf{R}_{2|3}$.

In addition to the derivation based on linear regression, partial correlation can also be derived from the inverse covariance matrix, also known as the precision matrix. Let Σ be the $M \times M$ covariance matrix based on \mathbf{X} and let $\Omega = \Sigma^{-1} = \{\omega_{ij}\}_{M \times M}$ be the precision matrix. The partial correlation between node i and j can be derived from the precision matrix as Peng et al. (2009):

$$\rho_{ij} = -\omega_{ij} / \sqrt{\omega_{ii}\omega_{jj}}. \quad (1)$$

Under the Gaussian assumption, one can infer that node i and j are conditionally independent given the other nodes when ρ_{ij} equals 0. Therefore, partial correlation provides a way to assess the direct connection between nodes and allows correct estimation of the true network by removing all the confounding effects (Smith et al., 2011).

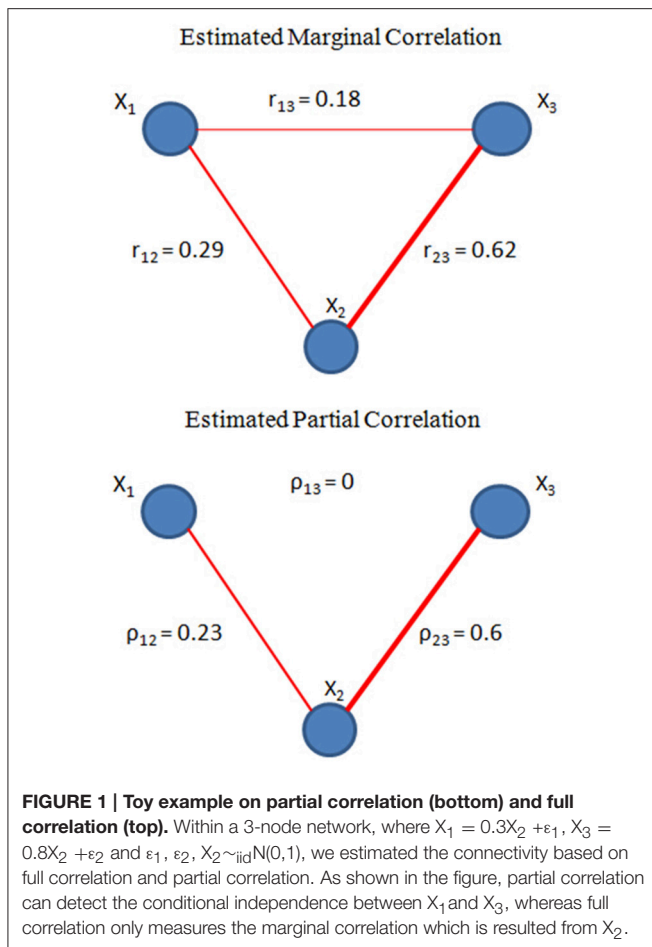
To illustrate the difference between full correlation and partial correlation, we provide a toy example using a 3-node network. X_1, X_2, X_3 represent the measurements from the 3 nodes, where

$$X_1 = \alpha_1 X_2 + \varepsilon_1, X_3 = \alpha_2 X_2 + \varepsilon_2, \varepsilon_1, \varepsilon_2, X_2 \sim_{\text{iid}} N(0, 1). \quad (2)$$

Here both X_1 and X_3 are directly associated with X_2 , but X_1 and X_3 are not directly related to each other given X_2 . We then estimated both the full correlation and partial correlation based on the time series generated from (2) with $\alpha_1 = 0.3$ and $\alpha_2 = 0.8$. The results are presented in **Figure 1**. Both correlation methods were able to detect the true connectivity between nodes 1 and 2, and between nodes 2 and 3. However, for nodes 1 and 3, the full correlation estimate implies that they were also associated. From the data-generating model (2), we know that this association is not due to the true connection between nodes 1 and 3 but rather caused by their common connection with node 2. The partial correlation estimate for this connection had a value of zero, correctly reflecting that there was no direct connection between nodes 1 and 3. This toy example demonstrates the ability of partial correlation in removing spurious associations due to a third-party node, and hence provides a more reliable measure for direct connectivity in brain networks.

The Proposed Procedure for Estimating Partial Correlation Using Neuroimaging Data

Unlike full correlation which can be readily calculated from the observed fMRI data, the estimation of partial correlation is less straightforward and more computationally challenging. The precision matrix method provides an efficient way to obtain the full set of partial correlations between all node pairs in a network. However, since estimating the precision matrix commonly involves inverting the covariance matrix, this approach becomes challenging as the number of nodes (and the dimension of the covariance matrix) increases. In particular, direct inversion of the covariance matrix is not feasible when the number of nodes is larger than the number of observations at each node, such as the case of estimating large-scale brain networks in relatively short fMRI scanning sessions. Various approaches based on regularization methods such as Graphical lasso have been applied to address this issue in neuroimaging studies (Friedman et al.,



2008). The issues with the existing approaches are that they require long computation time and often fail when the number of nodes is large. Another difficulty is that the regularization methods require the selection of a tuning parameter to control the sparsity of the estimated precision matrix, and in current neuroimaging applications, this selection is often conducted in a fairly subjective manner.

In this section, we propose a new statistical procedure for estimating the partial correlations in a brain network. Our proposed procedure consists of three parts: (1) estimating the precision matrix using Constrained L1-minimization for Inverse Matrix Estimation (CLIME), which is a recently developed statistical method that is computationally more efficient and demonstrates better performance as compared to many existing algorithms; (2) choosing the tuning parameter for the CLIME algorithm based on our proposed *Dens*-based method, which is fast and can be easily understood and controlled by the users; and (3) deriving the full set of partial correlations from the estimated precision matrix.

A Constrained L1 Approach (CLIME) to Sparse Precision Matrix Estimation

The CLIME method is an approach that has been recently developed in the statistical community for estimating a sparse

precision matrix (Cai et al., 2011). The CLIME estimator of the precision matrix Ω is derived using the following procedure. First, we find the solution Ω^1 of the following optimization problems:

$$\Omega^1 = \arg \min \|\Omega\|_1 \text{ subject to } |\hat{\Sigma}\Omega - \mathbf{I}|_\infty \leq \lambda, \quad (3)$$

here, $\Omega^1 = \{\tilde{\omega}_{ij}\}_{M \times M}$ is an initial estimator of the precision matrix Ω , $\hat{\Sigma}$ is the estimated covariance matrix, λ is a tuning parameter ranging from 0 to 1, where a larger λ imposes a stronger sparsity regularization and hence yields a more sparse Ω^1 . Because Ω^1 is not necessarily symmetric, the final CLIME estimator $\hat{\Omega}_*$ is obtained by symmetrizing Ω^1 as follows.

$$\hat{\Omega}_* = \{\hat{\omega}_{ij}\}_{M \times M}, \quad (4)$$

with $\hat{\omega}_{ij} = \min(\tilde{\omega}_{ij}, \tilde{\omega}_{ji})$.

A unique feature of the CLIME method is that it develops an approach to solve the convex program (3) by decomposing it into M vector minimization problems that estimate each column of Ω^1 one at a time. It can be shown that solving the optimization problem in (3) is equivalent to solving the M vector minimization problem, which can be achieved via linear programming. By estimating the precision matrix column-by-column, CLIME significantly reduces the computational and statistical difficulties in its estimation. Another appealing feature is that the final CLIME estimator $\hat{\Omega}_*$ is shown to be positive definite with high probability (Cai et al., 2011). This means that the CLIME method has a high chance of producing a valid precision matrix estimate for brain network modeling.

Regularization Selection

As with other regularization methods, the CLIME approach also requires the specification of a tuning parameter, i.e., λ in (3). This parameter controls the sparsity of the estimated precision matrix and the subsequent estimate of the partial correlation matrix. An advantage of the CLIME method is that the tuning parameter is selected within the finite range of 0–1, whereas the tuning parameter in other regularization methods does not have a finite range. For example, graphical lasso involves a tuning parameter that ranges from 0 to 8. From (3), a smaller λ yields a denser graph and larger λ yields a sparser graph. When λ approaches toward 1, which means imposing strongest sparsity regularization, Ω^* will approach an empty matrix which corresponds to an empty network without any edges. When λ approaches toward 0, the minimum sparsity regularization, $\hat{\Omega}_*$ will approach the precision matrix estimate that is obtained without the sparsity constraint.

Two common ways to select the tuning parameter in regularization methods are AIC and BIC (Schwarz, 1978; Akaike, 1998). Let $\hat{\Omega}_*(\lambda)$ be the estimated precision matrix based on tuning parameter λ . AIC selects λ such that:

$$\hat{\lambda} = \arg \min_{\lambda} \{-2 \log |\hat{\Omega}_*(\lambda)| + 2 \text{trace}(\hat{\Sigma} \hat{\Omega}_*(\lambda)) + 2d(\lambda)\},$$

and BIC selects λ such that:

$$\hat{\lambda} = \arg \min_{\lambda} \{-2 \log |\hat{\Omega}_*(\lambda)| + 2 \text{trace}(\hat{\Sigma} \hat{\Omega}_*(\lambda)) + d(\lambda) \log(T)\}.$$

Here, $d(\lambda)$ denotes the degrees of freedom of the underlying Gaussian model. The $d(\lambda)$ is difficult to estimate in the high-dimensional setting where the number of nodes in the network (M) exceeds the number of observations (T) at each node. In this case, the $d(\lambda)$ is often estimated by the number of non-zero elements in $\hat{\Omega}_*(\lambda)$. It has been shown that AIC and BIC methods tend to yield an overly dense precision matrix in the high-dimensional case (Liu et al., 2010).

Another commonly used method for selecting λ is the k-fold cross-validation (K-CV) method (Efron, 1982). In this type of procedure, the observed data are partitioned into k blocks, where $k-1$ blocks are used as training data to estimate the precision matrix and the remaining block is retained as validation data. For each λ value in the search grid, one estimates the precision matrix and corresponding partial correlations using the $k-1$ blocks of training data and then evaluates a loss function of the estimates using the validation data. Two typical loss functions are the negative log-likelihood and Trace L2 defined below.

Negative log-likelihood: $-\log|\hat{\Omega}_*(\lambda)| + \text{trace}(\hat{\Sigma}\hat{\Omega}_*(\lambda)) - M$

Trace L2: $\text{trace}(\text{diag}(\hat{\Sigma}\hat{\Omega}_*(\lambda) - \mathbf{I}_M)^2)$

The K-CV methods based on these two loss functions are implemented in the CLIME R package (Cai et al., 2012). One issue with K-CV methods is that they are typically computationally expensive. Furthermore, it has been shown that K-CV based on the negative log-likelihood loss function tends to select overly dense graphs (Wasserman and Roeder, 2009).

In this paper, we present a new method for selecting λ . Specifically, we propose a *Dens* criterion function that measures how dense the estimated precision matrix is. Then we consider a series of λ within the finite range $(0, 1)$. We start with a large value of λ which results in an extremely sparse graph with little or no edges, then decrease λ so that the estimated precision matrix becomes denser and more edges are allowed to appear in the graph. We continue to decrease λ until the density of the precision matrix, measured by the proposed criterion function, reaches its plateau and remains stable. Finally, we examine the profile of the *Dens* criterion function across the series of λ values and select the value of λ that corresponds to the desired density level that the investigator would like to achieve.

To measure how dense an estimated precision matrix is, we propose the following *Dens* criterion function:

$$Dens(\Omega) = \sum_{ij} |\omega_{ij}|, \text{ where } \Omega = \{\omega_{ij}\}. \quad (5)$$

That is, *Dens* is the sum of the absolute values of all elements in the estimated precision matrix, and measures the density level of the precision matrix. Essentially, *Dens* is the matrix-wise L1 norm of Ω .

For the CLIME procedure, we consider a monotonically decreasing sequence $\{\lambda_n, n = 0, 1, \dots\}$ within the range $(0, 1)$ with $\lambda_0 \rightarrow 1$ and $\lambda_n \rightarrow 0$ as n increases. For simplicity, we denote $Dens(\hat{\Omega}_*(\lambda))$ as $Dens(\lambda)$. For $\lambda_0 \rightarrow 1$, the CLIME estimator $\hat{\Omega}_*(\lambda)$ approaches a zero matrix which corresponds to

an empty network without any edges; hence, $Dens(\lambda_0)$ is close to zero. As λ_n decreases, $\hat{\Omega}_*(\lambda)$ becomes denser and more elements become non-zero, resulting in the increase in $Dens(\lambda_n)$. As n increases and $\lambda_n \rightarrow 0$, $Dens(\lambda_n)$ reaches a plateau and becomes stabilized with further decrease in λ_n . With a finite sequence of $\{\lambda_n\}$ we can find the maximum of $Dens(\lambda_n)$, and denote it as $Dens_{max}$. In practice, it is not necessary to select λ_{max} that corresponds to $Dens_{max}$, because it is somewhat arbitrary and depends on the smallest value specified in the finite sequence of $\{\lambda_n\}$. Instead, based on the profile of $Dens(\lambda_n)$, users can choose the value in the sequence that corresponds to the plateau point in the profile, which is denoted as λ_{platu}^* . After λ_{platu}^* , $Dens(\lambda)$ becomes stabilized and only increases by a trivial amount when further decreasing the tuning parameter. Specifically, we define λ_{platu}^* as the largest λ_n in the sequence such that for any $\lambda_k \leq \lambda_n$, we have

$$\frac{|Dens(\lambda_k) - Dens_{max}|}{Dens_{max}} \leq \varepsilon,$$

where ε is a user-specified small value such as 0.01. Since the estimated network is close to the maximum density level at λ_{platu}^* , $\hat{\Omega}_*(\lambda_{platu}^*)$ corresponds to the estimate of the precision matrix that is obtained under minimum sparsity constraint.

As the number of nodes in the network increases, it may be desirable to impose a certain sparsity regularization to reduce the number of false positive edges in the estimated precision matrix. In this case, we propose the following method to select the tuning parameter based on a user-specified *Dens* level for the precision matrix estimate. Suppose the user would like to obtain an precision matrix estimate that would reach p percent of the maximum density level, that is $Dens(\lambda_n) = p \times Dens_{max}$, then the corresponding tuning parameter λ_p^* can be selected from $\{\lambda_n\}$ as follows:

$$\lambda_p^* = \arg\min_{\lambda_n} \{|Dens(\lambda_n) - p \times Dens_{max}|\}. \quad (6)$$

After we select the tuning parameter and obtain the CLIME estimate Ω^* of the precision matrix, we can derive the partial correlation matrix estimate, $Pcorr = \{\rho_{ij}\} M \times M$, via the following equation:

$$Pcorr = -\text{diag}(\Omega)^{-1/2} \Omega \text{diag}(\Omega)^{-1/2} + 2\mathbf{I}_M. \quad (7)$$

In summary, we have proposed a novel tuning parameter selection criterion for the sparse precision matrix estimation in brain network modeling. A detailed summary of the steps of our procedure is provided in **Table 1**.

SIMULATION STUDIES AND RESULTS

In this section, we investigate the empirical performance of the proposed tuning parameter selection method using synthetic data. We simulated spatially- and temporally-dependent data that mimic real fMRI data. Specifically, to induce spatial dependence between the nodes, we generated data from specified networks and considered various sparsity levels for the network.

TABLE 1 | Proposed *Dens*-based partial correlation estimation approach.**Summary steps:**

Input: Estimate the sample covariance matrix $\hat{\Sigma}$ based on the observed fMRI time series from M nodes in the brain. If one would like to impose sparsity regularization on the precision matrix estimate, specify a percentage p , where $p \in (0, 1)$, for selecting the tuning parameter based on the desired density level of the precision matrix estimate.

Step 1, Select the Tuning Parameter

- Specify a monotonically decreasing sequence $\{\lambda_n, n = 0, 1, \dots\}$ within the range $(0, 1)$ with $\lambda_0 \rightarrow 1$ and $\lambda_1 \rightarrow 0$ as n increases.
- Obtain CLIME estimate based on each value in $\{\lambda_n\}$ starting from λ_0 . Keep decreasing λ_n until $Dens(\lambda_n)$ reaches its plateau and remains stable afterwards.
- Denote the maximum $Dens(\lambda_n)$ in its profile as $Dens_{\max}$ and the tuning parameter that corresponds to the plateau point in the profile as λ_{platu} .
- If the user specifies a percentage p that represents the desired $Dens$ level, find the corresponding tuning parameter λ_{*p} from the sequence $\{\lambda_n\}$ based on (4).

Step 2, Estimate the precision matrix using CLIME

Based on the selected tuning parameter λ , obtain CLIME estimate $\hat{\Omega}(\lambda)$ through the procedure in (3) and (4)

Step3, Derive estimate for the partial correlation matrix

Obtain \hat{P}_{corr} from $\hat{\Omega}(\lambda)$ using Equation (7)

We then evaluated the performance of the proposed tuning parameter selection method based on the *Dens* criterion and compared that to the existing selection methods.

Synthetic Data

We generated time series data for M nodes over T time points. Real fMRI data, which are collected over a series of time points, demonstrate both temporal and spatial dependence. In order to mimic this complex covariance structure, we first specified a precision matrix Ω that represent the network connectivity among the M nodes, the spatial covariance matrix Σ_s can be derived from Ω . We then induced temporal correlation in the node time series via an AR(1) model. The detailed procedure is presented as follows. Let \mathbf{Y} be the $T \times M$ data matrix. Based on a pre-specified precision matrix Ω , \mathbf{Y}_Ω were generated as:

$$\mathbf{Y}_\Omega = \mathbf{X} + \mathbf{Z} \quad (8)$$

where $\mathbf{X} = \{\mathbf{X}_1, \dots, \mathbf{X}_M\}^T$ was a $T \times M$ matrix where each row \mathbf{X}_i 's $\sim_{\text{iid}} N_M(0, \Sigma_s)$. Here $\Sigma_s = \Omega^{-1} - \tau^2 \mathbf{I}_M$ is the spatial covariance matrix derived from Ω . $\mathbf{Z} = \{\mathbf{Z}_1, \dots, \mathbf{Z}_M\}$ was also a $T \times M$ matrix where each column \mathbf{Z}_i 's $\sim_{\text{iid}} N_T(0, \Sigma_T)$ with $\Sigma_T = \{\Sigma_{T,ij}\} = \{\tau^2 \gamma^{|i-j|}\}$ being the temporal covariance matrix based on an AR(1) model.

In the data generation model (8), \mathbf{X} induces the spatial covariance structure in the data which is controlled by the precision matrix, and \mathbf{Z} induces the temporal correlations in the data which are AR(1) time series with variance τ^2 and adjacent correlation γ . In order to ensure that the spatial covariance matrix Σ_s is positive definite, the variance τ^2 is set to be half of the inversed largest eigenvalue of Ω . As a result, \mathbf{Y}_Ω generated from (8) has a matrix normal distribution and the precision matrix of \mathbf{Y}_Ω in the spatial domain is Ω .

In our simulation, we generated data from (8) with $M = 10$ and $T = 50$. To examine the performance of the proposed *Dens* criterion under various scenarios, we considered 9 sparsity levels ranging from 0.29 to 0.93, where the sparsity level represents the percentage of non-zero elements in the off-diagonal. For each scenario, we had 100 simulation runs.

In the next section, we evaluated the performance of the proposed *Dens*-based regularization selection method, and compared to four existing selection methods including the AIC,

BIC, and K-CV approaches with the negative log likelihood and trace L2 loss functions. For our proposed *Dens*-based selection method, we selected three tuning parameters corresponding to different density levels: λ_{platu}^* which leads to an estimate which corresponds to the plateau point in the *Dens* profile, and $\lambda_{0.45}^*$ and $\lambda_{0.75}^*$ which lead to estimates that reach 45 and 75% of the maximum density level, respectively. For K-CV methods, we used 5-fold cross validation for selecting λ .

To evaluate the performance of the various methods in estimating the partial correlation matrix, we calculated the MSE, sensitivity, and specificity by comparing the true and estimated partial correlations from different methods. Here, the MSE is obtained as the average MSE across all off-diagonal edges in the partial correlation matrix.

Results from the Simulation Study

We present detailed simulation results for each of the 9 sparsity levels in **Tables 2–4**. We also present the average results across all sparsity levels and as well as the average computation time for these methods in **Tables 5, 6**.

Compared with the existing methods, the proposed *Dens*-based method is much more computationally efficient, especially compared to the K-CV methods (**Tables 5, 6**). The computational efficiency provides an important advantage in estimating brain networks based on high-dimensional fMRI data. In addition, our proposed method provided the most accurate estimation in terms of the average MSE and the number of times it achieved the lowest MSE value across different sparsity levels (see **Tables 2, 5**). This indicates that our method has better accuracy, on average, across different sparsity levels. In terms of sensitivity and specificity, AIC, BIC, and K-CV with a negative log likelihood loss function tended to select an overly dense network with extremely low specificity, which was consistent with previous findings in the literature. In comparison, K-CV based on Trace L2 loss function provided more balanced performance in terms of sensitivity and specificity. For our method, λ_{platu}^* also tended to select an overly dense network, which is expected since it imposes the minimum sparsity regularization. For the $\lambda_{0.45}^*$ and $\lambda_{0.75}^*$ which applied sparsity constraints, we achieved much better balance between sensitivity and specificity. In particular, $\lambda_{0.45}^*$ offers the best average of sensitivity and specificity at

TABLE 2 | Comparison of MSE for estimated partial correlation matrix based on different regularization selection methods across various sparsity levels with the simulated data.

Methods	Sparsity level								
	0.29	0.36	0.47	0.49	0.56	0.76	0.87	0.89	0.93
K-CV log like	0.016	0.031	0.026	0.014	0.015	0.014	0.019	0.011	0.010
K-CV Trace L2	0.012	0.018	0.017	0.009	0.010	0.011	0.011	0.011	0.011
AIC	0.016	0.031	0.026	0.014	0.015	0.014	0.019	0.011	0.010
BIC	0.016	0.031	0.026	0.014	0.015	0.014	0.019	0.011	0.010
$\lambda^*_{0.45}$	0.010	0.015	0.012	0.006	0.010	0.012	0.011	0.015	0.016
$\lambda^*_{0.75}$	0.010	0.020	0.017	0.008	0.009	0.009	0.012	0.009	0.008
λ^*_{platu}	0.014	0.027	0.023	0.012	0.012	0.012	0.016	0.010	0.009

Based on simulated data, we examined our proposed *Dens* method and commonly used regularization selection methods including K-CV with negative log likelihood, Trace L2, AIC, BIC. For *Dens* method, we adopt three different density level: 45, 75% and plateau, corresponding to $\lambda^*_{0.45}$, $\lambda^*_{0.75}$, and λ^*_{platu} separately. The MSE values in bold are the optimal result across the different methods at each sparsity level.

TABLE 3 | Comparison of Sensitivity for identifying connections based on different regularization selection methods across various sparsity levels with the simulated data.

Methods	Sparsity level								
	0.29	0.36	0.47	0.49	0.56	0.76	0.87	0.89	0.93
K-CV log like	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
K-CV Trace L2	0.95	0.86	0.81	0.86	0.85	0.83	0.81	0.86	0.84
AIC	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BIC	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$\lambda^*_{0.45}$	0.85	0.79	0.63	0.74	0.67	0.65	0.62	0.57	0.61
$\lambda^*_{0.75}$	0.96	0.92	0.86	0.93	0.87	0.91	0.90	0.91	0.93
λ^*_{platu}	0.99	0.97	0.96	0.98	0.97	0.98	0.97	0.99	0.99

0.706, which is much higher than those of the four existing methods (see **Table 5**). In summary, our proposed *Dens*-based method provided comparable or better performance with respect to the existing methods but only used a small fraction of computation time required by the other methods (see **Table 6**). Furthermore, the *Dens*-based method provides investigators an intuitive and flexible way to select the tuning parameter according to desired density level they would like to impose on the network estimates.

APPLICATION TO RS-FMRI DATA FROM THE PHILADELPHIA NEURODEVELOPMENTAL COHORT (PNC)

PNC Study and Description

The PNC is a collaborative project between the Brain Behavior Laboratory at the University of Pennsylvania and the Children's Hospital of Philadelphia (CHOP), funded by NIMH through the American Recovery and Reinvestment Act of 2009 (Satterthwaite et al., 2014, 2015). The PNC study includes a population-based sample of over 9500 individuals aged 8–21 years selected among those who received medical care at the Children's Hospital of Philadelphia network in the greater Philadelphia area; the sample

is stratified by sex, age and ethnicity. A subset of participants from the PNC were recruited for a multimodality neuroimaging study which included resting-state fMRI (rs-fMRI). In this paper, we considered rs-fMRI data from 881 participants in the PNC study that were released in the dbGaP database. Compared to many other large-scale publicly available rs-fMRI datasets, the PNC data has a major advantage that all the images were acquired on a single MRI scanner using the same scanning protocol. Hence, the images from the PNC data do not suffer from extra variation caused by different scanners or protocols.

All images from the PNC study were acquired on a Siemens Tim Trio 3 Tesla, Erlangen, Germany using the same imaging sequences. The rs-fMRI scans were acquired with 124 volumes, TR 3000ms, TE 32 ms, flip angle 90°, FOV 192 × 192 mm, matrix 64 × 64 and effective voxel resolution 3.0 × 3.0 × 3.0 mm. More details about experimental settings and image acquisition can be found in Satterthwaite et al. (2015).

Prior to analysis, we performed a quality control procedure on the rs-fMRI. Specifically, we removed subjects who had more than 20 volumes with relative displacement >0.25 mm to avoid images with excessive motion (Satterthwaite et al., 2015). Among the 881 subjects who had rs-fMRI scans, 515 participants' data met the inclusion criterion and were used in our following

TABLE 4 | Comparison of Specificity for identifying connections based on different regularization selection methods across various sparsity levels with the simulated data.

Methods	Sparsity level								
	0.29	0.36	0.47	0.49	0.56	0.76	0.87	0.89	0.93
K-CV log like	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
K-CV Trace L2	0.22	0.35	0.30	0.39	0.29	0.37	0.40	0.28	0.41
AIC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BIC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$\lambda^*_{0.45}$	0.75	0.61	0.64	0.77	0.72	0.73	0.75	0.81	0.82
$\lambda^*_{0.75}$	0.20	0.13	0.16	0.19	0.25	0.21	0.23	0.22	0.24
λ^*_{platu}	0.04	0.03	0.04	0.04	0.05	0.04	0.05	0.04	0.05

TABLE 5 | Averaged performance of regularization methods across various sparsity levels with the simulated data.

Methods	MSE	Sensitivity	Specificity	Average of Sens. and Spec.	Computation time (Secs)
K-CV log like	0.017	1.000	0.000	0.500	2.7268
K-CV Trace L2	0.012	0.853	0.333	0.593	2.7079
AIC	0.017	1.000	0.000	0.500	0.0027
BIC	0.017	1.000	0.000	0.500	0.0016
$\lambda^*_{0.45}$	0.012	0.680	0.731	0.706	0.0004
$\lambda^*_{0.75}$	0.011	0.910	0.204	0.557	0.0004
λ^*_{platu}	0.015	0.978	0.041	0.509	0.0004

The values in bold are the optimal result across the different methods.

TABLE 6 | Comparison of computational time to select the tuning parameter for one randomly selected subject from the PNC study using different regularization selection methods.

Methods	K-CV log like	K-CV TraceL2	AIC	BIC	$\lambda^*_{0.45}$	$\lambda^*_{0.75}$	λ^*_{platu}
Computational time (Secs)	8575.93	8257.72	0.242	0.229	0.004	0.004	0.004

The values in bold are the optimal result across the different methods.

analysis. Among these 515 subjects, 290 (56%) were female and the mean age was 14.51 years (SD = 3.32).

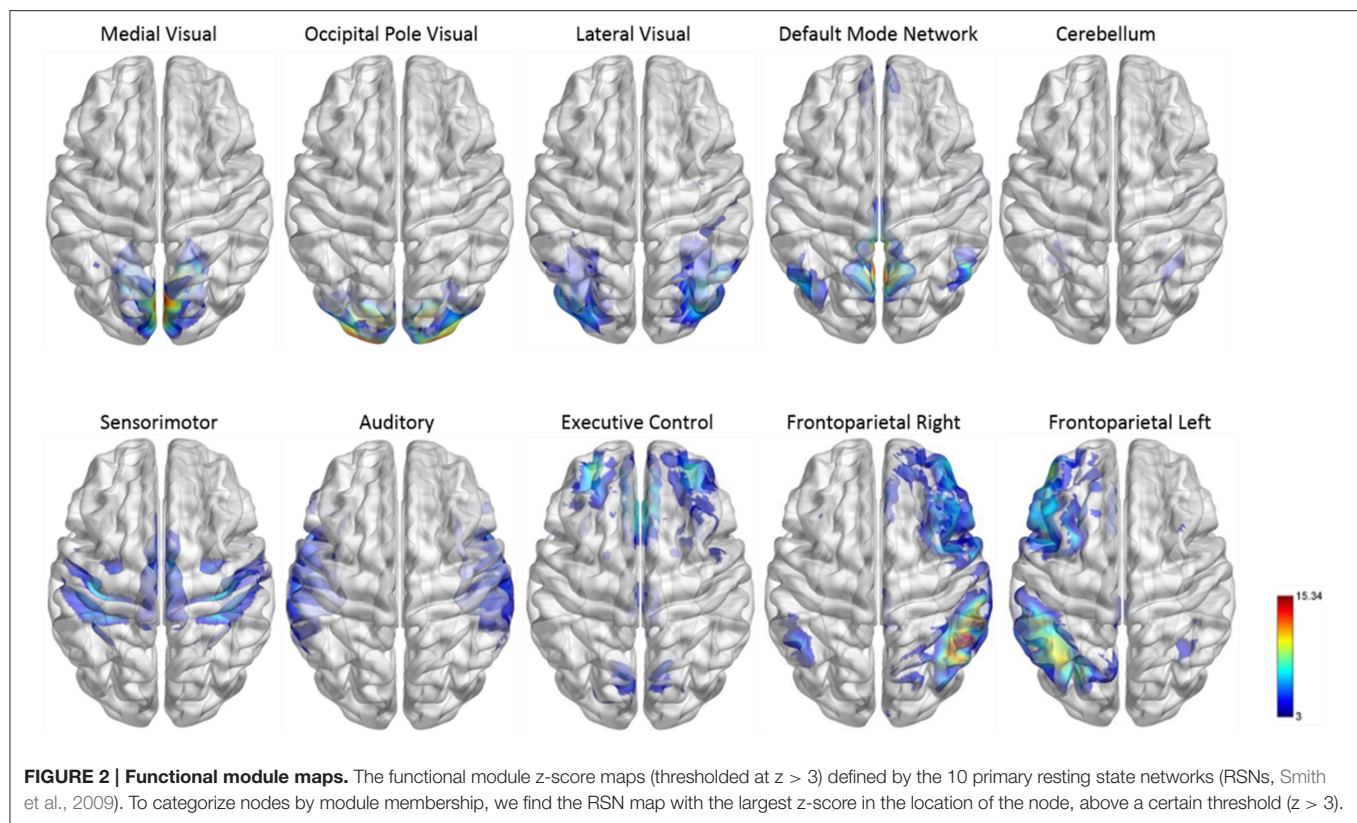
Rs-fMRI Data Preprocessing

The rs-fMRI data were preprocessed using the preprocessing script released from the 1000 Functional Connectomes Project. Specifically, skull stripping was performed on the T1 images to remove extra-cranial material, then the first four volumes of the functional time series were removed to stabilize the signal, leaving 120 volumes for subsequent preprocessing. The anatomical image was registered to the 8th volume of the functional image and subsequently spatially normalized to the MNI standard brain space. These normalization parameters from MNI space were used for the functional images, which were smoothed with a 6 mm FWHM Gaussian kernel. Motion corrections were applied on the functional images. A validated confound regression procedure (Satterthwaite et al., 2015) was performed on each subject’s time series data to remove confounding factors including motions, global effects, white matter (WM) and cerebrospinal fluid (CSF) nuisance signals.

The confound regression contained nine standard confounding signals (6 motion parameters plus global/WM/CSF) as well as the temporal derivative, quadratic term and temporal derivative of the quadratic of each. Furthermore, motion-related spike regressors were included to bound the observed displacement. Lastly, the functional time series data were band-pass filtered to retain frequencies between 0.01 and 0.1 Hz which is the relevant frequency range for rs-fMRI.

Brain Network Construction

In fMRI, brain activity is measured at voxel level, which are regions a few cubic millimeters in size. A typical 3D fMRI scan contains hundreds of thousands of voxels across the brain. The first step in brain network construction is usually to select a set of network nodes across the brain. Using individual voxels as network nodes has several issues: it results in an extremely high-dimensional connectivity matrix that is computationally challenging to estimate, and the voxel-based network tends to be very noisy due to the high noise level of fMRI BOLD signals in individual voxels. Additionally, a voxel-based network is highly

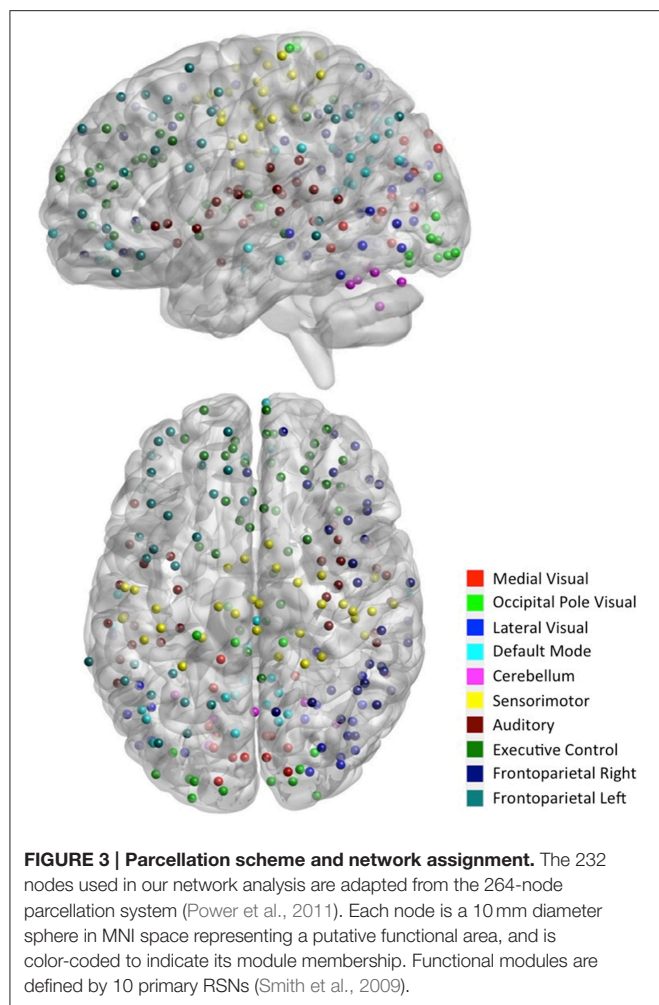


variable across subjects due to the difficulty of matching different subjects' brains at the voxel level. On the other hand, defining nodes by a coarse parcellation of the brain into large functionally homogenous regions can cause a loss in spatial resolutions when investigating the connectivity between brain locations. In our paper, we adopted the 264-node cortical parcellation system defined by Power et al. (2011). This system of nodes was determined using a combination of meta-analysis of task-based fMRI studies and resting state functional connectivity mapping techniques. In this network, each node is a 10 mm diameter sphere in standard MNI space representing a putative functional area, and the collection of nodes provides good coverage of the whole brain (see **Figure 2**). This node system provides a good balance of spatial resolution and dimension reduction. It is a finer spatial resolution than the commonly used Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002), but is not as granular as using a system of single voxels. This kind of intermediate node scheme is recommended to balance the trade-off between increased spatial resolution and attenuate signal-to-noise ratio (Fornito et al., 2010; Power et al., 2011).

To facilitate the understanding of the functional roles of the nodes, we assigned them to 10 functional networks or "modules" that correspond to the major resting state networks (RSNs) described by Smith et al. (2009) (see **Figure 2**). The RSN maps, determined by ICA decomposition of a large database of activation studies (BrainMap) and rs-fMRI data, are coherent during both task activity and at rest. The functional modules include medial visual network ("Med Vis," 15 nodes), occipital

pole visual network ("OP Vis," 15 nodes), lateral visual network ("Lat Vis," 19 nodes), default mode network ("DMN," 20 nodes), cerebellum ("CB," 6 nodes), sensorimotor network ("SM," 31 nodes), auditory network ("Aud," 29 nodes), executive control network ("EC," 39 nodes), and right and left frontoparietal networks ("FPR" and "FPL," 32 and 26 nodes, respectively). To determine the module membership at each node, we found the RSN map with the largest z-value in the location of the node, above a certain threshold ($z > 3$). Thirty two of the 264 nodes were not strongly associated with any RSN maps, and were therefore not included. A visualization of the remaining 232 nodes, classified by functional module, is shown in **Figure 3**. All brain visualizations were created using BrainNet Viewer (Xia et al., 2013).

To construct the network, we extracted the time series from each node with the following steps. First, the time series at each voxel were detrended, demeaned, and whitened. We then averaged the time series for all the voxels in each node to represent the node-specific time series. These node-specific time series were then used in subsequent analyses to estimate connectivity in the network. We note that using the within-node average or SVD time series in network construction is only appropriate when such summarized time series sufficiently represent the temporal dynamics within each node. When one uses a coarse brain parcellation such as the AAL regions in network construction, such this dimension reduction can cause problems in accurate estimation of the conditional independence structure in a network (Han et al., 2014). In the next section,



we describe the estimation of the 232×232 connectivity matrix using partial correlation to measure direct brain connectivity. For comparison, we also estimated a connectivity matrix based on full correlation for each subject to examine marginal connectivity between the nodes.

Estimation of the Partial Correlation Matrix

We applied the proposed method to estimate the partial correlation matrix based on the rs-fMRI data from the PNC study. For a given subject, we first obtained the sample covariance matrix based on the time series from each node. We then estimated the precision matrix from the sample covariance matrix using the CLIME method.

Comparison between *Dens*-Based Method and Existing Methods for Selecting the Tuning Parameter

To choose the tuning parameter for CLIME, we applied the proposed method based on the *Dens* criterion and considered λ_{platu}^* , $\lambda_{0.45}^*$ and $\lambda_{0.75}^*$. In comparison, we also considered other existing methods including AIC, BIC, and a 5-fold K-CV approach with the negative log likelihood and Trace L2 loss functions. In **Figure 4**, we plotted the profiles of the objective functions adopted by these methods for choosing the tuning

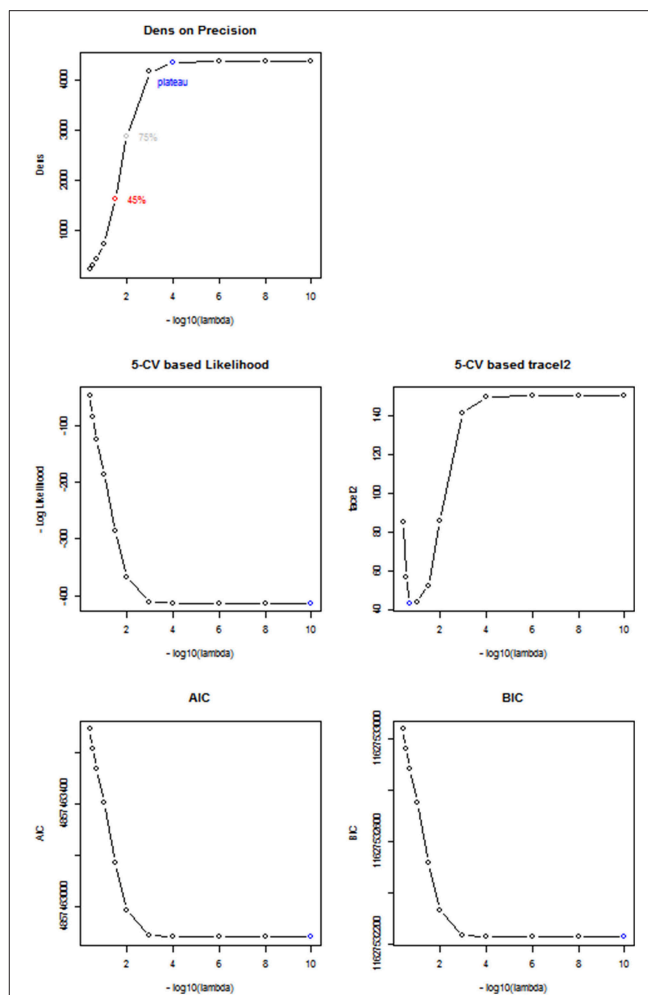


FIGURE 4 | Comparison between different Regularization methods.

Based on a randomly selected subject, we compared the performance between 5 different regularization methods including *Dens* method, 5-CV based negative log likelihood, 5-CV based TraceL2, AIC, and BIC, where λ values are on the $-\log_{10}$ scale, ranging from 10^{-10} to 0.4. The selected λ under each method is in blue.

parameter across a series of λ values, ranging from $1e-10$ to 0.4, for a randomly selected subject in the PNC study. In **Table 5**, we present the selected tuning parameter and the associated computation time based on each of these methods. From **Figure 4**, we can see the profile of the objective function based on *Dens* shows a similar pattern with the profiles of the AIC, BIC and the negative log likelihood 5-CV. All four of these profiles show that the objective functions improve significantly when λ was decreased from 0.4 to $1e-3$, reach a plateau around $1e-4$, and then only had very small changes when λ was further decreased. However, since these three existing methods (negative log likelihood 5-CV, AIC, BIC) all choose the λ that maximizes their corresponding objective function, they ended up choosing the minimum λ , i.e., $1e-10$, in the series. In contrast, the Trace L2-based 5-CV method had a different pattern which selected a value of $\lambda = 0.2$. This corresponds to a fairly strong sparsity constraint in CLIME and leads to the

sparsest estimate of the partial correlation matrix among all these methods.

Our proposed *Dens*-based selection method was the most efficient of all the methods considered. In particular, it showed a dramatic reduction in computation time as compared with the cross-validation methods. Unlike the AIC, BIC and negative log likelihood-based 5-CV methods which always selected the minimum $\lambda = 1e-10$, our proposed *Dens*-method was much more flexible in terms of selecting tuning parameters that correspond to various density levels that users may be interested in. Specifically, we found that the network corresponding to $\lambda_{\text{platu}}^* = 1e-4$ was extremely close to the estimated network based on $\lambda = 1e-10$ chosen by the AIC, BIC and negative log likelihood-based methods. We found that $\lambda_{0.45}^* = 0.032$ and it induced less stringent sparsity control as compared to $\lambda = 0.2$ as selected by the Trace L2 method. For $\lambda = 0.2$, we can see from **Figure 4** that it only reaches 10% of the *Dens* level in the unconstrained estimates of the network.

We also investigated the consistency of the results based on the *Dens*-based selection method across subjects. We randomly selected 100 subjects from the PNC study and applied the proposed method for choosing the CLIME tuning parameter for estimating subject-specific precision matrices. **Figure 5** displays the profiles of the *Dens* objective function across subjects. The results show that the proposed *Dens* objective function demonstrates a consistent pattern across subjects, and we also found consistent values across all 100 subjects for λ_{platu}^* , $\lambda_{0.45}^*$, and $\lambda_{0.75}^*$. Based on this finding, it is well-justified for us to apply the same tuning parameter to estimate partial correlation matrices for all subjects in the PNC study. This greatly facilitates between-subject comparisons and also allows the construction of a group-level partial correlation matrix by combining subject-specific estimates.

Comparison between the Proposed Method and Existing Methods for Estimating Partial Correlation

Using the fMRI data from the PNC study, we compared the performance of the proposed *Dens*-based method with two existing methods for estimating partial correlation. We first compared to the method referred to as the L1 precision method (Schmidt, 2006) which was used to obtain partial correlation in the well-known network modeling paper by Smith et al. (2011). The L1 precision method requires selection of a regularization-controlling parameter $\lambda^\#$. We considered the values within the range used in Smith et al. (2011) which includes $\lambda^\# = 1$ and 5. For these regularization values, the L1 precision method produced a diagonal matrix, which is an overly sparse estimate for the precision matrix. To fix this issue, we decreased $\lambda^\#$ to 0.1 and 0.5 to obtain a less sparse precision matrix. However, in these cases, the L1 precision algorithm failed to provide valid estimates and produced precision matrices with complex values. Furthermore, the L1-precision method is much more time-consuming than the proposed approach, using 1573 s for estimating a single subject's precision matrix at $\lambda^\# = 5$. When we specified $\lambda^\# = 0.1$ in order to obtain a less sparse precision matrix, the computation time dramatically increased to 12,849 s per subject. In comparison,

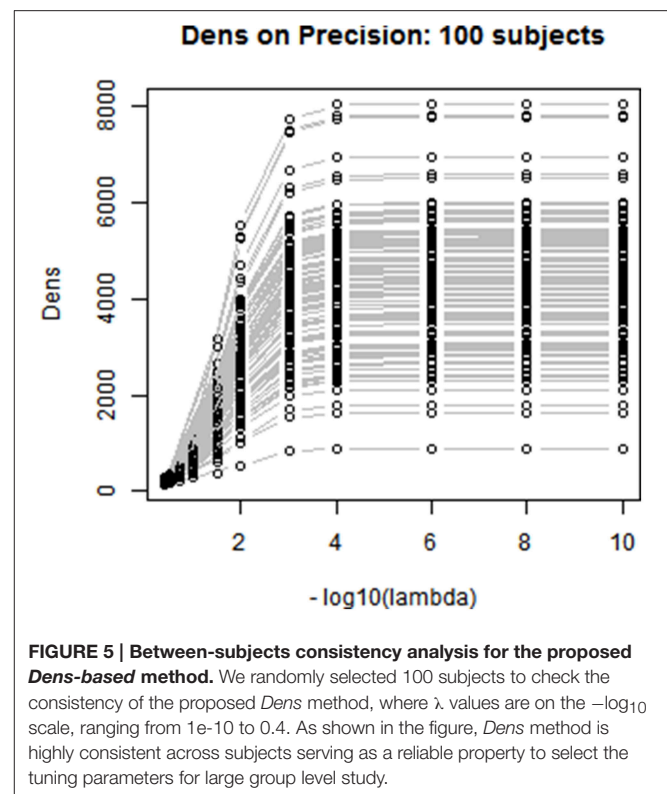
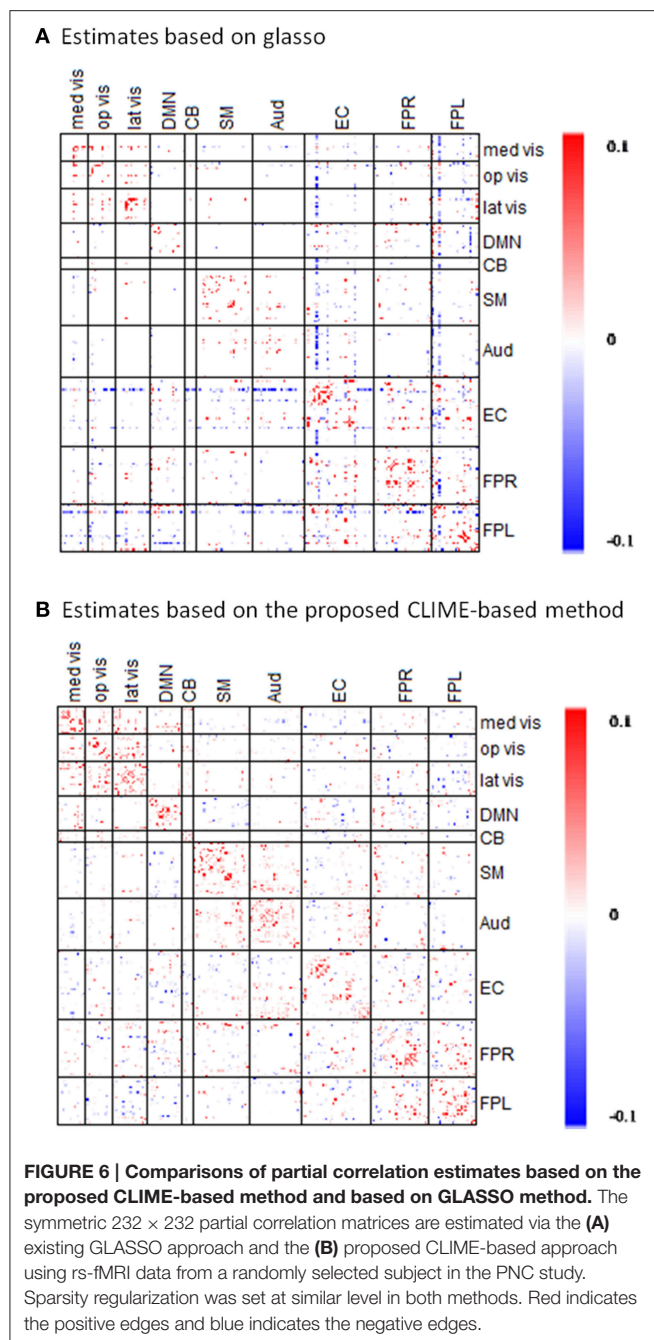


FIGURE 5 | Between-subjects consistency analysis for the proposed *Dens*-based method. We randomly selected 100 subjects to check the consistency of the proposed *Dens* method, where λ values are on the $-\log_{10}$ scale, ranging from $1e-10$ to 0.4 . As shown in the figure, *Dens* method is highly consistent across subjects serving as a reliable property to select the tuning parameters for large group level study.

our proposed method produced valid estimates of the partial correlation matrix for all 515 subjects in PNC data. Our method was also significantly faster than the L1 precision method, only taking about 58–60 s per subject. In addition, we also considered another existing method for estimating partial correlation based on the glasso R package (Schmittmann et al., 2015). When comparing the results (**Figure 6**), one major distinction is that connectivity matrix based on our proposed method showed more positive connections within-modules nodes suggesting within-module are more densely connected to one another than to the rest of the network. In comparison, glasso-based connectivity matrix showed less within-module positive connections and in some cases even produced strong negative connections within the same functional module. Based on the network comparison criterion in the literature (Power et al., 2011), these results suggest that the connectivity matrix based on our proposed method more accurately reflects the brain organization in the sense that it better captures the strong positive functional connections within the established functional modules.

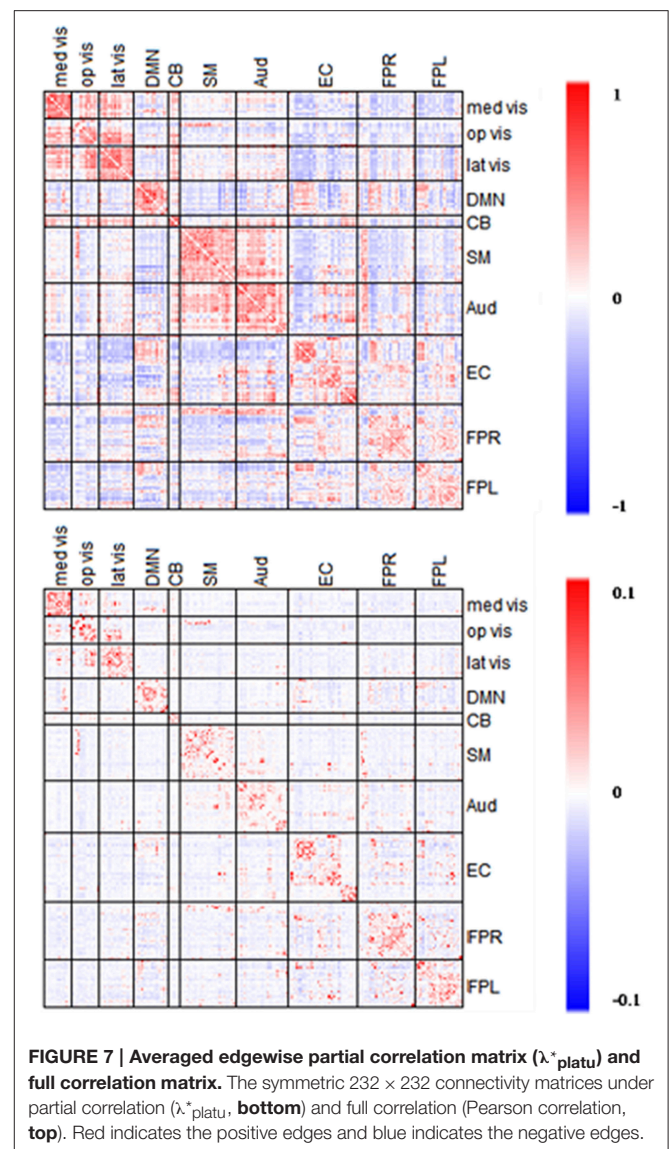
Comparison of Network Connectivity Based on Partial Correlation and Full Correlation

In the section, we compare the partial-correlation-based network connectivity and full correlation-based connectivity for the PNC study. Following the method from Satterthwaite et al. (2014), we did not threshold the correlation matrix, yielding a fully connected correlation matrix. Thus, to ensure comparability, we imposed minimum sparsity control in the partial correlation

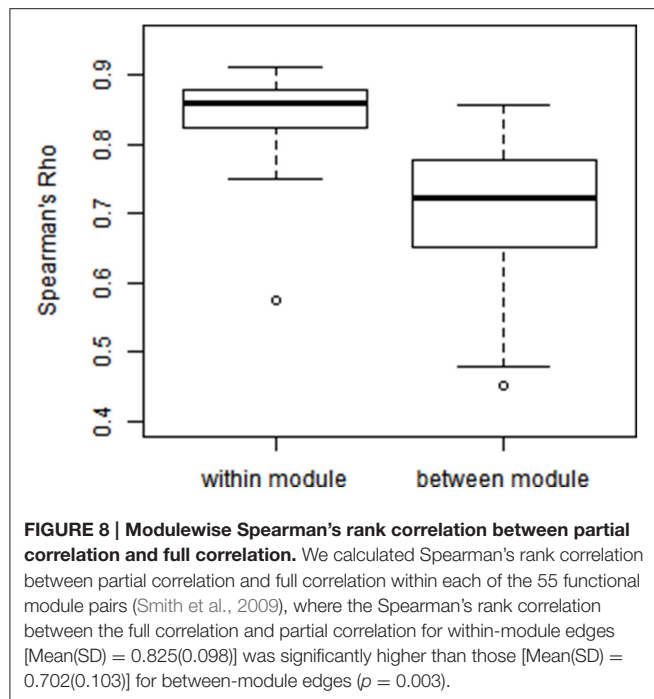


estimation and selected λ_{platu}^* for the CLIME. **Figure 7** displays the partial correlation matrix and correlation matrix averaged across the 515 subjects in PNC data.

Full correlation values ranged between -0.45 and 0.83 , and in comparison, partial correlation values ranged between -0.03 and 0.18 . As expected, the magnitude of partial correlation was much smaller than full correlation since the partial correlation reflects the direct connections between nodes after removing the confounding effects from all the other nodes. Based on the 10 functional module system defined by Smith et al. (2009), we divided the upper-triangle of the 232×232 edgewise connectivity



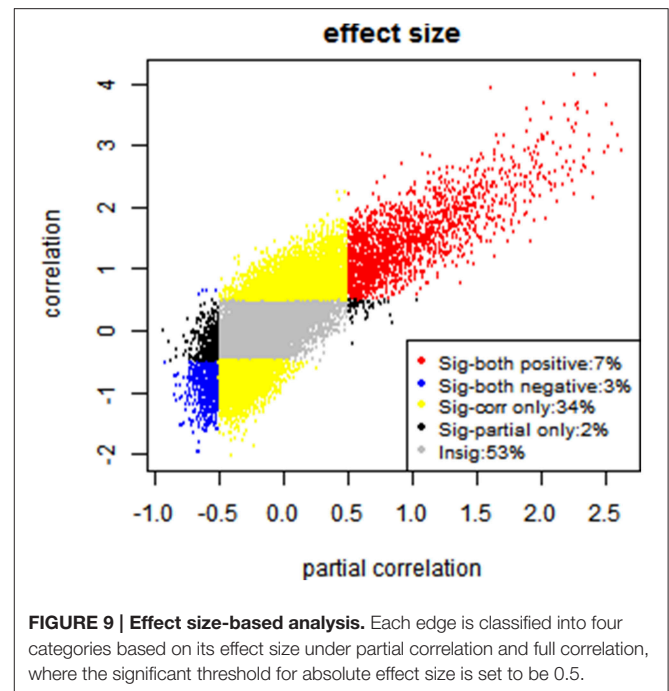
matrices into 55 module-wise blocks including the 10 within-module blocks and 45 between-module blocks. In the full correlation-based connectivity matrix, we can see the majority of positive marginal connections were found in within-module blocks, that is the diagonal blocks in the connectivity matrix. We also found positive connections in several between-module blocks, in particular between the three visual networks (Med Vis, Op Vis, Lat Vis) and also between the Auditory (Aud) and Sensorimotor (SM) network. In the partial correlation-based connectivity matrix, the strong positive connection in within-module blocks became even more prominent, indicating that the most significant positive direct connections in the brain are observed within functional modules, and for between-module node pairs we observed fewer positive connections as compared to the full correlation matrix. For example, we observed fewer positive connections between the Auditory (Aud) and Sensorimotor (SM) network. Similarly, the connections between the three visual networks had dropped considerably



too as compared to the full correlation matrix. These findings suggest that a lot of the marginal connections for between-module node pairs are mainly due to some confounding factors and not necessarily due to the direct connections between modules. Another important finding is that in the full correlation-based connectivity matrix, there were considerable negative functional connections in the between-module blocks. Several of these negative marginal connections disappeared in the partial correlation matrix, indicating that many of the negative connections may be caused by confounding factors. This finding agrees with some recent findings in the neuroimaging community that showed many negative functional connections in rs-fMRI may be due to non-neurological reasons such as global signal removal performed during imaging pre-processing (Giove et al., 2009; Murphy et al., 2009; Weissenbacher et al., 2009; Chen et al., 2011) or inhomogeneous cerebral circulation across the brain (Goelman et al., 2014).

We examined the consistency between partial correlation and full correlation findings across all edges in the network. Since the measures have different scales, we utilized Spearman's rank correlation coefficient (Spearman's Rho) to measure their association at all edges. As shown in **Figure 8**, the Spearman's Rho between the full correlation and partial correlation for within-module edges (Mean \pm SD = 0.825 \pm 0.098) was significantly higher than those for between-module edges (Mean \pm SD = 0.702 \pm 0.103; $p = 0.003$). This demonstrates that partial correlation and full correlation were more consistent for within-module edges compared to between-module edges.

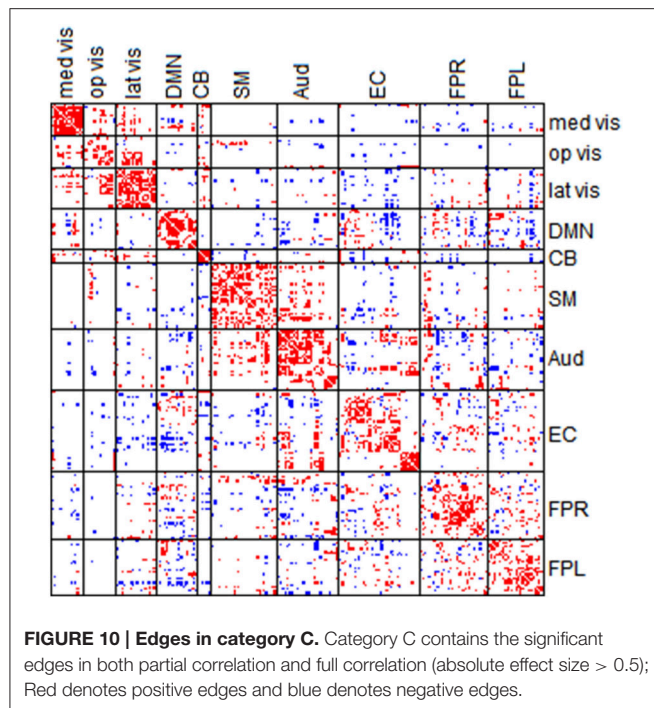
Furthermore, since researchers are mostly interested in significant connections, we examine the consistency between the partial correlation and full correlation for these significant edges. Given the large sample size of the PNC data, we have



high statistical power to detect even very small deviations from zero in the correlations. Therefore, even edges with very small effect size demonstrated highly significant p -values in hypothesis testing. Therefore, we used the effect size instead of p -values for thresholding purpose. Specifically, we first performed the Fisher's Z transformation on both the partial correlation and full correlation values. We then calculated the effect size for the connectivity at each edge by dividing the mean of z-transformed full correlations or partial correlations to its standard deviation (Kemmer et al., 2015). The effect sizes ranged from -2 to 4 for full correlation and -1 to 2.5 for partial correlation (see **Figure 9**). We then defined significant edges as those with an effect size of greater than 0.5.

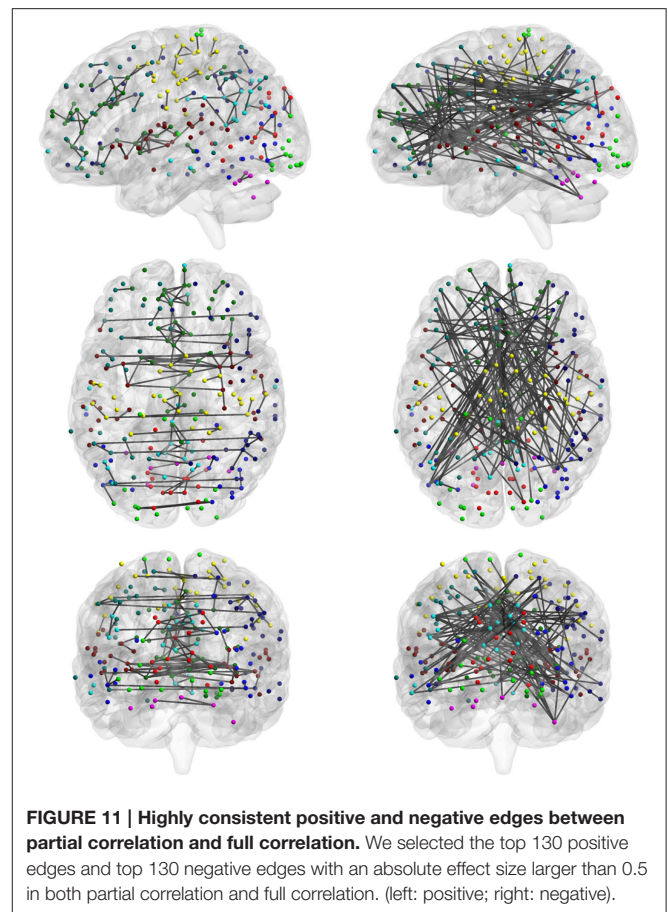
After the thresholding to define the significant edges, each edge is classified into one of the following four categories: (A) significant in partial correlation but insignificant in full correlation (2%); (B) significant in full correlation but insignificant in partial correlation (34%); (C) significant in both (10%); (D) insignificant in both (53%), shown in **Figure 9**. Moreover, we evaluated the sign consistency between the full correlation and partial correlation on the edge level. The percentage of edges with sign consistency within each of those four categories are A: 83.54%, B: 86.52%, C: 100% and D: 66.05%.

Among the four categories, category C reveals the significantly consistent edges based on full correlation and partial correlation. **Figure 10** displays edges mapped to the module-wise blocks. Results show that consistently significant positive edges were more concentrated at the within-module regions and consistently significant negative edges were more concentrated at the between-module regions. In particular, we found that considerable consistently negative connections were observed between the default mode network and other modules,



especially with the executive control module. To provide better visualization of these consistently significant edges, we selected the top 130 positive edges and top 130 negative edges from category C and mapped them onto the brain (**Figure 11**). An important observation from **Figure 11** is that the strongest positive connections based on both partial correlation and full correlation were the connections between homologous brain locations in the left and right hemisphere. This finding is consistent with some previous findings based on PET resting-state data collected on rats which also showed that the largest partial correlation coefficients in rat brain were between homologous brain regions (Horwitz et al., 1984). Another important observation from **Figure 11** is that the strongest negative connections based on both partial correlation and full correlation tend to have longer spatial distance than strongest positive connections, which is consistent with the previous findings showing that the percentage of negative functional connectivity and spatial distance are significantly correlated (Chen et al., 2011).

We further examined edges in category B, which represented edges that were significant based on full correlation but insignificant based on partial correlation. We examined the proportion of category B edges in each of the module-wise blocks and found that these inconsistent edges were more likely to be observed in between-module connections than for within-module connections. In particular, we found that the following three between-module pairs showed the highest inconsistency between the marginal and direct connectivity: Med vis and FPL for which 56% of all edges between these two networks were in category B, that is only significant based on full correlation; Lat vis and EC for which 50 of all edges between them were in Category B.



Comparison between Network Connectivity Using Partial Correlation Matrix Based on Different $Dens$ level

In this section, we explore the difference in the estimated direct connectivity based on the proposed partial correlation method using different levels of sparsity control. Specifically, we compared partial correlation estimates obtained with λ_{platu}^* where minimum sparsity control was applied vs. partial correlation estimates obtained with $\lambda_{0.45}^*$ where some sparsity regularization were applied such that the partial correlation matrix reached about 45% of the maximum density level.

The estimated partial correlation matrices based on λ_{platu}^* and $\lambda_{0.45}^*$ are presented in **Figure 12**. As expected, the partial correlation matrix based on $\lambda_{0.45}^*$ was sparser than that based on λ_{platu}^* . Furthermore, in the between-module regions the majority of negative (blue) connections under λ_{platu}^* disappeared using $\lambda_{0.45}^*$, while in the within-module regions the positive (red) connections under λ_{platu}^* were retained using $\lambda_{0.45}^*$. Marginally, the partial correlations ranged between -0.03 to 0.18 based on λ_{platu}^* and -0.02 to 0.22 based on $\lambda_{0.45}^*$. Therefore, the limit of the estimated correlations shrank slightly in the negative edges but increased in the positive edges.

To further explore this shrinkage effect, we examined the edges with an absolute effect size larger than 0.3. As shown in

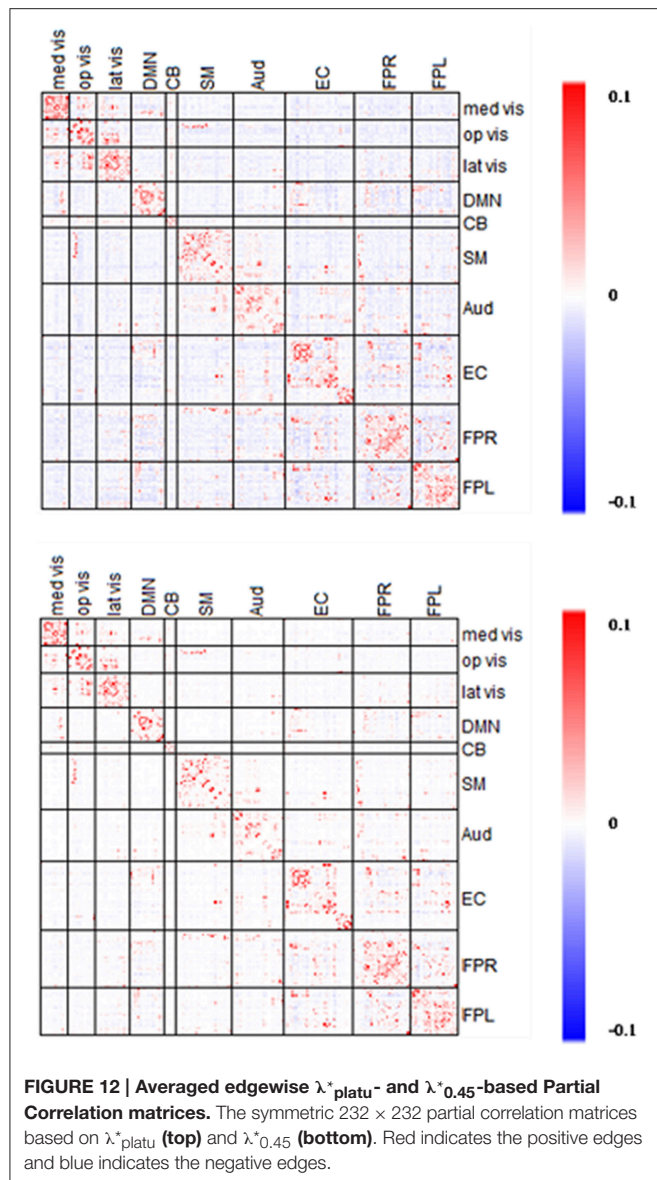
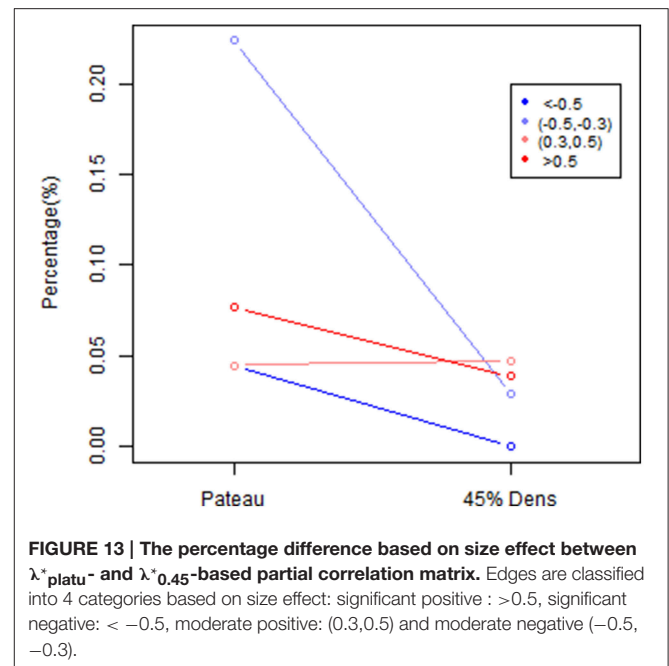


Figure 13, the majority of the negative edges with medium (0.3–0.5) to large (>0.5) effect sizes disappeared under a stronger sparsity control, whereas the percentage of the negative edges with medium effect sizes decreased from 22.4 to 2.9%, and the percentage of the negative edges with large effect sizes was decreased to 0.04%. However, the positive edges with medium to large effect sizes were mostly retained under a stronger sparsity control. These results suggest that for positive edges, the edges with medium and large effect size remained fairly robust under shrinkage. However, the negative edges were more likely to disappear under the stronger sparsity control, so the shrinkage effects were much stronger for negative edges than for positive edges. This result suggests that the when applying more sparsity regularization in our proposed procedure, we will still maintain the ability to detect the significant positive edges while the negative edges would experience more shrinkage in the estimates.



Again this may be mainly due to the fact that a lot of the negative connections observed in rs-fMRI data were not due to direct connection or neurophysiological effect but rather due to artifacts from imaging processing or biological reasons (Chen et al., 2011; Goelman et al., 2014).

DISCUSSION

In this paper, we propose a more efficient and reliable statistical method for estimating partial correlation in brain network modeling, which provides a useful tool to investigate direct brain functional connectivity. Compared to existing methods used in the neuroimaging community, the proposed method is shown to be more reliable and computationally efficient. Another major advantage of this technique is that it is scalable to large-scale brain networks with a large number of nodes, for which the existing methods often fail to generate reliable network estimates. Thus, the proposed method can provide a powerful tool for investigating whole brain connectivity in both task-related as well as resting state fMRI studies.

When estimating the partial correlation matrix under the regularization framework, a major challenge is how to select an appropriate tuning parameter to control the sparsity level. Existing selection approaches are often made based on subjective choices or by considering only a few candidates. We propose a new *Dens*-based selection method which considers a series of values across the range of the tuning parameter, and we evaluate the proposed *Dens* criterion for the estimated precision matrix at each value. Hence, we can have a more comprehensive picture of the whole profile of the criterion function across the range of the tuning parameter. Based on the *Dens* profile, users can now have better understanding on the implications on the sparse level of the estimated networks based on different

tuning parameters. Thus, they can make more informed choices of the tuning parameter based on the desired *Dens* level they would like to achieve in the estimated partial correlation matrix. Our proposed *Dens*-based selection method is also much faster than the existing selection methods. This will allow users to perform the selection process across many or even all subjects to evaluate the consistency in the selection of the tuning parameter across subjects and to select a common tuning parameter that has good performance across different subjects. In comparison, some of the existing selection methods, such as the cross-validation based method, are very time consuming and hence it is very difficult to conduct such consistency checks across a large number of subjects. Our results from the PNC data showed that the proposed selection procedure leads to a fairly consistent choice of the tuning parameter across different subjects. Therefore, we can apply the same regularization across all subjects, which facilitates performing group analysis of the partial correlations.

When comparing the partial correlation-based and full correlation-based connectivity matrices, we note that the partial correlation removed considerable marginal correlations found in the full correlation matrix that may be due to non-neurophysiological confounding factors. For example, in the partial correlation matrix, many of the significant marginal connections in between-module pairs were not present suggesting these connections between different brain modules were likely caused by global effects or common connection to a third party (Smith et al., 2011; Smith, 2012). Furthermore, the full-correlation-based connectivity matrix demonstrated considerable amount of negative functional connectivity in between-module pairs. Neuroimaging literature has shown that many negative connection findings in rs-fMRI may be caused by non-neurophysiological reasons such as artifacts from global signal removal or inhomogeneous cerebral circulation across the brain (Chen et al., 2011; Goelman et al., 2014). There are considerable controversies in terms of origin and

interpretations for these negative connections (Giove et al., 2009; Murphy et al., 2009; Weissenbacher et al., 2009). Hence, many network analyses simply ignore all negative connection (Buckner et al., 2009; Meunier et al., 2009; Satterthwaite et al., 2015). When applying the partial correlation to investigate the direct functional connectivity, we observed that many of the negative connections disappear and those that remain tend to be well-established negative connections such as those between default mode network and other networks. Moreover, based on our *Dens*-based method, we demonstrated that the moderate negative connections were less robust than the positive connections and the strong negative connections, further indicating that a lot of the moderate negative functional connectivity may be caused by non-neurophysiological reasons. By using the proposed partial correlation method with appropriate sparsity control, we can potentially perform meaningful network analysis for negative connections as well in brain network modeling.

An R package “DensParcorr” for implementing the proposed statistical methods can be downloaded from CRAN and the website of Center for Biomedical Imaging Statistics (CBIS) of Emory University.

AUTHOR CONTRIBUTIONS

YW and YG developed the methodology and performed the analysis of PNC data; YW, YG, and KJ developed and conducted the simulation study; YW and PK preprocessed the data; YW, YG, PK, and JK wrote the manuscript.

ACKNOWLEDGMENTS

This work was supported by NIMH R01 grants (2R01MH079448-04A1 and 1R01MH105561-01). PNC study was supported by NIH RC2 grants (MH089983, MH089924). The center for Applied Genomics at The Children’s Hospital in Philadelphia recruited all subjects.

REFERENCES

- Akaike, H. (1998). “Information theory and an extension of the maximum likelihood principle,” in *Selected Papers of Hirotugu Akaike*, eds E. Parzen, K. Tanabe, and G. Kitagawa (New York, NY: Springer), 199–213.
- Ben-Israel, A., and Greville, T. N. (2003). *Generalized Inverses: Theory and Applications*. New York, NY: Springer Science & Business Media.
- Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409
- Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., et al. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer’s disease. *J. Neurosci.* 29, 1860–1873. doi: 10.1523/JNEUROSCI.5062-08.2009
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Cai, T., Liu, W., and Luo, X. (2011). A constrained l_1 minimization approach to sparse precision matrix estimation. *J. Am. Stat. Assoc.* 106, 594–607. doi: 10.1198/jasa.2011.tm10155
- Cai, T. T., Liu, W., and Luo, X. (2012). *clime: Constrained L_1 -Minimization for Inverse (covariance) Matrix Estimation*. R package version 0.4.1. Available online at: <http://CRAN.R-project.org/package=clime>
- Chen, G., Chen, G., Xie, C., and Li, S.-J. (2011). Negative functional connectivity and its dependence on the shortest path length of positive network in the resting-state human brain. *Brain Connect.* 1, 195–206. doi: 10.1089/brain.2011.0025
- Church, J. A., Fair, D. A., Dosenbach, N. U., Cohen, A. L., Miezin, F. M., Petersen, S. E., et al. (2009). Control networks in paediatric Tourette syndrome show immature and anomalous patterns of functional connectivity. *Brain* 132, 225–238. doi: 10.1093/brain/awn223
- Deco, G., Jirsa, V. K., and McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* 12, 43–56. doi: 10.1038/nrn2961
- Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., et al. (2010). Prediction of individual brain maturity using fMRI. *Science* 329, 1358–1361. doi: 10.1126/science.1194144
- Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: SIAM.
- Fornito, A., Zalesky, A., and Bullmore, E. T. (2010). Network scaling effects in graph analytic studies of human resting-state fMRI data. *Front. Syst. Neurosci.* 4:22. doi: 10.3389/fnsys.2010.00022
- Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441. doi: 10.1093/biostatistics/kxm045

- Giove, F., Gili, T., Iacovella, V., Macaluso, E., and Maraviglia, B. (2009). Images-based suppression of unwanted global signals in resting-state functional connectivity studies. *Magn. Reson. Imaging* 27, 1058–1064. doi: 10.1016/j.mri.2009.06.004
- Goelman, G., Gordon, N., and Bonne, O. (2014). Maximizing negative correlations in resting-state functional connectivity MRI by time-lag. *PLoS ONE* 9:e111554. doi: 10.1371/journal.pone.0111554
- Han, F., Liu, H., and Caffo, B. (2013). Sparse median graphs estimation in a high dimensional semiparametric model. *arXiv preprint* arXiv:1310.3223.
- Han, F., Qiu, H., Liu, H., and Caffo, B. (2014). On the impact of dimension reduction on graphical structures. *arXiv preprint* arXiv:1404.7547.
- Horwitz, B., Duara, R., and Rapoport, S. I. (1984). Intercorrelations of glucose metabolic rates between brain regions: application to healthy males in a state of reduced sensory input. *J. Cereb. Blood Flow Metab.* 4, 484–499. doi: 10.1038/jcbfm.1984.73
- Kemmer, P. B., Guo, Y., Wang, Y., and Pagnoni, G. (2015). Network-based characterization of brain functional connectivity in Zen practitioners. *Front. Psychol.* 6:603. doi: 10.3389/fpsyg.2015.00603
- Liu, H., Han, F., Yuan, M., Lafferty, J., and Wasserman, L. (2012). High-dimensional semiparametric Gaussian copula graphical models. *Ann. Stat.* 40, 2293–2326. doi: 10.1214/12-AOS1037
- Liu, H., Roeder, K., and Wasserman, L. (2010). Stability approach to regularization selection (stars) for high dimensional graphical models. *Adv. Neural Inf. Process. Syst.* 24, 1432–1440.
- Liu, W., and Luo, X. (2012). High-dimensional sparse precision matrix estimation via sparse column inverse operator. *arXiv preprint* arXiv:1203.3896.
- Marrelec, G., Krainik, A., Duffau, H., Pelegrini-Issac, M., Lehericy, S., Doyon, J., et al. (2006). Partial correlation for functional brain interactivity investigation in functional MRI. *Neuroimage* 32, 228–237. doi: 10.1016/j.neuroimage.2005.12.057
- Meinshausen, N., and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Stat.* 34, 1436–1462. doi: 10.1214/009053606000000281
- Meunier, D., Achard, S., Morcom, A., and Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *Neuroimage* 44, 715–723. doi: 10.1016/j.neuroimage.2008.09.062
- Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., and Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage* 44, 893–905. doi: 10.1016/j.neuroimage.2008.09.036
- Ogawa, S., Lee, T.-M., Kay, A. R., and Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. Natl. Acad. Sci. U.S.A.* 87, 9868–9872. doi: 10.1073/pnas.87.24.9868
- Peng, J., Wang, P., Zhou, N., and Zhu, J. (2009). Partial correlation estimation by joint sparse regression models. *J. Am. Stat. Assoc.* 104, 735–746. doi: 10.1198/jasa.2009.0126
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011). Functional network organization of the human brain. *Neuron* 72, 665–678. doi: 10.1016/j.neuron.2011.09.006
- Satterthwaite, T. D., Elliott, M. A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M. E., et al. (2014). Neuroimaging of the Philadelphia neurodevelopmental cohort. *Neuroimage* 86, 544–553. doi: 10.1016/j.neuroimage.2013.07.064
- Satterthwaite, T. D., Wolf, D. H., Roalf, D. R., Ruparel, K., Erus, G., Vandekar, S., et al. (2015). Linked sex differences in cognition and functional connectivity in youth. *Cereb. Cortex* 25, 2383–2394. doi: 10.1093/cercor/bhu036
- Schäfer, J., and Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.* 4, 1175–1189. doi: 10.2202/1544-6115.1175
- Schmidt, M. (2006). *L1precision: Matlab Code for MAP Estimation of Gaussian Graphical Model Precision with L1-Regularizer*. Available online at: <http://www.cs.ubc.ca/~schmidtm/Software/L1precision.html>
- Schmittmann, V. D., Jahfari, S., Borsboom, D., Savi, A. O., and Waldorp, L. J. (2015). Making large-scale networks from fMRI data. *PLoS ONE* 10:e0129074. doi: 10.1371/journal.pone.0129074
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136
- Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L., and Greicius, M. D. (2009). Neurodegenerative diseases target large-scale human brain networks. *Neuron* 62, 42–52. doi: 10.1016/j.neuron.2009.03.024
- Smith, S. M. (2012). The future of FMRI connectivity. *Neuroimage* 62, 1257–1266. doi: 10.1016/j.neuroimage.2012.01.022
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045. doi: 10.1073/pnas.0905267106
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for FMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Wasserman, L., and Roeder, K. (2009). High dimensional variable selection. *Ann. Stat.* 37, 2178. doi: 10.1214/08-aos646
- Weissenbacher, A., Kasess, C., Gerstl, F., Lanzenberger, R., Moser, E., and Windischberger, C. (2009). Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. *Neuroimage* 47, 1408–1416. doi: 10.1016/j.neuroimage.2009.05.005
- Xia, M., Wang, J., and He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS ONE* 8:e68910. doi: 10.1371/journal.pone.0068910
- Zhang, J., Zhou, L., Wang, L., and Li, W. (2015). Functional brain network classification with compact representation of SICE matrices. *IEEE Trans. Biomed. Eng.* 62, 1623–1634. doi: 10.1109/TBME.2015.2399495

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Wang, Kang, Kemmer and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fMRI data

Yuting Xu and Martin A. Lindquist*

Department of Biostatistics, Johns Hopkins University, Baltimore, MD, USA

OPEN ACCESS

Edited by:

Jian Kang,
Emory University, USA

Reviewed by:

Shuo Chen,
University of Maryland, USA
Yize Zhao,
Statistical and Applied Mathematical
Sciences Institute (SAMS), USA (in
collaboration with Fei Zou),
Fei Zou,
The University of North Carolina at
Chapel Hill, USA

*Correspondence:

Martin A. Lindquist,
Department of Biostatistics, Johns
Hopkins University, 615 N. Wolfe
Street, E3634, Baltimore, MD 21205,
USA
mlindqui@jhsp.edu

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 14 May 2015

Accepted: 29 July 2015

Published: 04 September 2015

Citation:

Xu Y and Lindquist MA (2015)
Dynamic connectivity detection: an
algorithm for determining functional
connectivity change points in fMRI
data. *Front. Neurosci.* 9:285.
doi: 10.3389/fnins.2015.00285

Recently there has been an increased interest in using fMRI data to study the dynamic nature of brain connectivity. In this setting, the activity in a set of regions of interest (ROIs) is often modeled using a multivariate Gaussian distribution, with a mean vector and covariance matrix that are allowed to vary as the experiment progresses, representing changing brain states. In this work, we introduce the Dynamic Connectivity Detection (DCD) algorithm, which is a data-driven technique to detect temporal change points in functional connectivity, and estimate a graph between ROIs for data within each segment defined by the change points. DCD builds upon the framework of the recently developed Dynamic Connectivity Regression (DCR) algorithm, which has proven efficient at detecting changes in connectivity for problems consisting of a small to medium (<50) number of regions, but which runs into computational problems as the number of regions becomes large (>100). The newly proposed DCD method is faster, requires less user input, and is better able to handle high-dimensional data. It overcomes the shortcomings of DCR by adopting a simplified sparse matrix estimation approach and a different hypothesis testing procedure to determine change points. The application of DCD to simulated data, as well as fMRI data, illustrates the efficacy of the proposed method.

Keywords: functional connectivity, dynamic functional connectivity, resting state fMRI, change point detection, network dynamics

1. Introduction

Functional connectivity (FC) is the study of the temporal dependencies between distinct, possibly spatially remote, brain regions (Friston, 1994). Assessing FC using functional Magnetic Resonance Imaging (fMRI) data, has proven particularly useful for discovering patterns indicating how brain regions are related, and comparing these patterns across groups of subjects (Lindquist, 2008; Friston, 2011). In recent years, it has become one of the most active research areas in the neuroimaging community, and it is a central concept in the long term goal of understanding the human connectome (Sporns et al., 2005). The hope is that increased knowledge of networks and connections will help facilitate research into a number of common brain disorders.

FC is fundamentally a statistical concept, and is typically assessed using statistical measures such as correlation (Biswal et al., 1995), cross-coherence (Sun et al., 2004), and mutual information (Jeong et al., 2001). In the past few years it has become increasingly common to assume that the fMRI time series data follows a multivariate Gaussian distribution, and quantify FC using

the estimated covariance, correlation or precision (inverse covariance) matrix (Varoquaux et al., 2010; Cribben et al., 2012, 2013). In this setting there is a well-known relationship between the estimated precision matrix and the underlying network graph of interest, and the use of algorithms for estimating sparse precision matrices (and thus graphs) have become critical (Friedman et al., 2008).

Most functional connectivity analyses performed to date have generally assumed that the relationship within functional networks is stationary across time. However, in recent years there has been an increased interest in studying dynamic changes in FC over time. These analyses have shown that rather than being static, functional networks appear to fluctuate on a time scale ranging from seconds to minutes (Chang and Glover, 2010). Here changes in both the strength and directionality of functional connections have been observed to vary across experimental runs (Hutchison et al., 2013), and it is believed that these changes may provide insight into the fundamental properties of brain networks.

When the precise timing and duration of state-related changes in FC are known before hand, it is possible to apply methods such as the psychophysiological interactions (PPI) technique (Friston et al., 1997) or statistical parametric networks analysis (Ginestet and Simmons, 2011). However, in many research settings the nature of the psychological processes being studied is unknown, particularly in resting-state fMRI (rfMRI), and it is therefore important to develop methods that can describe the dynamic behavior in connectivity without requiring prior knowledge of the experimental design. In the past couple of years, a number of such approaches have been suggested in the neuroimaging literature, including the use of sliding window correlations (Chang and Glover, 2010; Handwerker et al., 2012; Hutchison et al., 2013; Allen et al., 2014), change point models (Cribben et al., 2012, 2013), and volatility models (Lindquist et al., 2014).

One example is dynamic connectivity regression (DCR), which is a data-driven technique for partitioning a time course into segments and estimating the different connectivity networks within each segment (Cribben et al., 2012). It applies a greedy search strategy to identify possible changes in FC using the Bayesian Information Criteria (BIC). While optimizing the BIC value within each subsequence, DCR utilizes the GLASSO algorithm to estimate a sparse inverse covariance matrix. This is followed by a secondary analysis of the candidate split points,

where a permutation test is performed to decide whether or not the reduction in BIC at that time point is significant enough to be deemed a true change point. The structure of the DCR algorithm is briefly demonstrated in **Figure 1**.

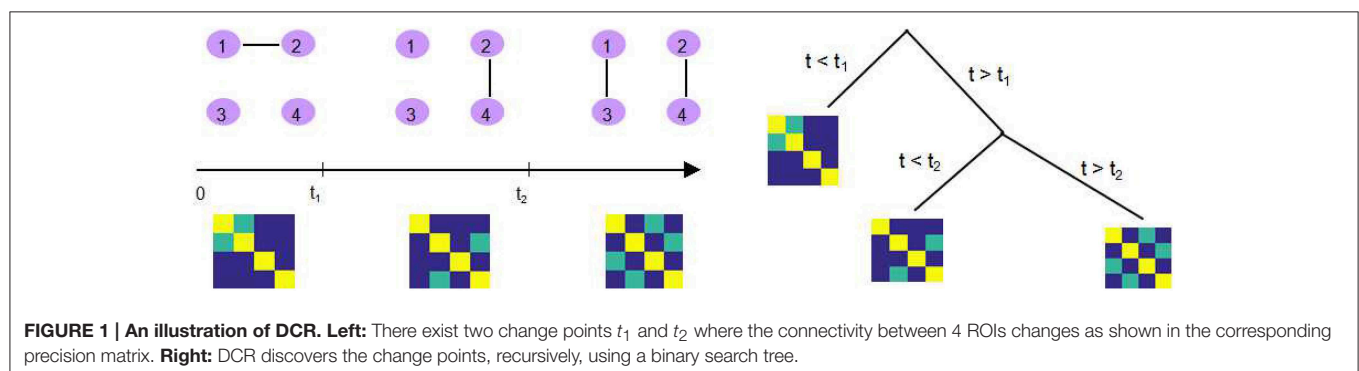
While the DCR algorithm has proven useful for detecting changes in FC, it has two major drawbacks. First, the computational cost of the algorithm increases rapidly with the number of ROIs. As the number of ROIs surpasses 50, the computation time can become prohibitive. Second, DCR requires a number of user-specified input parameters, some of which may be difficult to optimize without in-depth knowledge of the experiment and familiarity with the algorithm.

In this work, we introduce the Dynamic Connectivity Detection (DCD) algorithm for change point detection in fMRI time series data, as well as the estimation of a graph representing connectivity within each partition. It builds upon the basic DCR framework, using the same binary search tree structure to recursively identify potential change points. However, it replaces a number of critical components of DCR, including the manner in which the sparse matrix estimation is performed and significant change points determined. An adaptive thresholding approach is used to estimate a sparse covariance matrix, which provides a significant speed up in computation time compared to the GLASSO algorithm, and improves scalability. In addition, the permutation test used to detect significant change points is replaced by an alternative hypothesis test. Because of these changes, all the input parameters in the DCD algorithm have a clear interpretation in the context of hypothesis testing, allowing users to specify the desired control of Type I and Type II errors.

This paper is organized as follows. In Section 2 we begin by briefly reviewing the basic steps of DCR, followed by a discussion of sparse parameter estimation, and a description of the new DCD algorithm for single-subject change point detection and graph estimation. Thereafter we demonstrate the performance of DCD in Sections 3 and 4 by applying the method to a series of simulation studies and experimental data. The obtained results are contrasted with similar results obtained using DCR. The paper concludes with a discussion.

2. Methods

Consider fMRI data from a single subject consisting of multivariate time series, where each dimension corresponds to



activity from a single region of interest (ROI). Assume that the measurement vector at each time point follows a multivariate Gaussian distribution, whose parameters may vary across time. Throughout, we denote the measurement at time t as $\mathbf{y}(t)$ ($1 \leq t \leq T$), which represents a J -dimensional Gaussian random vector whose distribution is $\mathcal{N}(\boldsymbol{\mu}(t), \Sigma(t))$.

The goal of DCD is to detect temporal change points in functional connectivity and estimate a sparse connectivity graph for each segment, where the vertices are ROIs and the edges represent the relationship between ROIs. More specifically, we seek to partition the time series into several distinct segments, within which the data follows a multivariate Gaussian distribution with a different mean vector or covariance matrix from its neighboring segments. Further, for each segment we seek to estimate a graph representing connectivity between ROIs in the segment.

The DCR algorithm (Cribben et al., 2012, 2013) was previously developed to deal with the same problem. While, DCR has proven efficient at detecting changes in connectivity for problems consisting of a small to medium (<50) number of regions, it runs into computational problems as the number of regions becomes large (>100). The proposed DCD algorithm seeks to circumvent these issues by updating how (i) the underlying mechanisms by which change points are determined, and (ii) network structures are identified. Before discussing DCD in detail, we begin by giving a brief overview of DCR and sparse parameter estimation.

2.1. Dynamic Connectivity Regression (DCR)

The original DCR algorithm (Cribben et al., 2012), dealt with detecting change points in a group of subjects, but here we concentrate on the single subject case (Cribben et al., 2013). DCR aims at detecting temporal change points in functional connectivity and estimating a graph of the conditional dependencies between ROIs, for data that falls between each pair of change points. The measured signal is modeled as a Gaussian random vector where each element represents the activity of one region. The partitions in DCR are found using a regression tree approach. It attempts to first identify a candidate change point using the Bayesian Information Criterion (BIC), and then perform a permutation test to decide whether it is significant. If a significant change points is found, the same procedure is recursively applied to search for more changes points by further splitting the subset; see **Figure 1** for an illustration.

The required user specified input parameters for the algorithm are:

- 1) Δ : the minimum possible distance between adjacent changes points, chosen based on prior knowledge about the fMRI experiment.
- 2) $\lambda - list$: the full regularization path of tuning parameters λ required by GLASSO.
- 3) ξ : the mean block size of the stationary bootstrap.
- 4) α : the significance level for the permutation test.
- 5) N_b : the number of bootstrap samples.

Suppose we have a J -dimensional time series $\mathbf{Y} = \{\mathbf{y}(t)\}_{1 \leq t \leq T}$, where the $\mathbf{y}(t)$'s are assumed to be independent identically

distributed random variables which follow a multivariate Gaussian distribution. Here the mean vector can be estimated using the sample mean, and a sparse precision matrix can be estimated using the GLASSO technique (see next section for more detail). In order to choose the appropriate tuning parameter λ needed for GLASSO, the full regularization path $\lambda - list$ is run, and the optimal value is selected based on the value that minimizes the BIC. Finally, the model is refit without regularization, but keeping the zero elements fixed, and the optimized baseline BIC for the original time series, b_0 , is recorded.

For all possible split points t ($\Delta \leq t \leq T - \Delta$), the same procedure is repeated, and the BIC score for the two subsequences $\mathbf{Y}_1 = \{\mathbf{y}(t')\}_{1 \leq t' \leq t}$ and $\mathbf{Y}_2 = \{\mathbf{y}(t')\}_{t+1 \leq t' \leq T}$, denoted $b_1(t)$ and $b_2(t)$, respectively, are computed. A time point t_0 is chosen as a candidate change point, if it (i) produces the smallest combined BIC score $b_1(t_0) + b_2(t_0)$ for all possible split points t , and (ii) the combined BIC score is smaller than the baseline b_0 . In the continuation we let $\delta_b = b_0 - (b_1(t_0) + b_2(t_0))$ represent the decrease in BIC at t_0 .

Because change points are defined by a decrease in BIC, a random permutation procedure is used to create a $100(1 - \alpha)\%$ confidence interval for BIC reduction at the candidate change point t_0 , to determine whether it should be deemed a significant change point. Using a stationary bootstrap procedure with mean block size ξ , permuted time series are repeatedly created. Each time course is partitioned at time t_0 and the BIC reduction is computed as described above. The procedure is performed N_b times, thus allowing for the creation of a permutation distribution for the BIC reduction. If δ_b is more extreme than the $(1 - \alpha)$ quantile of the permutation distribution, we conclude t_0 is a significant change point. This procedure is recursively applied to each individual partition until no further split reduces the BIC score.

2.2. Sparse Parameter Estimation

The estimation of the covariance and precision matrix is a critical step in identifying candidate change points in the DCR algorithm. While the number of ROIs J is moderate, and the length of time series T is large, the sample covariance matrix S is a consistent estimator of the covariance matrix Σ . However, in high dimensional settings, when J is large compared to the sample size T , S has an infinite determinant, leading to divergence in the numerical algorithm. Thus, sparsity constraints are required to estimate the covariance, or precision matrix, consistently.

In this section we discuss two methods for performing sparse matrix estimation. While the original DCR method imposes sparsity on the precision matrix, the proposed DCD algorithm instead seeks to estimate a sparse covariance matrix. By making this shift, we can use a newly developed adaptive thresholding approach that provides a faster, more scalable solution to the change point problem described above. Statistically this changes the interpretation of the problem, as zeros in the precision matrix correspond to conditional independence between variables, while zeros in a covariance matrix correspond to marginal independence between variables. In a series of simulations and

an application to real data we examine the implications of this choice.

2.2.1. Graphical LASSO (GLASSO)

The Least Absolute Shrinkage and Selection Operator (LASSO) technique (Tibshirani, 1996), is often used for shrinkage and feature selection in regression problems. It adds an L_1 penalty term to the objective function, thus producing more interpretable models with some coefficients forced to be exactly zero. The Graphical LASSO (GLASSO) (Friedman et al., 2008) is an extension of this idea to graphical models, aimed at estimating sparse precision matrices. Based on the assumption that the observed data vectors $\{y(t)\}_{1 \leq t \leq T}$ follow a multivariate Gaussian distribution with covariance matrix Σ , it adds an L_1 norm penalty to the elements of the precision matrix $\Omega = \Sigma^{-1}$, and estimates the mean vector μ and precision matrix Ω by maximizing the penalized log-likelihood. After substituting the sample mean (the MLE of μ) into the objective function, this reduces to:

$$\log \det(\Omega) - \text{tr}(S\Omega) - \lambda \|\Omega\|_1$$

where S is the empirical covariance matrix, and the parameter λ controls the amount of regularization. Maximizing the penalized profile log-likelihood gives a sparse estimate of Ω .

If the ij^{th} element of matrix Ω is zero, the variables $y_i(t)$ and $y_j(t)$ are conditionally independent, given the other variables. We can therefore define a connectivity graph $G = (V, E)$ with the ROIs the vertices in V , and prune the edge between vertices i and j if the variables are conditionally independent. Thus, increasing the sparsity of Ω provides a sparser graphical representation of the relationship between the variables.

2.2.2. Adaptive Thresholding Approach

Here we introduce an adaptive thresholding approach that allows one to estimate a sparse covariance matrix. Again, assume the data $\{y(t)\}_{1 \leq t \leq T}$ follows an i.i.d. multivariate Gaussian distribution $\mathcal{N}(\mu, \Sigma)$. In this setting, the sample mean

$$\hat{\mu} = \frac{1}{T} \sum_{1 \leq t \leq T} y(t)$$

is a consistent estimator of μ .

To estimate the covariance matrix, we begin by using the empirical covariance matrix

$$\hat{\Sigma} = \frac{1}{T} \sum_{1 \leq t \leq T} (y(t) - \hat{\mu})(y(t) - \hat{\mu})^T$$

as a candidate estimator of Σ . To achieve sparsity we investigate whether individual elements should be set equal to zero following an idea of Cai and Liu (2011), where a method to model the distribution of $\hat{\Sigma}_{ij}$ is proposed.

Let $X_t^{ij} = (y_i(t) - \mu_i)(y_j(t) - \mu_j)$, where a subscript represents a single dimension of a vector, then the ij^{th} element of $\hat{\Sigma}$ is:

$$\hat{\Sigma}_{ij} = \frac{1}{T} \sum_{1 \leq t \leq T} X_t^{ij} = \bar{X}^{ij} \quad (1)$$

Now $X_1^{ij}, X_2^{ij}, \dots, X_T^{ij}$ is a sequence of i.i.d. random variables with $E[X_t^{ij}] = E[(y_i(t) - \mu_i)(y_j(t) - \mu_j)] = \Sigma_{ij}$ by definition, and further assume $\text{Var}[X_t^{ij}] = \delta_{ij}^2 < \infty$. Then by the Central Limit Theorem,

$$\sqrt{T}(\hat{\Sigma}_{ij} - \Sigma_{ij}) \rightarrow \mathcal{N}(0, \delta_{ij}^2)$$

A natural estimate of δ_{ij}^2 is given by:

$$\hat{\delta}_{ij}^2 = \frac{1}{T} \sum_{1 \leq t \leq T} (X_t^{ij} - \bar{X}^{ij})^2 \quad (2)$$

Alternatively, one can use the Jackknife technique to estimate the variance of estimator $\hat{\Sigma}_{ij}$ directly (see Appendix B).

Using this result, we can test $H_0 : \Sigma_{ij} = 0$ vs. $H_1 : \Sigma_{ij} \neq 0$ at significance level η as follows:

$$\left| \frac{\sqrt{T}\hat{\Sigma}_{ij}}{\hat{\delta}_{ij}} \right| = \frac{T|\hat{\Sigma}_{ij}|}{\sqrt{\sum_{t=1}^T (X_t^{ij} - \bar{X}^{ij})^2}} > z_{1-\eta/2}$$

If we successfully reject the null hypothesis, we can conclude that $\Sigma_{ij} \neq 0$ and keep $\hat{\Sigma}_{ij}$ as the estimator for Σ_{ij} . Otherwise we modify the candidate estimator and set $\hat{\Sigma}_{ij} = 0$. Similarly, using the diagonal elements of $\hat{\Sigma}$ as estimates of the variance of $\hat{\mu}$, we can perform a hypothesis testing for each element of μ and obtain a sparse estimate of $\hat{\mu}$. Since the testing procedure is performed for a potentially large number of parameters, we need to correct for multiple comparisons (Lindquist and Mejia, 2015).

2.3. Dynamic Connectivity Detection (DCD)

The DCD algorithm seeks to speed up the DCR algorithm, while achieving equivalent, or improved, results. The general procedure of DCD is similar to DCR, where a candidate split point is identified based on whether it further maximizes a likelihood-based function, and a hypothesis test is performed to decide whether this candidate split point is statistically significant. If a significant change point is found, the procedure is applied recursively to each of the two subsequences in order to find further split points.

The major improvement from DCR to DCD is that we incorporate the adaptive thresholding approach as our sparse matrix estimation method, which successfully improves upon the computational efficiency. In addition, during each step, a binary “mask” representing the non-zero parameter elements (in the mean vector and covariance matrix) is saved for each partition. If an additional change point is found for this partition, the “mask” is imposed on the parameters of both “child” partitions (the two subsets of time series created by splitting the data at the change point). This implies that if the estimate of one element of the covariance matrix for some partition is zero, then the estimate of corresponding element in any sub-partition will also be zero. The recursive sparsity feature is illustrated in **Figure 2**.

All input parameters in DCD have a clear statistical interpretation, enhancing its user-friendliness. The required user specified input parameters for the algorithm are:

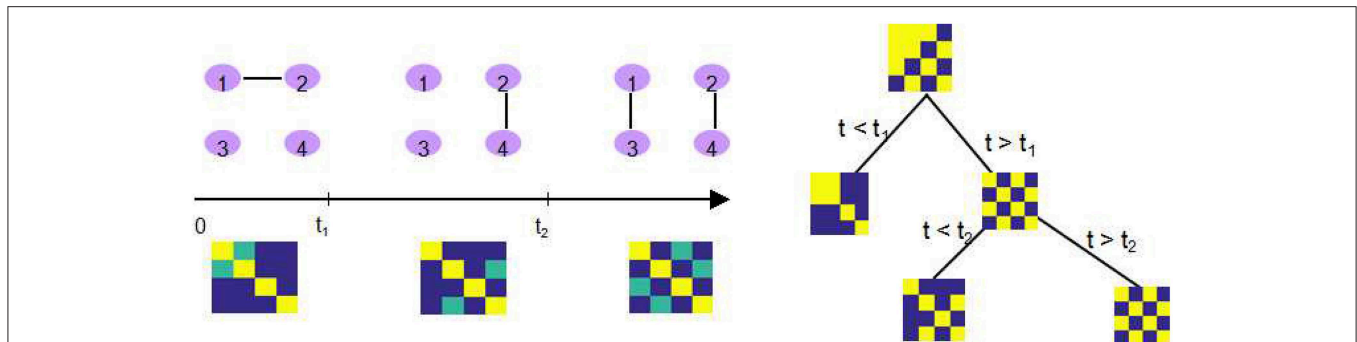


FIGURE 2 | An illustration of how the sparsity structure is recorded in DCD. Left: The split points t_1 and t_2 , and the corresponding covariance matrix within each partition. Yellow elements in the covariance matrix plot represents 1, green elements 0.5, and blue elements 0. **Right:** DCD uses a binary mask to record the sparsity structure at each node of the binary search tree.

- 1) α : the type I error bounds for the hypothesis tests used to determine significant splits.
- 2) β : the type II error bounds for the hypothesis test used to determine significant splits.
- 3) η : the type I error bound for the hypothesis test used to determine the sparsity of the covariance matrix.

Since the length of the time series partition affects statistical inference, we need to calculate the minimum partition length Δ needed to achieve the desired error bounds. We apply a power analysis based on a two sample t -test to calculate Δ from the inputs α and β ; for details please refer to Appendix A.

Given a J -dimensional time series $\mathbf{Y} = \{\mathbf{y}(t)\}_{1 \leq t \leq T}$, we begin by calculating the maximized baseline log-likelihood L_0 under the assumption that

$$\mathbf{y}(t) \stackrel{i.i.d}{\sim} \mathcal{N}(\boldsymbol{\mu}_0, \Sigma_0), \quad 1 \leq t \leq T.$$

Hence, the log-likelihood function is given by

$$L(\boldsymbol{\mu}_0, \Sigma_0 | \mathbf{Y}) \propto - \sum_{t=1}^T (\mathbf{y}(t) - \boldsymbol{\mu}_0)^T \Sigma_0^{-1} (\mathbf{y}(t) - \boldsymbol{\mu}_0) - T \log(\det \Sigma_0). \quad (3)$$

We first calculate the sample mean and sample covariance matrix as the maximum likelihood estimator of $\boldsymbol{\mu}_0$ and Σ_0 , and then further improve the estimator by performing the adaptive thresholding method described in Section 2.2.2, in order to obtain a sparse mean vector $\hat{\boldsymbol{\mu}}_0$ and sparse covariance matrix $\hat{\Sigma}_0$.

The maximized log-likelihood function can now be expressed as:

$$L_0 = -T \left(\text{tr}(\hat{\Sigma}_0^{-1} S) + \log(\det \hat{\Sigma}_0) \right)$$

where S is the normalized scatter matrix:

$$S = \frac{1}{T} \sum_{1 \leq t \leq T} (\mathbf{y}(t) - \hat{\boldsymbol{\mu}}_0)^T (\mathbf{y}(t) - \hat{\boldsymbol{\mu}}_0)$$

While calculating the sparse structure of parameter $\boldsymbol{\theta}_0 = (\boldsymbol{\mu}_0, \text{vec}\{\Sigma_0\})$, a binary array *mask* is saved, indicating the non-zero elements of $\boldsymbol{\theta}_0$. It is assumed that any subsequence of the time series will satisfy the parent sparsity property.

For any possible candidate split point t ($\Delta \leq t \leq T - \Delta$), assume the two subsequences $\mathbf{Y}_1 = \{\mathbf{y}(t')\}_{1 \leq t' \leq t}$ and $\mathbf{Y}_2 = \{\mathbf{y}(t')\}_{t+1 \leq t' \leq T}$ follow multivariate Gaussian distribution with parameters $\boldsymbol{\theta}_{1t} = (\boldsymbol{\mu}_{1t}, \text{vec}\{\Sigma_{1t}\})$ and $\boldsymbol{\theta}_{2t} = (\boldsymbol{\mu}_{2t}, \text{vec}\{\Sigma_{2t}\})$, respectively. Here only the upper triangular elements are used when vectorizing the covariance matrix. The dimension of the parameter vector is therefore $J + J*(J+1)/2 = (J+1)(J+2)/2$. Next, the maximum likelihood estimators $\hat{\boldsymbol{\theta}}_{it}^{ML}$ ($i = 1, 2$) are computed, imposing the parent sparsity structure by taking the Hadamard product with the *mask* vector:

$$\hat{\boldsymbol{\theta}}_{it} = \hat{\boldsymbol{\theta}}_{it}^{ML} \otimes \text{mask}, \quad i = 1, 2$$

Now the maximized log-likelihood under current split point t can be obtained as follows:

$$L_t = L(\hat{\boldsymbol{\theta}}_{1t} | \mathbf{Y}_1) + L(\hat{\boldsymbol{\theta}}_{2t} | \mathbf{Y}_2).$$

Similar to DCR we can now step through all possible candidate split points and find the one, denoted t_0 , which shows the maximum improvement in log-likelihood L_t compared to L_0 :

$$t_0 = \underset{t}{\text{argmax}} (L_t - L_0) +$$

If the maximum L_{t_0} is less than the baseline L_0 , the DCD procedure returns no detected split points; otherwise a set of hypothesis tests are performed to determine whether t_0 is a significant change point.

For the sake of clarity, denote the Gaussian distribution parameters of the two subsequences as $\boldsymbol{\theta}_i = (\boldsymbol{\mu}_i, \text{vec}\{\Sigma_i\}) =: \boldsymbol{\theta}_{it}$, ($i = 1, 2$). We now seek to test:

$$H_{j0} : \boldsymbol{\theta}_1(j) = \boldsymbol{\theta}_2(j) \text{ vs. } H_{j1} : \boldsymbol{\theta}_1(j) \neq \boldsymbol{\theta}_2(j), \quad j \in \{j' : \text{mask}(j') = 1\} \quad (4)$$

If any of the non-zero parameters are significantly different for the two subsequences, i.e., if we reject any of the null hypotheses, then we conclude that t_0 is a significant change point for partitioning the time series \mathbf{Y} . We use Bonferroni correction to control the family-wise error rate (FWER), and reject H_{j0} if the p -value is less than $\frac{\alpha}{\sum_j \text{mask}(j')}$.

To perform each test we use Welch's t -test (two-sample t -test for unequal variance). For $j \leq J$, use the diagonal element of $\hat{\Sigma}$ as an estimate of the variance of $\hat{\mu}$; and for $j > J$, use the estimator described in Equation (2) to estimate the variance of each element of $\hat{\Sigma}$. If t_0 is identified as a significant change point, continue searching for more change points by recursively repeating the above procedure on the two "child" subsequences until no further change points are returned; otherwise finish the DCD procedure by returning a null value.

The complete procedure for performing the DCD algorithm is summarized below:

1. Take the input parameters α , β , η , and calculate the minimum partition length Δ as described in Appendix A.
2. Consider the full multivariate time series with length T , calculate the sparsity structure of its multivariate normal distribution parameters as described in Section 2.2.2, estimate the mean vector and a sparse covariance matrix accordingly, and calculate the baseline likelihood function L_0 .
3. For each value of t ranging from Δ to $T - \Delta$, partition the time series into two subsequences $\{1 : t\}$ and $\{t + 1 : T\}$, calculate the sparsity structure of parameters based on the parent sparsity structure from Step (2), then calculate the combined likelihood function using the estimated sparse parameters.
4. Find the time point which produces the largest increase in combined likelihood function, perform the hypothesis test described in Equation (4) to determine whether it is a significant change point. If yes, split the time series into two partitions accordingly.
5. Apply Steps (2–4) recursively to each partition until no further change points are found.
6. After detecting all change points, estimate a connectivity graph for each partition using a sparse matrix estimation technique, such as Adaptive Thresholding Approach to obtain a covariance graph or GLASSO to obtain a connectivity graph.

3. Simulations

A series of simulations were performed to test the efficacy of the new DCD algorithm, and compare its performance to the DCR method. For this reason, we adopt simulation settings inspired by those found in the original DCR work (Cribben et al., 2012). However, in contrast to that work, for each simulation the connectivity pattern and strength between nodes remains the same across different subjects, since our focus is on the single subject case instead of on group-level inference. In addition, the object of each simulation in this paper is focused on identifying the timing of the connectivity change points, rather than explicitly assessing the quality of the estimation of the underlying graphs.

The descriptions and parameter settings for each simulation are listed below. Here N , T , and p represent the number of subjects, the length of the time series, and the number of regions, respectively. The true dependency between ROIs (i.e., the precision matrices) are shown as heat maps in **Figures 3–7**. More details regarding the exact strength of these connections can be found in Appendix C. Here the notation $(i, j) = k$ indicates that the (i, j) element of the precision matrix takes the value k . All unspecified diagonal elements are one and non-diagonal elements are zero. In the latter case, the ROIs were made up of i.i.d. Gaussian noise indicating a lack of functional connectivity. Hence, each simulation is created assuming sparsity in the precision matrix, which should theoretically benefit DCR over DCD, which imposes sparsity in the covariance matrix.

For each simulation, both the DCD and DCR approaches were applied to the N subjects individually. Since the DCR algorithm has many parameters, and according to previous work several are insensitive to change, we fix several of them as follows:

$$\Delta = 50, \lambda - \text{list} = (2^0, 2^{-1}, \dots, 2^{-9}), N_b = 50, \xi = \Delta/2.$$

For DCD, we fix $\eta = 0.05$. All remaining parameters are altered depending on the simulation setting.

Below we list a brief description of each simulation study.

• Simulation 1

Description: The data is white noise with no connectivity change points.

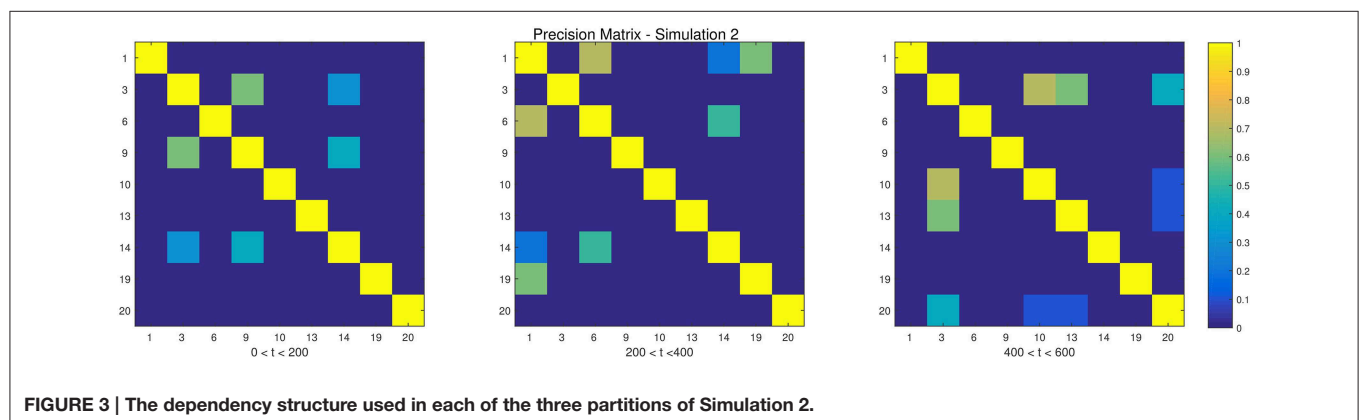
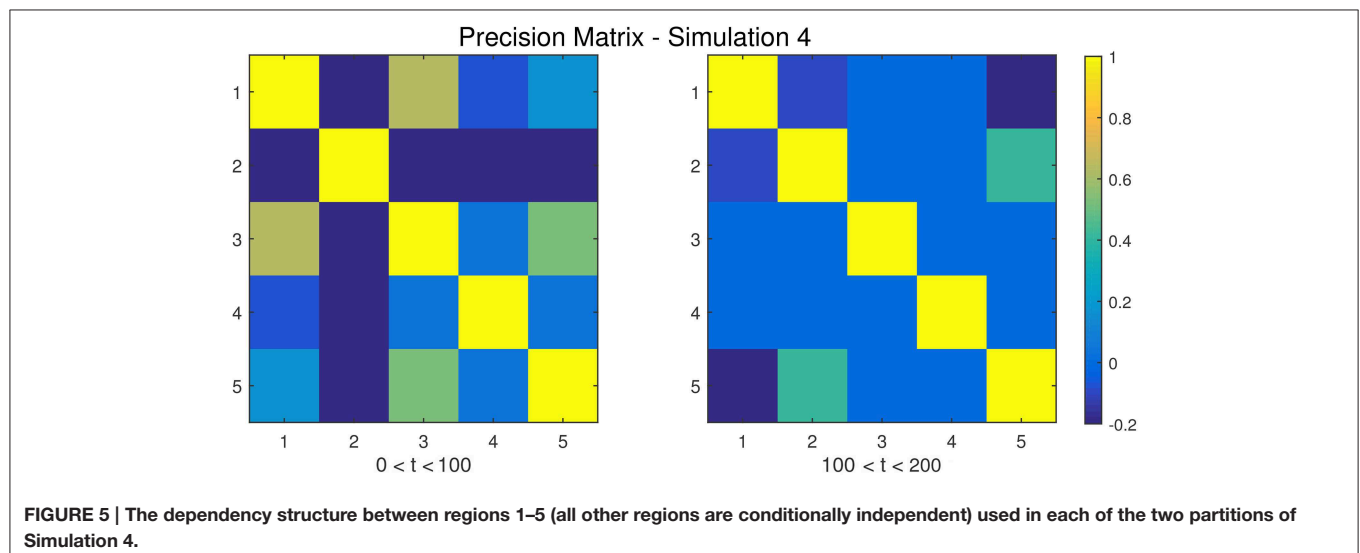
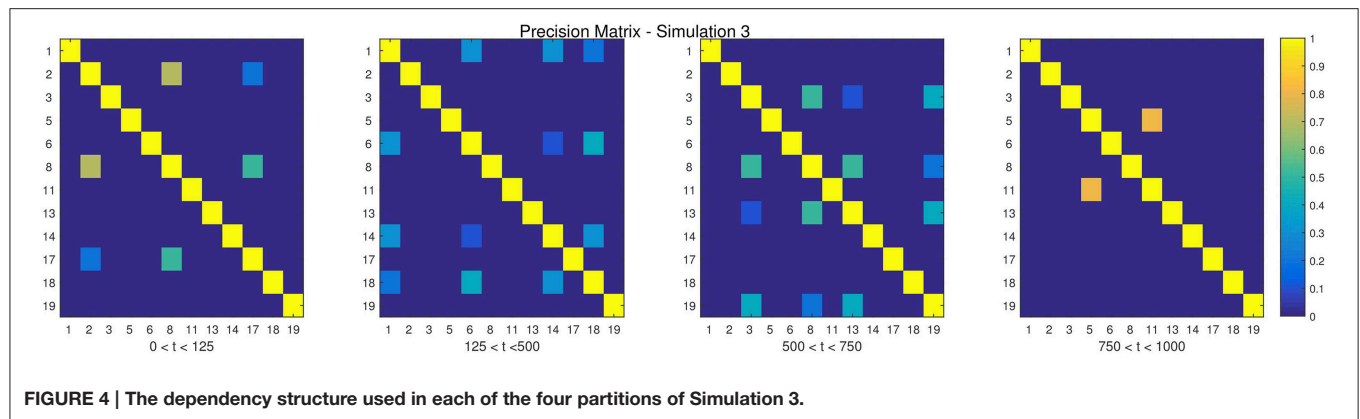


FIGURE 3 | The dependency structure used in each of the three partitions of Simulation 2.



Size: $N = 20$, $T = 1000$, $p = 20$

DCD parameters: $(\alpha, \beta) = (0.05, 0.1)$;

DCR parameters: $\alpha = 0.05$.

• Simulation 2

Description: There are two change points at times 200 and 400. Spikes are imposed onto the time series, imitating a common artifact found in fMRI data. For each subject there are 5 randomly placed spikes, each with magnitude 15.

Size: $N = 20$, $T = 1000$, $p = 20$

DCD parameters: $(\alpha, \beta) = (0.05, 0.1)$;

DCR parameters: $\alpha = 0.05$.

• Simulation 3

Description: There are three change points at times 125, 500, and 750.

Size: $N = 15$, $T = 1000$, $p = 20$

DCD parameters: $(\alpha, \beta) = (0.05, 0.05)$;

DCR parameters: $\alpha = 0.05$.

• Simulation 4

Description: There is a single change point at time 100.

Size: $N = 25$, $T = 200$, $p = 5$

DCD parameters: $(\alpha, \beta) = (0.05, 0.1)$;

DCR parameters: $\alpha = 0.05$.

• Simulation 5

Description: There are five change points at times 200, 300, 500, 600, and 800.

Size: $N = 20$, $T = 1000$, $p = 20$

DCD parameters: $(\alpha, \beta) = (0.05, 0.05)$;

DCR parameters: $\alpha = 0.05$.

• Simulation 6

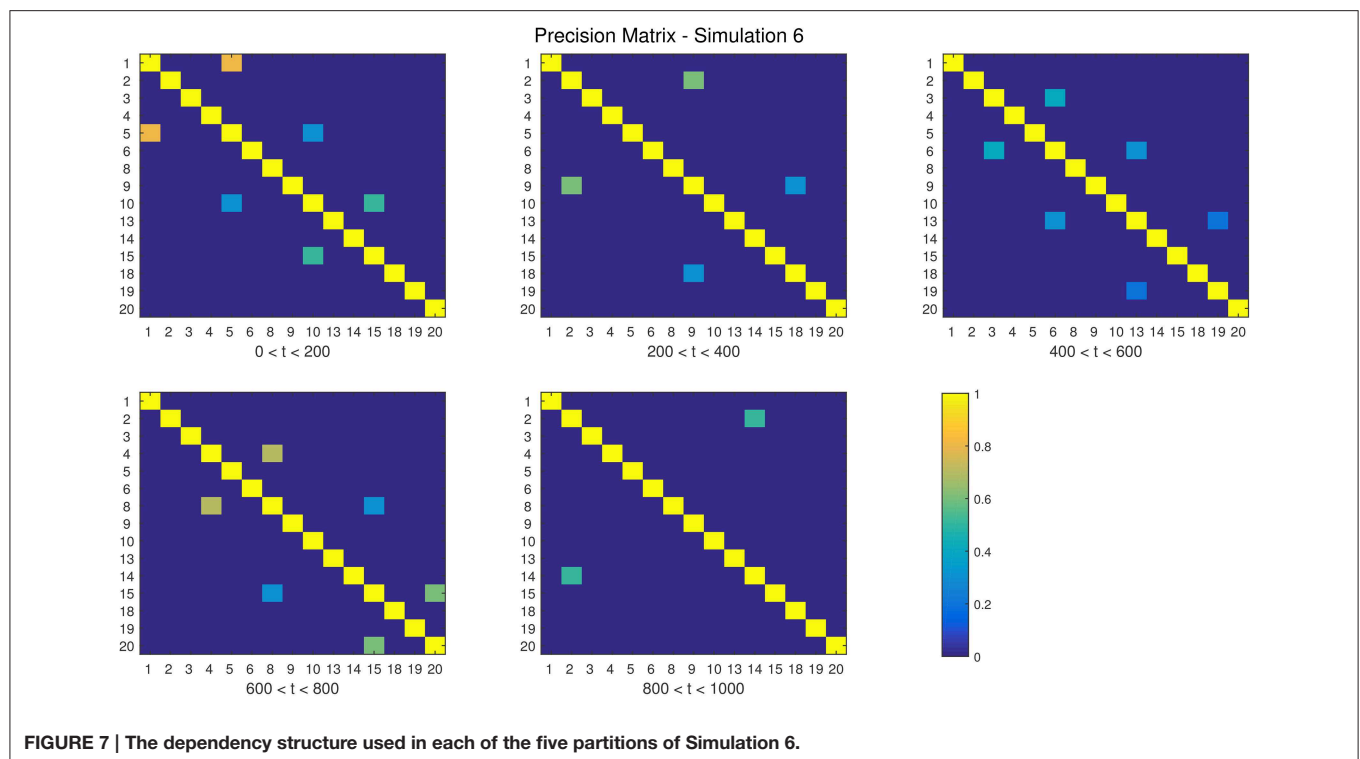
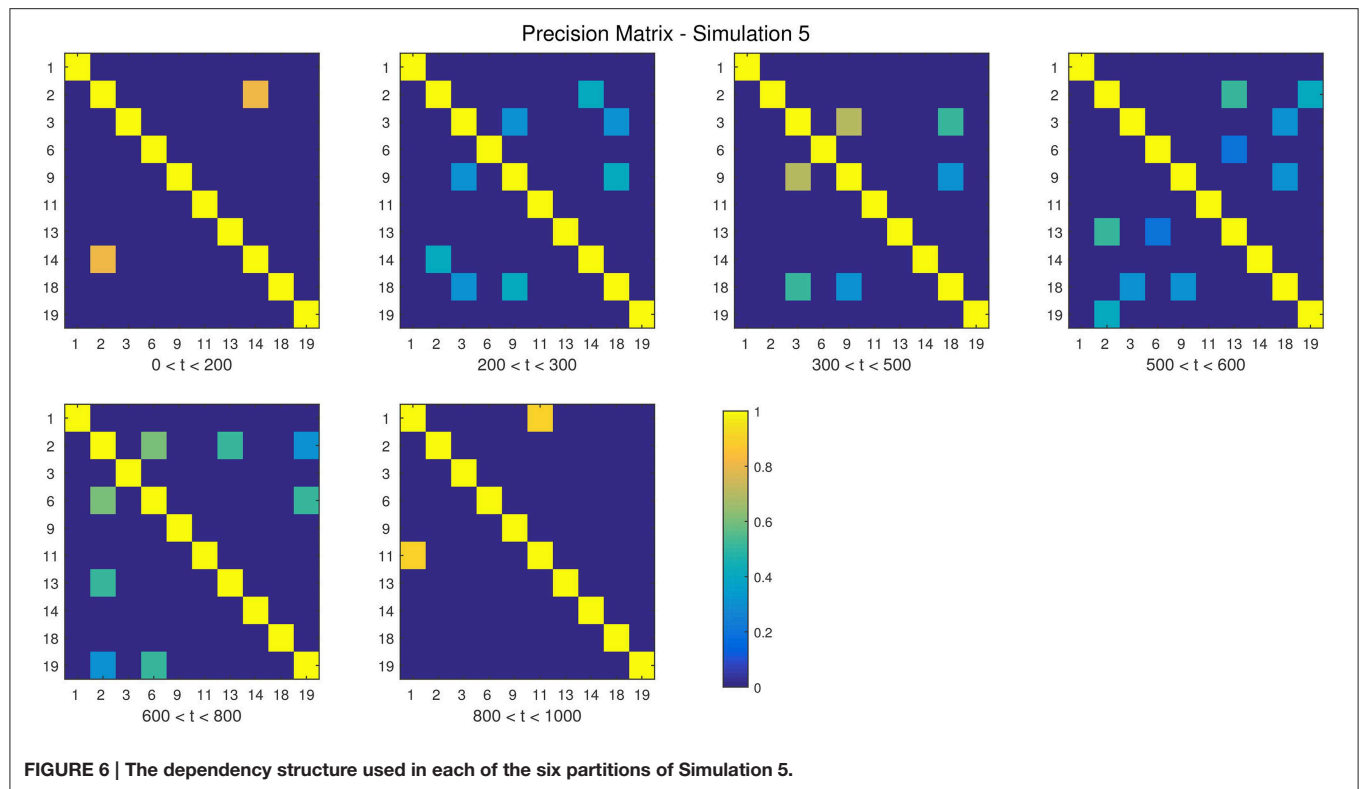
Description: There are four change points at times 200, 400, 600, and 800.

Size: $N = 20$, $T = 1000$, $p = 20$

DCD parameters: $(\alpha, \beta) = (0.05, 0.05)$;

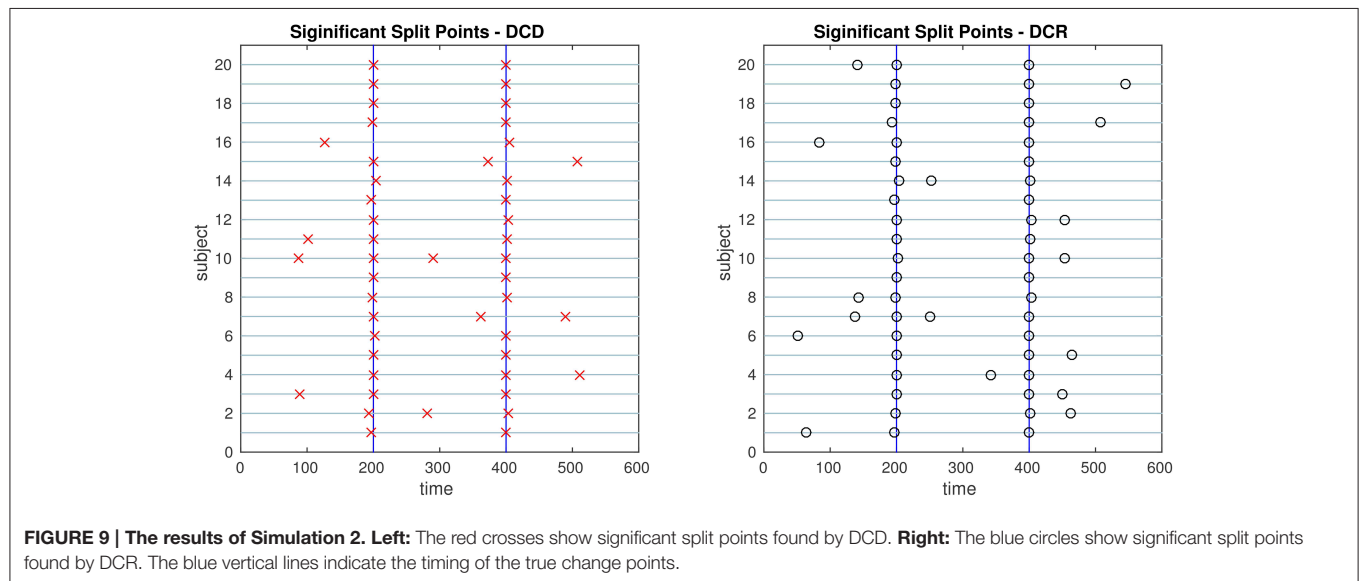
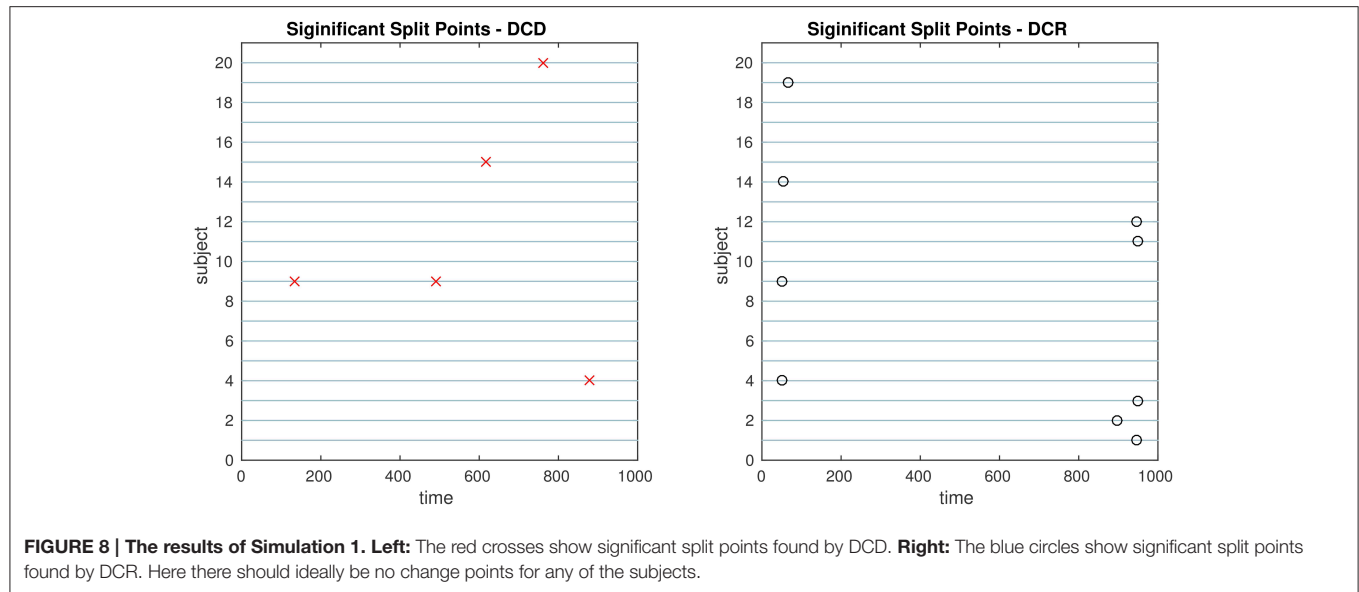
DCR parameters: $\alpha = 0.05$.

The results of the simulations are shown in **Figures 8–13**. In each figure, the y-axis represents the subject number, while the x-axis represents time points. All red crosses in the left sub figures represent change points detected for each subject by DCD, and the blue circles are those detected by DCR. The blue vertical line indicates the true change points for each simulation setting. In **Table 1**, we list the respective runtimes of DCD and DCR for each simulation. The computing platform used was an Intel Core i5-3210M CPU 2.5 GHz with 16.0 GB RAM.



The results of Simulation 1, where there are no true change points, are shown in **Figure 8**. The DCD algorithm finds 5 false positive change points, whereas the DCR algorithm finds 9.

Interestingly, the DCR false positives are primarily grouped at the time points Δ and $T - \Delta$. The reason for this is that when adding the BIC score from two sub-series of lengths n_1 and n_2 , where



$n_1 + n_2 = n$, and assuming the number of parameters $k_1 \approx k_2 \approx k$, the total penalty term is $k \log(n_1) + k \log(n_2) \propto \log(n_1(n - n_1))$, which favors small or large values of n_1 when minimizing the BIC. In addition, the runtime of DCD is approximately 30 times faster than DCR, providing a significant decrease in computation time.

The results of Simulation 2 are shown in **Figure 9**. Here there exist two true change points, the first at time 200, and the second at time 400. In addition, there are 5 spikes placed at random time points for each subject. Both algorithms do a good job of detecting the true change points in most cases, with a few instances of false positives for each. Here DCD is approximately 60 times faster than DCR in obtaining the results.

The results of Simulation 3 are shown in **Figure 10**. Here there exist three true change points, the first at time 125, the second at

time 500, and the third at time 750. Clearly, both algorithms do an excellent job of detecting the true change points. Here DCD is approximately 30 times faster than DCR in obtaining the results.

Figure 11 shows the results of Simulation 4. Again, both algorithms do an excellent job of detecting the true change point, which is located at time 100, but DCD does so with a 20-fold increase in speed.

Finally, the results of Simulations 5 and 6 are shown in **Figures 12, 13**, respectively. In both cases the algorithms do an excellent job of detecting the true change points. However, DCD does so with a 30-fold increase in speed in both cases.

Although the main goal of DCD is to detect change points, and the estimation of a connectivity graph seems a byproduct, the accuracy of the covariance matrix or precision matrix estimation leads to better change point detection, and vice versa. Using the

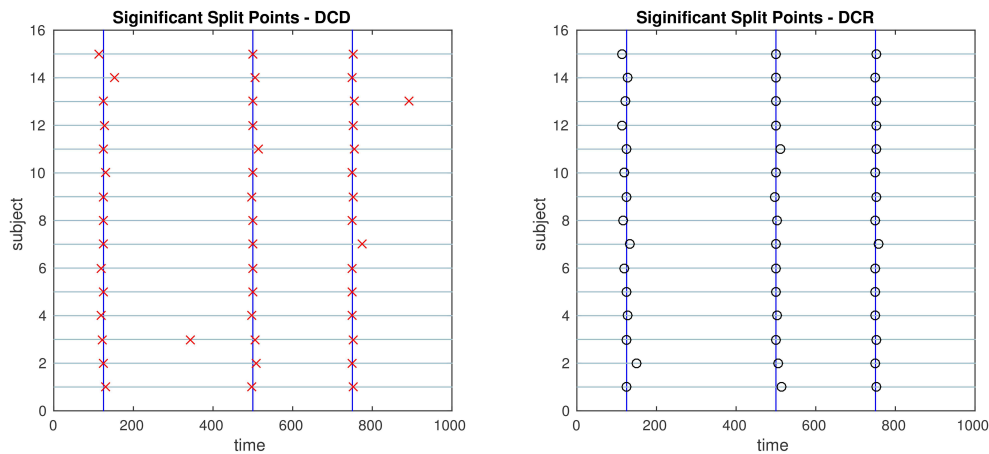


FIGURE 10 | The results of Simulation 3. Left: The red crosses show significant split points found by DCD. **Right:** The blue circles show significant split points found by DCR. The blue vertical lines indicate the timing of the true change points.

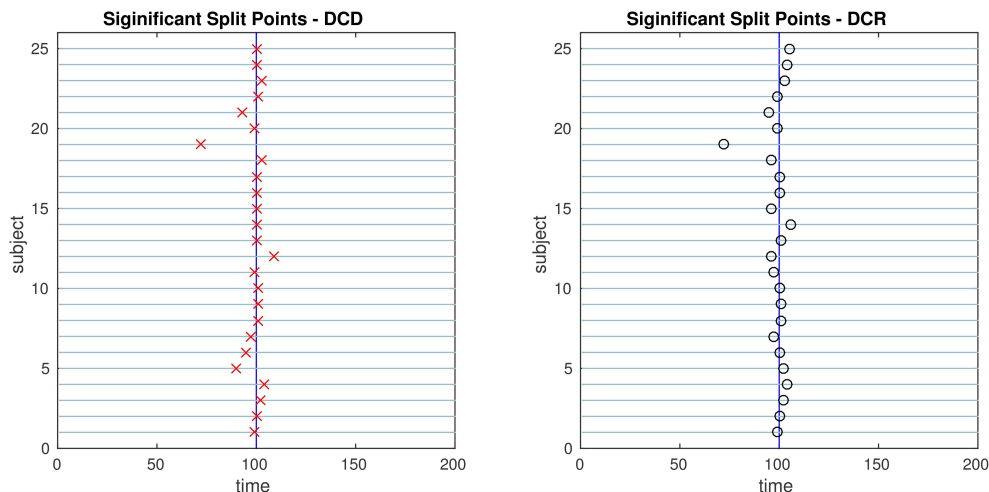


FIGURE 11 | The results of Simulation 4. Left: The red crosses show significant split points found by DCD. **Right:** The blue circles show significant split points found by DCR. The blue vertical line indicates the true change points.

Adaptive Thresholding Approach, we need to control the family-wise error rate or false discovery rate. The estimation of a J -dimensional covariance matrix requires $O(J^2)$ hypothesis tests. In our simulation examples, we adjust the significance level η by η/J , to guard against being as conservative as Bonferroni correction, while still obtaining adequate control over the family-wise error rate. Results show that the estimation of the sparsity structure is accurate in most simulations. The list of the average proportion of correctly identified zero/non-zero elements of the covariance matrices are listed in **Table 2**.

In summary, in each of the “low dimensional” simulations described above, with the number of ROIs ~ 20 , DCR achieves similar results as DCD with a significant speed-up in runtime. However, to investigate how well the methods scale to a more “high dimensional” settings, we expand upon two of the simulations to inspect how computational time

changes as a function of the number of ROIs for the two algorithms.

In the first (denoted 2B), we generated 80 ROIs data for 50 subjects under the same settings as described in Simulation 2. Here only the first 20 nodes contain information, and the remaining are simply white noise. We ran DCD and DCR using ROIs 1: r , where r ranged from 20 to 80 in increments of 5. In the second (denoted 4B), we generated 70 ROIs for 50 subjects under the same settings as described in Simulation 4. Here only the first 5 nodes contain information, while all remaining nodes are white noise. We ran DCD and DCR on a subset of ROIs numbered 1: r , where r ranged from 5 to 70 in increments of 5.

The results of Simulation 2B are summarized in **Figures 14, 15**. From **Figure 14** it is clear that the computation time for DCR increases exponentially with the number of ROIs, while the computation time for DCD is much shorter and nearly linear.

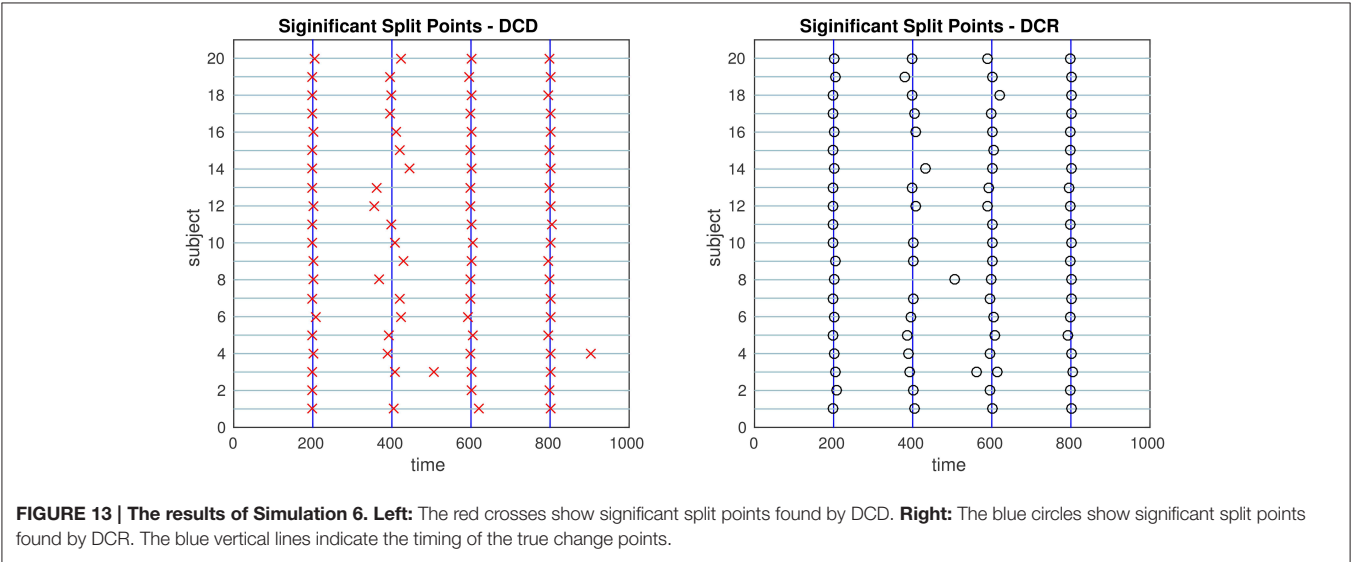
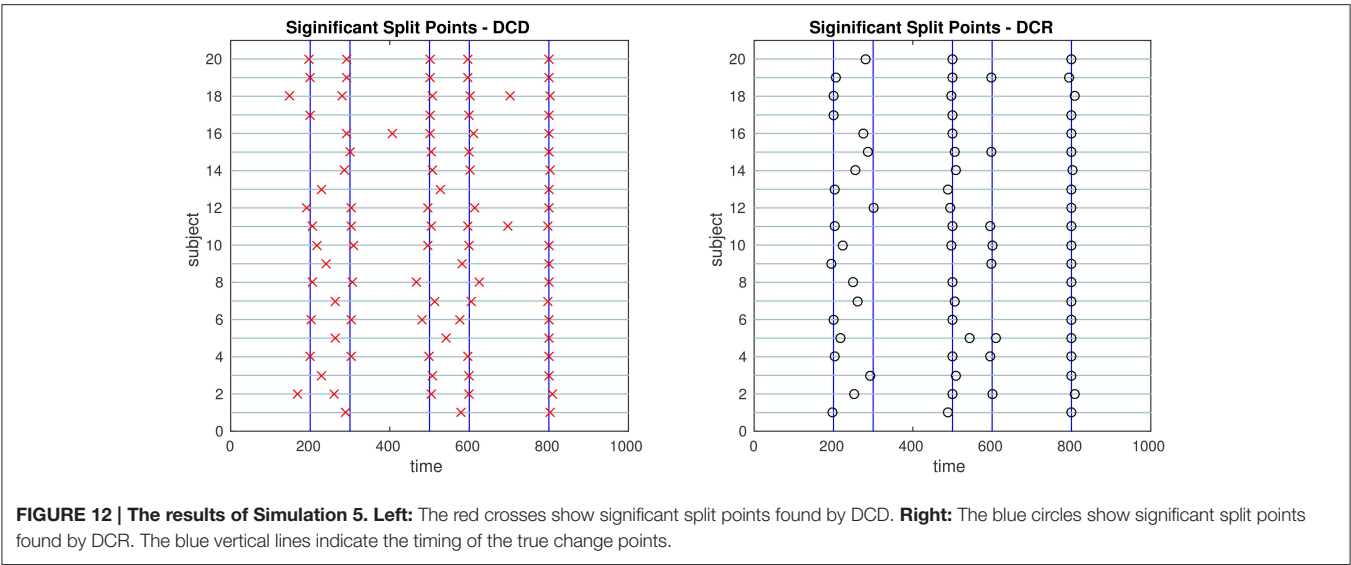


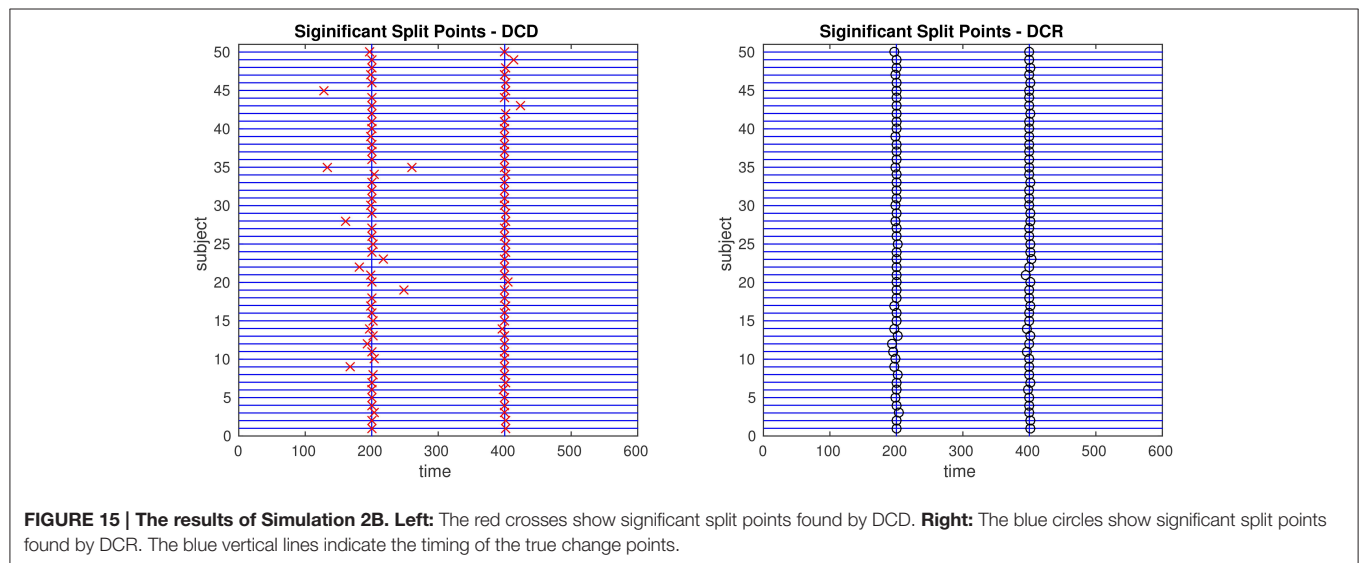
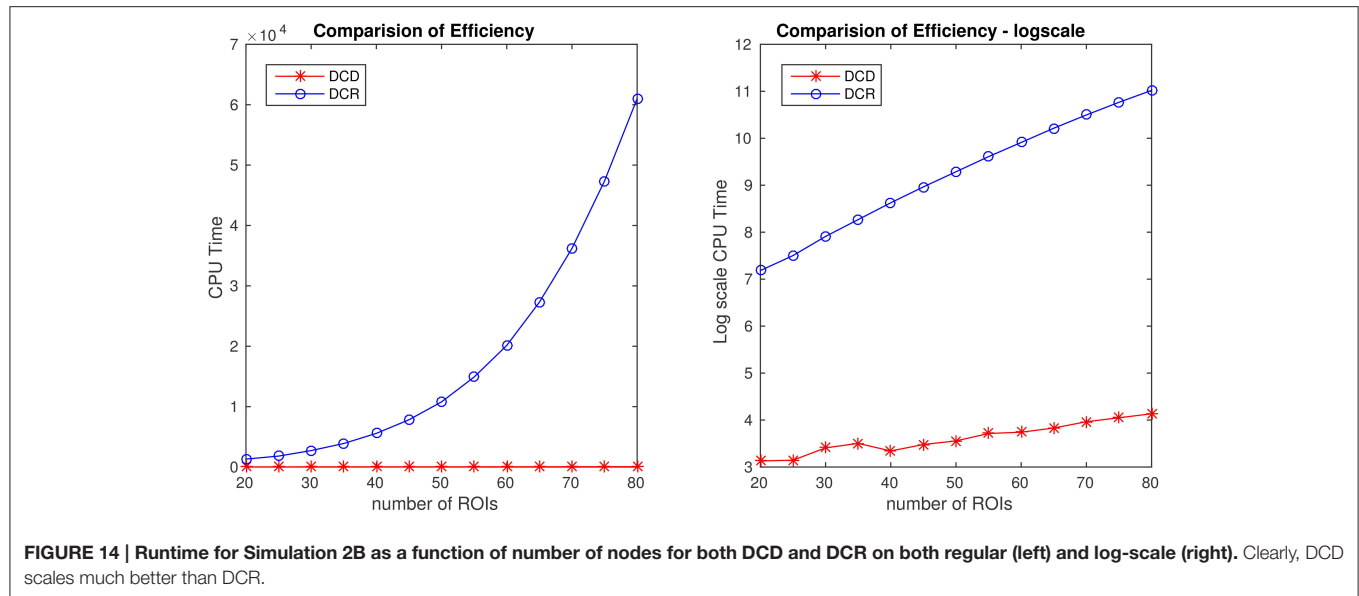
TABLE 1 | Runtime comparison between the DCD and DCR algorithms for each simulation.

	Simulation 1	Simulation 2	Simulation 3	Simulation 4	Simulation 5	Simulation 6
DCR run time	249.163458	351.531306	408.477891	13.262008	585.510362	581.070222
DCD run time	8.005733	5.756054	12.363800	0.712059	16.958910	18.126004
ratio = $\frac{DCRtime}{DCDtime}$	31.1231	61.0716	33.0382	18.6249	34.5252	32.0573

Runtime is measured in units of seconds.

TABLE 2 | Sparsity control results of covariance matrices.

Simulation	1	2	3	4	5	6
Correct zero rate	0.9497	0.9568	0.9474	0.9529	0.4955	0.9461
Correct non-zero rate(TP)	1	0.9763	0.9644	0.6535	0.9837	0.9628
False positive rate (average)	0.0503	0.0432	0.0526	0.0471	0.5045	0.0539



Though the results of DCR appear slightly better than DCD (see Figure 15), with less deviations from the true change points, this comes at a substantial computational cost.

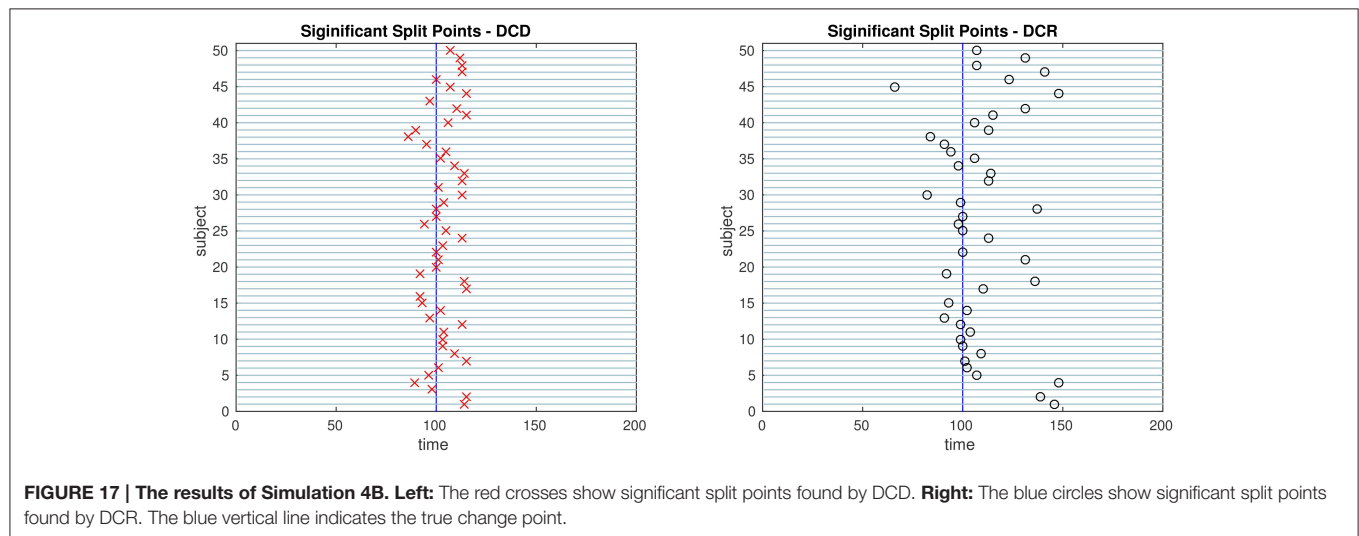
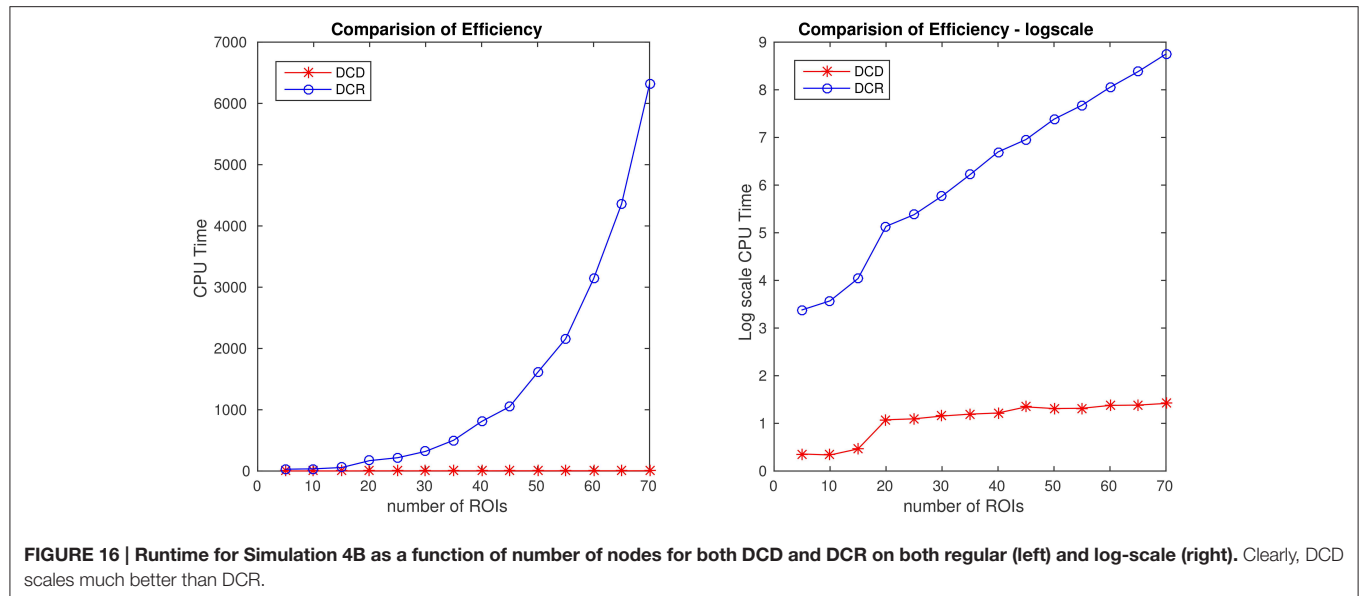
The results of Simulation 4B are summarized in Figures 16, 17. Based on Figure 16 it is clear that the computation time for DCR increases exponentially with the number of ROIs, while the computation time for DCD is much shorter and has a near linear increase. In addition, judging by Figure 17 the algorithm also appears to more accurately detect the timing of the true change points.

4. Application to Experimental Data

4.1. Social Evaluative Threat Experiment

The data was taken from an experiment where subjects performed an anxiety-inducing task while fMRI data was

acquired (Wager et al., 2009). This is the same data set used in the previous DCR papers (Cribben et al., 2012, 2013), as well as in other papers exploring mean change points (Lindquist et al., 2007; Robinson et al., 2010). The task was a variant of a well-studied laboratory paradigm for eliciting social threat, during which participants were asked to give a speech under evaluative pressure. It consisted of an off-on-off design, with an anxiety-provoking speech preparation task sandwiched between two lower-anxiety rest periods. Prior to the scanning session, subjects were informed that they were to be given 2 min to prepare a 7 min speech, the topic of which would be revealed to them during scanning, that would be delivered to a panel of expert judges after the scanning session. However, they were told that there was a small chance that they would be randomly selected not to give the speech. After the start of fMRI acquisition, during the initial 2 min resting period subjects viewed a fixation cross. At the end



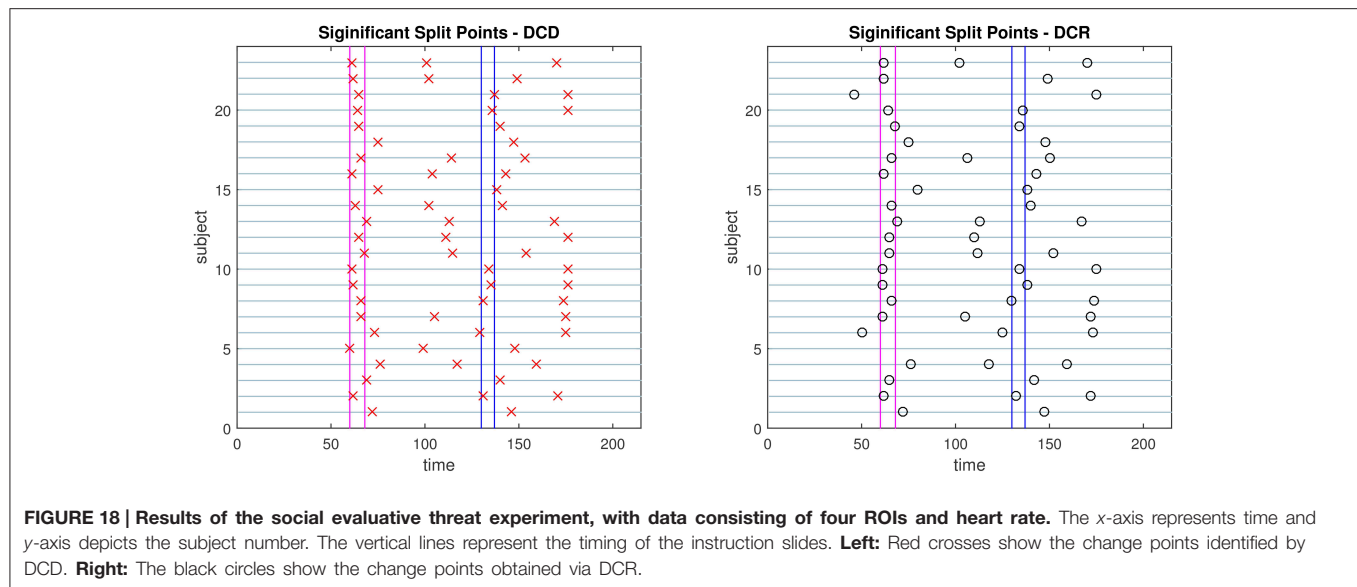
of this period, an instruction slide appeared describing the speech topic for 15 s (“why you are a good friend”). The slide instructed subjects to prepare enough for the entire 7 min period. After 2 min of silent preparation, a second instruction screen appeared for 15 s that informed subjects that they would not have to give the speech. The functional run concluded with an additional 2 min period of resting baseline.

During the course of the experiment a series of 215 functional images were acquired (TR = 2 s). A detailed description of the data acquisition and preprocessing can be found in previous work (Wager et al., 2009). In order to create ROIs, time series of voxels were averaged across pre-specified regions of interest. We used data consisting of 4 ROIs and heart rate for 23 subjects. The 4 ROIs were chosen due to the fact that they showed a significant relationship to heart rate in an independent data set. They included the ventral medial prefrontal cortex (VMPFC), the

anterior medial prefrontal cortex (mPFC), the striatum/pallidum, and the dorsal lateral prefrontal cortex (DLPFC)/inferior frontal junction (IFJ). The temporal resolution of the heart rate was 1 s compared to 2 s for fMRI data, so it was down-sampled by taking every other measurement.

Both the DCD and DCR approaches were applied to the 23 subjects individually. For the DCD algorithm, we used $(\alpha, \beta, \eta) = (0.1, 0.1, 0.05)$ as input parameters, and the runtime was 0.92 s. For the DCR algorithm, we adopted similar parameter settings used in Cribben et al. (2013), where we used the following settings: $\Delta = 40$, $\lambda - list = (1, 2^{-1}, \dots, 2^{-9})$, $\alpha = 0.1$, $N_b = 50$, and $\xi = 20$. The runtime for DCR was 32.14 s.

The change points detected by the two algorithms are displayed in Figure 18. Both consistently give rise to change points around the time of the first visual cue. In addition, there appear to be changes toward the middle of speech preparation

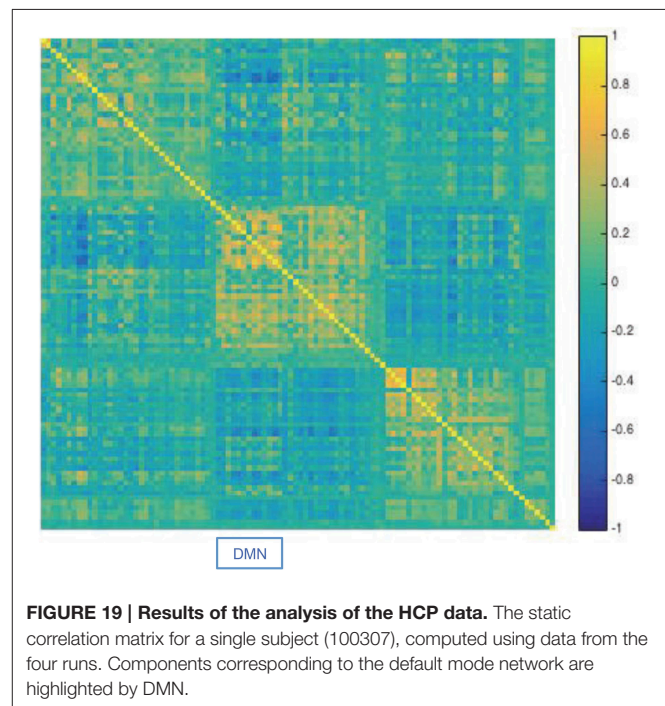


and around the time of the second visual cue, though these are less consistent across subjects. Interestingly, in contrast to the DCR algorithm, the first change points given by the DCD algorithm appears to coincide more closely to the timing of the first introduction cue. Otherwise the number, and placement, of the detected change points are roughly equivalent across methods.

4.2. Human Connectome Project

To study DCD's performance on high dimensional data, we applied the method to resting-state fMRI (rfMRI) data from the 2014 Human Connectome Project (HCP) data release (Van Essen et al., 2013). The data consists of 4 separate 15 min rfMRI runs, each consisting of 1200 time points, collected for each of 468 subjects. Each run was minimally preprocessed according to the procedure outlined in Glasser et al. (2013), with artifacts removed using FIX (FMRIB's ICA-based Xnoiseifier) (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Each data set was temporally demeaned with variance normalization applied according to Beckmann and Smith (2004). Group-PCA output was generated by applying MELODICs Incremental Group-PCA on the 468 subjects. This comprises the top 4500 weighted spatial eigenvectors from a group-averaged PCA. The output was fed into group-ICA using FSL's MELODIC tool (Beckmann and Smith, 2004), applying spatial-ICA with 100 distinct ICA components. The set of ICA spatial maps were mapped onto each subject's time series data to obtain a single representative time series per ICA component using the "dual-regression" approach, in which the full set of ICA maps are used as spatial regressors against the full data (Filippini et al., 2009).

For illustration purposes we applied DCD to data consisting of 100 ICA component time courses from a single subject (100307). We began by computing the static correlation matrix for the subject by concatenating data across the four runs. The resulting correlation matrix was sorted using the Louvain algorithm (Blondel et al., 2008), which has proven efficient



at identifying communities in large networks. The resulting correlation matrix can be seen in **Figure 19**. There are clear groupings of similar components that correspond to common networks seen in the resting-state literature, including the visual, somatomotor, cognitive control, and default mode networks.

Next, we applied DCD with input parameters $(\alpha, \beta, \eta) = (0.05, 0.05, 0.02)$ to each of the four runs. The runtime for each was less than 10 s. The correlation matrices for all partitions are displayed in **Figure 20**, along with the corresponding temporal partition listed above them. Each run consisted of either 6 or 7 partitions, and there are clear similarities in connectivity states

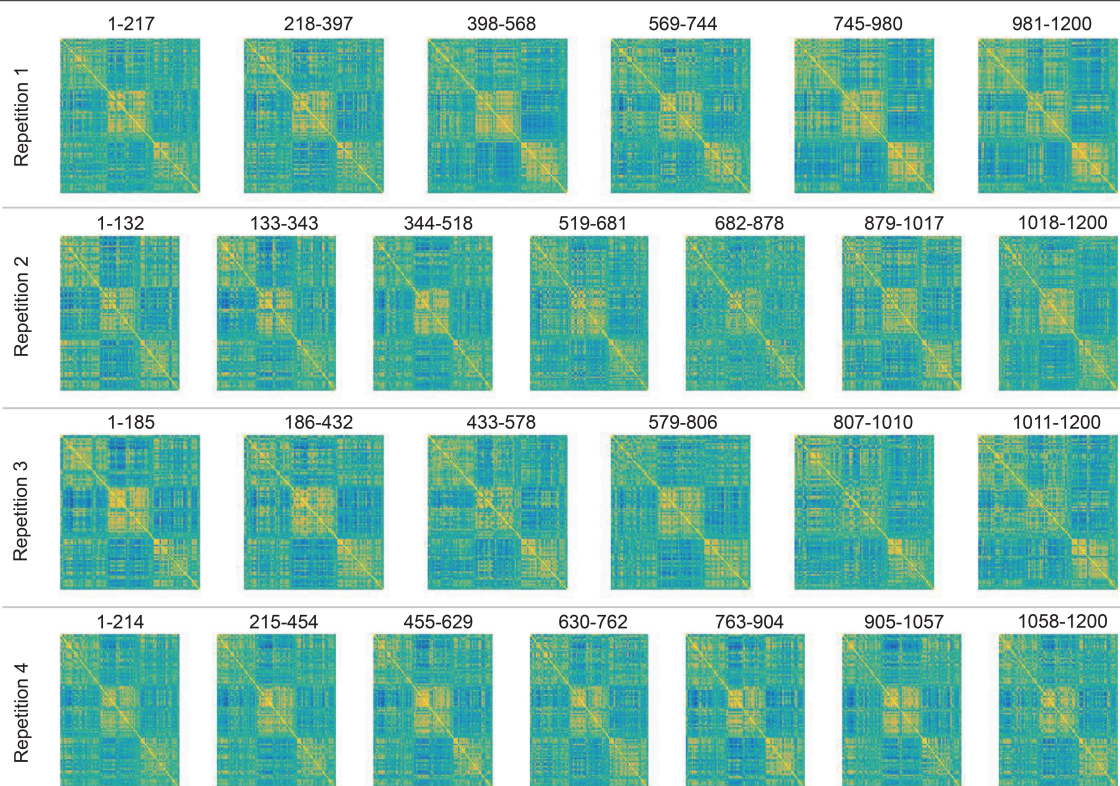


FIGURE 20 | Results of the analysis of the HCP data using DCD. Each row depicts the estimated correlation matrices for the time partitions detected by DCD for each of 4 runs for subject 100307. Above each matrix is the temporal information for each time partition.

between runs. Here one would not expect the timing of the change points to be similar across runs, as there is no explicit task designed to invoke state changes. Rather, this example is primarily meant to illustrate that DCD is able to detect change points in situations where there are 100 nodes.

That said, these results are consistent with results seen in previous literature (Allen et al., 2014), and suggest that dynamic behavior of functional connectivity is present in the resting state data. In particular states appear to be differentiated by connectivity between default mode components, and between default mode components and other components throughout the brain.

5. Discussion

In this work, we have developed a novel algorithm for change point detection in fMRI data. It partitions the fMRI time series into sequences based upon functional connectivity changes between ROIs or voxels, as well as mean activation changes. DCD can be applied to time series data from ROI studies, or to temporal components obtained from either a principal components or independent components analysis. Its data-driven design means it does not require any prior knowledge of the nature of the experiment. In addition, the accuracy of the result on single subject data allows for analysis on experiments where one expects large heterogeneity in connectivity across

subjects and between runs, such as in resting state fMRI data.

To reduce the burden on users, all three input parameters to the DCD algorithm have a clear statistical interpretation, making it easy to use even for those unfamiliar with the intrinsic details of the algorithm. As long as the user has a basic understanding of hypothesis testing, they should have the appropriate knowledge necessary to alter the parameters in order to improve the performance of the algorithm.

We contrast the approach to the previously introduced DCR technique, which also seeks to find connectivity change points. The most significant advantage of DCD compared to DCR is its computational efficiency, driven in large part by the newly proposed adaptive thresholding schema for sparse covariance matrix estimation. Based on the results of two high-dimensional simulation studies, as well as further empirical studies, we found that the computation time for DCR grows rapidly with an increased number of ROIs. Thus, when the number of regions exceeds 50, the computational burden of DCR can be intimidating for most users. In contrast, the computation time of DCD increases roughly linearly, and can easily handle hundreds of ROIs, in a matter of minutes for most general fMRI experimental settings.

In the DCD algorithm, we choose to maximize the total likelihood function instead of the Bayesian information criterion (BIC) that is used in the DCR algorithm. The design of the DCD

algorithm frees the user from performing model selection from a list of regularization parameters, so that we can use the likelihood function as a more natural criterion. Furthermore, utilizing the likelihood function avoids a common problem arising when applying the BIC; namely that when adding the BIC score of two subsets of lengths n_1 and n_2 ($n_1 + n_2 = n$), consisting of roughly the same number of parameters $k_1 \approx k_2 \approx k$, the total penalty term is $k \log(n_1) + k \log(n_2) \propto \log(n_1(n - n_1))$, which tends to favor small or large n_1 when minimizing the BIC. This is the reason for the apparent cluster of false positives obtained using DCR at time points Δ and $T - \Delta$, shown in **Figure 8**.

Another critical difference between the two algorithms is the manner in which sparsity is enforced. DCR uses GLASSO, and thus places sparsity constraints on the precision matrix, while DCDs adaptive thresholding approach places them on the covariance matrix. The former may be more natural in the fMRI setting, due to the relationship between the precision matrix and the connectivity graph where zero elements correspond to conditional independence. However, we found in our simulation studies that when estimating connectivity change points it does not appear to be critical upon which matrix we impose sparsity, and the computational advantages of operating on the covariance matrix becomes increasingly attractive. However, in settings where the precision matrix is sparse, and the corresponding covariance matrix is dense, DCD can potentially run into problems and alternative approaches should be explored.

One limitation preventing us from further improving the runtime of the DCD algorithm comes from the nature of greedy method we used for maximizing the likelihood. The greedy search strategy makes the locally optimal choice at each step, but cannot ensure the global optimum solution is obtained. However, as a data-driven method, the results from DCD will still provide a reasonable starting point for exploring the experimental data. Another disadvantage of DCD are limits on the types of experiments it may be applied to. In this work, we have demonstrated its effectiveness using both blocked-design

task fMRI experiments as well as resting state data. However, for event-related designs, the brain connectivity and activity level may change too rapidly to be able to obtain a valid estimate from DCD. Hence, when the DCD algorithm detects no significant change points, it may in fact be the case that the activity pattern changes too frequently to be detected.

Similar to group-level DCR, there is also a simple variant of DCD for group inference, which stacks subjects and calculates the summation of the likelihood function in each step. This approach can be used in experiments where one expects subjects to change states at similar time points (e.g., in the social evaluative threat experiment), and is not recommended for resting-state experiments where subjects are not expected to behave in a similar manner. In general, we suggest one first performs single-subject DCD, and if the resulting change points show synchronization across a subset of subjects, then apply group-level DCD to obtain more accurate results. Due to the flexibility of the DCD algorithm, we can also incorporate the GLASSO technique for sparse precision matrix estimation in place of adaptive thresholding method, which may also lead to improved accuracy at the cost of slower runtime.

In sum, the newly proposed DCD algorithm is a fast and efficient approach toward detecting changes in functional connectivity, especially for experiments where the nature, timing or duration of the involved psychological processes are unknown.

Acknowledgments

This research was partially supported by NIH grant R01EB016061. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

- Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676. doi: 10.1093/cercor/bhs352
- Beckmann, C. F., and Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *Med. Imaging IEEE Trans.* 23, 137–152. doi: 10.1109/TMI.2003.822821
- Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* 2008:P10008. doi: 10.1088/1742-5468/2008/10/P10008
- Cai, T., and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *J. Am. Stat. Assoc.* 106, 672–684. doi: 10.1198/jasa.2011.tm10560
- Chang, C., and Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fmri. *Neuroimage* 50, 81–98. doi: 10.1016/j.neuroimage.2009.12.011
- Cribben, I., Haraldsdottir, R., Atlas, L. Y., Wager, T. D., and Lindquist, M. A. (2012). Dynamic connectivity regression: determining state-related changes in brain connectivity. *Neuroimage* 61, 907–920. doi: 10.1016/j.neuroimage.2012.03.070
- Cribben, I., Wager, T. D., and Lindquist, M. A. (2013). Detecting functional connectivity change points for single-subject fmri data. *Front. Comput. Neurosci.* 7:143. doi: 10.3389/fncom.2013.00143
- Filippini, N., MacIntosh, B. J., Hough, M. G., Goodwin, G. M., Frisoni, G. B., Smith, S. M., et al. (2009). Distinct patterns of brain activity in young carriers of the apoe-ε4 allele. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7209–7214. doi: 10.1073/pnas.0811879106
- Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441. doi: 10.1093/biostatistics/kxm045
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6, 218–229. doi: 10.1006/nimg.1997.0291
- Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Hum. Brain Mapp.* 2, 56–78. doi: 10.1002/hbm.460020107
- Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008
- Ginestet, C. E., and Simmons, A. (2011). Statistical parametric network analysis of functional connectivity dynamics during a working

- memory task. *Neuroimage* 55, 688–704. doi: 10.1016/j.neuroimage.2010.11.030
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the human connectome project. *Neuroimage* 80, 105–124. doi: 10.1016/j.neuroimage.2013.04.127
- Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., et al. (2014). Ica-based artefact removal and accelerated fmri acquisition for improved resting state network imaging. *Neuroimage* 95, 232–247. doi: 10.1016/j.neuroimage.2014.03.034
- Handwerker, D. A., Roopchansingh, V., Gonzalez-Castillo, J., and Bandettini, P. A. (2012). Periodic changes in fmri connectivity. *Neuroimage* 63, 1712–1719. doi: 10.1016/j.neuroimage.2012.06.078
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079
- Jeong, J., Gore, J. C., and Peterson, B. S. (2001). Mutual information analysis of the eeg in patients with alzheimer's disease. *Clin. Neurophysiol.* 112, 827–835. doi: 10.1016/S1388-2457(01)00513-2
- Lindquist, M. A. (2008). The statistical analysis of fmri data. *Stat. Sci.* 23, 439–464. doi: 10.1214/09-STS282
- Lindquist, M. A., and Mejia, A. (2015). Zen and the art of multiple comparisons. *Psychosom. Med.* 77, 114–125. doi: 10.1097/PSY.0000000000000148
- Lindquist, M. A., Waugh, C., and Wager, T. D. (2007). Modeling state-related fmri activity using change-point theory. *Neuroimage* 35, 1125–1141. doi: 10.1016/j.neuroimage.2007.01.004
- Lindquist, M. A., Xu, Y., Nebel, M. B., and Caffo, B. S. (2014). Evaluating dynamic bivariate correlations in resting-state fmri: a comparison study and a new approach. *Neuroimage* 101, 531–546. doi: 10.1016/j.neuroimage.2014.06.052
- Robinson, L. F., Wager, T. D., and Lindquist, M. A. (2010). Change point estimation in multi-subject fmri studies. *Neuroimage* 49, 1581–1592. doi: 10.1016/j.neuroimage.2009.08.061
- Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., and Smith, S. M. (2014). Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* 90, 449–468. doi: 10.1016/j.neuroimage.2013.11.046
- Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput. Biol.* 1:e42. doi: 10.1371/journal.pcbi.0010042
- Sun, F. T., Miller, L. M., and D'Esposito, M. (2004). Measuring interregional functional connectivity using coherence and partial coherence analyses of fmri data. *Neuroimage* 21, 647–658. doi: 10.1016/j.neuroimage.2003.09.056
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* 58, 267–288.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., et al. (2013). The wu-minn human connectome project: an overview. *Neuroimage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041
- Varoquaux, G., Gramfort, A., Poline, J.-B., and Thirion, B. (2010). “Brain covariance selection: better individual functional connectivity models using population prior,” in *Advances in Neural Information Processing Systems*, eds J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta (Vancouver, BC), 2334–2342.
- Wager, T. D., Waugh, C. E., Lindquist, M., Noll, D. C., Fredrickson, B. L., and Taylor, S. F. (2009). Brain mediators of cardiovascular responses to social threat: part i: reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity. *Neuroimage* 47, 821–835. doi: 10.1016/j.neuroimage.2009.05.043

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Xu and Lindquist. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Appendix

A. Minimum Partition Length

We need to calculate a minimum partition length Δ to control the type II error based on a pre-specified bound β . Consider two time series, each of length Δ . Denote the test statistic $t_{stat} = \frac{\delta x}{\sqrt{\frac{2}{\Delta}s}}$, where s represents the pooled variance. Under

the null hypothesis, t_{stat} follows a Student's t-distribution with $2\Delta - 2$ degrees of freedom, and we reject H_0 if $|t_{stat}| \geq t_{1-\alpha/2}(2\Delta - 2)$.

If the alternative hypothesis H_1 is true, and the actual difference in mean between the two groups is $\delta\mu$, then the statistic $t' = \frac{\delta x - \delta\mu}{\sqrt{\frac{2}{\Delta}s}}$ follows a Student's t-distribution with

$2\Delta - 2$ degrees of freedom. Without loss of generality, assume that $\delta\mu > 0$. Then the type II error of this hypothesis test satisfies:

$$Pr(|\delta x|) \leq s \cdot t_{1-\alpha/2}(2\Delta - 2) \approx Pr(t' \leq t_{1-\alpha/2}(2\Delta - 2) - \frac{\delta\mu}{\sqrt{\frac{2}{\Delta}s}}) \leq \beta$$

In practice, we set the effect size as $\frac{\delta\mu}{s} = 1$ and since we are comparing time courses from J regions we use Bonferroni correction to set $\alpha \rightarrow \alpha/J$ and $\beta \rightarrow \beta/J$. Beginning at $\Delta = 10$, if $Pr(t' \leq t_{1-\alpha/2M}(2\Delta - 2) - \sqrt{\frac{\Delta}{2}})$ is larger than β , increase Δ by 1 until the above equation is satisfied.

B. Jackknife Resampling

The jackknife is a useful technique for variance estimation. It “bootstraps” the estimator by systematically leaving out each observation and re-calculating the estimate. Suppose we have a sequence of data $\{X_t\}_{1 \leq t \leq T}$, and we want to estimate the variance of an estimator:

$$\hat{\theta} = \frac{1}{T} \sum_t X_t$$

First we calculate the jackknife estimate of θ as

$$\theta_{jack} = \frac{1}{T} \sum_t \tilde{\theta}_t$$

where $\tilde{\theta}_t$ is the estimator for a subsample omitting the t^{th} observation,

$$\tilde{\theta}_t = \frac{1}{T-1} \sum_{s \neq t} X_s$$

Hence,

$$\begin{aligned} \theta_{jack} &= \frac{1}{T} \sum_t \frac{1}{T-1} \left(\sum_{s=1}^T X_s - X_t \right) \\ &= \frac{1}{T-1} \sum_t \left(\frac{1}{T} \sum_{s=1}^T X_s \right) - \frac{1}{T(T-1)} \sum_t X_t \\ &= \frac{T}{T-1} \hat{\theta} - \frac{1}{T-1} \hat{\theta} = \hat{\theta} \end{aligned} \quad (A1)$$

Now calculate an estimate of the variance of $\hat{\theta}$ using the jackknife technique:

$$\begin{aligned} Var(\hat{\theta}) &= \frac{T-1}{T} \sum_t (\tilde{\theta} - \theta_{jack})^2 \\ &= \frac{T-1}{T} \sum_t \left(\frac{1}{T-1} \left(\sum_{s=1}^T X_s - X_t \right) - \hat{\theta} \right)^2 \\ &= \frac{T-1}{T} \sum_t \left(\frac{T}{T-1} \hat{\theta} - \frac{1}{T-1} X_t - \hat{\theta} \right)^2 \\ &= \frac{T-1}{T} \sum_t \frac{1}{(T-1)^2} (X_t - \hat{\theta})^2 \\ &= \frac{1}{(T-1)T} \sum_{t=1}^T (X_t - \hat{\theta})^2 \end{aligned} \quad (A2)$$

Applying the result to Equation (1), we can estimate the variance of $\hat{\Sigma}_{ij}$ as

$$Var(\hat{\Sigma}_{ij}) = \frac{1}{(T-1)T} \sum_{1 \leq t \leq T} (X_t^{ij} - \hat{\Sigma}_{ij})^2 \approx \frac{1}{T} \delta_{ij}^2$$

which is similar to that obtained using the central limit theorem.

C. Simulation Setting

Below is a more detailed list of simulation studies, including the exact value of precision matrices used in simulation 2–6.

• Simulation 1

Description: The data is white noise with no connectivity change points.

Size: $N = 20$, $T = 1000$, $p = 20$

• Simulation 2

Description: There are two change points at times 200 and 400. Spikes are imposed onto the time series, imitating a common artifact found in fMRI data. For each subject there are 5 randomly placed spikes, each with magnitude 15.

Size: $N = 20$, $T = 1000$, $p = 20$

Dependency Structure:

$$\begin{aligned} t \in [1, 200]: & \quad (3, 14) = 0.3, (3, 9) = 0.6, (9, 14) = 0.4 \\ t \in (200, 400]: & \quad (1, 6) = 0.7, (6, 14) = 0.5, (1, 19) = 0.6 \\ t \in (400, 600]: & \quad (3, 10) = 0.7, (3, 13) = 0.6, (3, 20) = 0.4, \\ & \quad (10, 20) = 0.1, (13, 20) = 0.1 \end{aligned}$$

• Simulation 3

Description: There are three change points at times 125, 500, and 750.

Size: $N = 15$, $T = 1000$, $p = 20$

Dependency Structure:

$$\begin{aligned} t \in [1, 125]: & \quad (2, 8) = 0.7, (8, 17) = 0.5, (2, 17) = 0.2 \\ t \in (125, 500]: & \quad (6, 14) = 0.1, (1, 6) = 0.3, (1, 18) = 0.2, \\ & \quad (1, 14) = 0.3, (6, 18) = 0.4 \\ t \in (500, 750]: & \quad (3, 8) = 0.5, (8, 13) = 0.5, (13, 19) = \\ & \quad 0.4, (3, 19) = 0.4, (3, 13) = 0.1, (8, 19) = 0.2 \\ t \in (750, 1000]: & \quad (5, 11) = 0.8 \end{aligned}$$

• Simulation 4

Description: There is a single change point at time 100.

Size: $N = 25$, $T = 200$, $p = 5$

Dependency Structure:

$$\begin{aligned} t \in [1, 100]: & \quad (1, 3) = 0.7, (3, 5) = 0.6, (1, 5) = 0.3, (3, 4) \\ & \quad = 0.2, (4, 5) = 0.2, (1, 4) = 0.1 \\ t \in (100, 200]: & \quad (1, 2) = -0.1, (1, 5) = -0.2, (2, 5) = 0.4 \end{aligned}$$

• Simulation 5

Description: There are five change points at times 200, 300, 500, 600, and 800.

Size: $N = 20$, $T = 1000$, $p = 20$

Dependency Structure:

$$\begin{aligned} t \in [1, 200]: & \quad (2, 14) = 0.8 \\ t \in (200, 300]: & \quad (2, 14) = 0.4, (3, 9) = 0.3, (9, 18) = 0.4, \\ & \quad (3, 18) = 0.3 \\ t \in (300, 500]: & \quad (3, 9) = 0.7, (3, 18) = 0.5, (9, 18) = 0.3 \\ t \in (500, 600]: & \quad (2, 19) = 0.4, (3, 18) = 0.3, (2, 13) = \\ & \quad 0.5, (6, 13) = 0.2, (9, 18) = 0.3 \\ t \in (600, 800]: & \quad (2, 6) = 0.6, (6, 19) = 0.5, (2, 19) = 0.3, \\ & \quad (2, 13) = 0.5 \\ t \in (800, 1000]: & \quad (1, 11) = 0.9 \end{aligned}$$

• Simulation 6

Description: There are four change points at times 200, 400, 600, and 800.

Size: $N = 20$, $T = 1000$, $p = 20$

Dependency Structure:

$$\begin{aligned} t \in [1, 200]: & \quad (1, 5) = 0.8, (5, 10) = 0.3, (10, 15) = 0.5 \\ t \in (200, 400]: & \quad (2, 9) = 0.6, (9, 18) = 0.3 \\ t \in (400, 600]: & \quad (3, 6) = 0.4, (6, 13) = 0.3, (13, 19) = 0.2 \\ t \in (600, 800]: & \quad (4, 8) = 0.7, (8, 15) = 0.3, (15, 20) = 0.6 \\ t \in (800, 1000]: & \quad (2, 14) = 0.5 \end{aligned}$$



Big Data Approaches for the Analysis of Large-Scale fMRI Data Using Apache Spark and GPU Processing: A Demonstration on Resting-State fMRI Data from the Human Connectome Project

Roland N. Boubela^{1,2†}, Klaudius Kalcher^{1,2†}, Wolfgang Huf^{1,2}, Christian Našel³ and Ewald Moser^{1,2,4*}

¹ Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria, ² MR Centre of Excellence, Medical University of Vienna, Vienna, Austria, ³ Department of Radiology, Tulln Hospital, Karl Landsteiner University of Health Sciences, Tulln, Austria, ⁴ Brain Behaviour Laboratory, Department of Psychiatry, University of Pennsylvania Medical Center, Philadelphia, PA, USA

OPEN ACCESS

Edited by:

Brian Caffo,
Johns Hopkins University, USA

Reviewed by:

Xi-Nian Zuo,
Chinese Academy of Sciences, China
Xin Di,
New Jersey Institute of Technology,
USA

*Correspondence:

Ewald Moser
ewald.moser@meduniwien.ac.at

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 15 August 2015

Accepted: 10 December 2015

Published: 06 January 2016

Citation:

Boubela RN, Kalcher K, Huf W,
Našel C and Moser E (2016) Big Data
Approaches for the Analysis of
Large-Scale fMRI Data Using Apache
Spark and GPU Processing: A
Demonstration on Resting-State fMRI
Data from the Human Connectome
Project. *Front. Neurosci.* 9:492.
doi: 10.3389/fnins.2015.00492

Technologies for scalable analysis of very large datasets have emerged in the domain of internet computing, but are still rarely used in neuroimaging despite the existence of data and research questions in need of efficient computation tools especially in fMRI. In this work, we present software tools for the application of Apache Spark and Graphics Processing Units (GPUs) to neuroimaging datasets, in particular providing distributed file input for 4D NIfTI fMRI datasets in Scala for use in an Apache Spark environment. Examples for using this Big Data platform in graph analysis of fMRI datasets are shown to illustrate how processing pipelines employing it can be developed. With more tools for the convenient integration of neuroimaging file formats and typical processing steps, big data technologies could find wider endorsement in the community, leading to a range of potentially useful applications especially in view of the current collaborative creation of a wealth of large data repositories including thousands of individual fMRI datasets.

Keywords: fMRI, big data analytics, distributed computing, graph analysis, Apache Spark, scalable architecture, machine learning, statistical computing

1. INTRODUCTION

The pressure to continuously analyze fast growing datasets has led internet companies to engage in the development of specialized tools for this new field of Big Data analysis, at first strongly focused on the specific data structures used by their applications, but increasingly taking more generalized forms. One of the most fundamental developments in this area is Google's MapReduce paradigm (Dean and Ghemawat, 2004), designed for efficient distributed computations on datasets too large to fit on a single machine, which are instead stored in a distributed file system in a cluster environment. The computation concept behind MapReduce is to use the individual cluster nodes where the data are stored as efficiently as possible by transferring as much of the computation as possible to the individual storage nodes instead of transferring their data to a designated compute node, and only perform subsequent aggregation steps of the computation to master compute

nodes. Thus, there exists a strong link between the distributed data storage and the computation. For example, Apache's open source implementation of the paradigm consists of Hadoop, the implementation of the actual MapReduce computation engine, and the Hadoop Distributed File System (HDFS) for data storage. The Hadoop ecosystem is further complemented by a variety of toolkits for specialized applications like machine learning.

The principles of the MapReduce paradigm can best be illustrated using the distributed algorithm for counting the number of occurrences of words in large documents, the canonical example for MapReduce computations. As the name suggests, these computations consist of two steps, termed Map and Reduce, with Map being performed on each node separately, and the Reduce step computed on a central node, aggregating the individual Map results. In the word count example, the Map step would consist in generating, for each part of the document stored on the distributed file system, a set of keys and values, with words being the keys and the number of occurrences of each word being the associated value. The Reduce step would then aggregate these partial results by building the sum of all values from all individual nodes associated with each word, thus gaining the overall number of occurrences of this word in the entirety of the dataset.

While the approach proves to be flexible enough for a wide range of computations, this brief description should also make it apparent that not all kinds of computations can be performed in this way. For example, many data analysis applications, like iterative machine learning algorithms, need to access data multiple times, which would be very inefficient if implemented in pure MapReduce terms. Addressing this issue and providing a more general framework for distributed computations on large datasets was the main motivation behind the introduction of the Spark framework (Zaharia et al., 2012; The Apache Software Foundation, 2015). The counterpart of data stored in the Hadoop distributed filesystem in the Spark framework are so-called resilient distributed datasets (RDD), which, unlike files in the HDFS, can be held entirely in memory if space allows (and cached to disk if memory is not sufficient), and provide a high-level interface for the programmer. The details of the distributed storage and computation on this distributed dataset are thus abstracted, making the writing of distributed code much easier in practice. Furthermore, Spark encompasses higher-level libraries for many applications including machine learning (MLlib) and graph analysis (GraphX), further facilitating the development of analyses in these specific domains. Spark can be used interactively from a Scala shell or via its Application Programming Interface (API), with APIs existing for Scala, Java, python and, most recently, R. With Spark being written in Scala and the interactive shell being a Scala shell, the connection between Spark and Scala is the strongest, and the other languages' APIs do not yet have the full functionality of the Scala API; for example, there is no interface to many functions of GraphX in python, and the R API is currently only in an early stage of development.

A further approach to accelerating computations on large datasets by parallelization, though not directly related to the Big Data technologies in the stricter sense mentioned above, concerns optimization of computations on a single machine, where in particular the use of Graphics Processing Units (GPUs)

can make an enormous difference in terms of computational efficiency and thus rendering possible the analysis of even larger datasets in a reasonable amount of time.

Both the big data frameworks and GPU acceleration can prove useful in the field of neuroimaging in general and functional MRI in particular, where increasing spatial and temporal resolutions as well as larger sample sizes lead to a rapid increase in the amount of data that needs to be processed in a typical study. GPU computing has been embraced not only to provide faster programs for standard algorithms (Eklund et al., 2014), but also to make some more complex analyses possible at all (Boubela et al., 2012; Eklund et al., 2012, 2013). Apart from such special tools, GPU acceleration has in some cases already be harnessed in standard neuroimaging data analysis libraries like, for example, in FSL (Jenkinson et al., 2012). In contrast to Big Data technologies in the narrower sense, however, these technologies do not scale arbitrarily, but are instead limited to the amount of data that can be held in memory on a single machine. But while GPUs have slowly been picked up by the neuroimaging community, the spread of Hadoop and Spark is more limited. In the context of the Human Connectome Project, Marcus et al. (2011) describe the infrastructure for the storage and exploration of such a large dataset, but do not employ big data tools for efficient analyses on the whole dataset of 1400 subjects. Only two published papers have yet used them in the field of neuroimaging: Wang et al. (2013) used Hadoop to use random forests for machine learning on a large imaging genetics dataset, and Freeman et al. (2014) provide an analysis framework based on Apache Spark and highlight applications for two-photon imaging and light-sheet imaging data.

The dearth in this domain is all the more surprising in view of the emergence of a number of data sharing initiatives and large-scale data acquisition projects covering a wide array of topics in human neuroimaging (Biswal et al., 2010; ADHD-200 Consortium, 2012; Nooner et al., 2012; Assaf et al., 2013; Jiang, 2013; Mennes et al., 2013; Van Essen et al., 2013; Satterthwaite et al., 2016). Certainly, the opportunities offered by the availability of neuroimaging data from a large number of subjects are coming with some challenges (Akil et al., 2011). As has previously been noted, the sheer size of the datasets and their complexity require new approaches to harvest the full benefit of "human neuroimaging as a big data science" (Van Horn and Toga, 2014). For example, Zuo et al. (2012), when computing network centrality measures at a voxel-wise level, resampled all datasets to a 4 mm (isotropic) resolution and stated two reasons for this choice. The first reason is the average resolution of the datasets available from the 1000 Functional Connectomes dataset in the largest voxel dimension, which was not much below 4 mm, leading to the conclusion that using a higher resolution might not be worth the effort on this dataset. The second stated reason was the computational demands that a higher resolution would require: while the voxelwise network at a 4 mm resolution had 22,387 nodes, this number would increase to 42,835 when using a 3 mm resolution. Since then, even higher resolutions than 3 mm have become more and more common—the Human Connectome Project dataset for example uses isotropic 2 mm

voxels—and the need to address the computational demands that accompany this increase in data size becomes obvious.

Still, while large-scale data repositories could provide a good model on how to use big data technologies in human neuroimaging, they have not yet been explored with these methods. One reason for the neuroimaging community not embracing big data tools more readily might be the lack of reasonably efficient I/O from (and, to a lesser extent, to) standard neuroimaging file formats like NIfTI. Removing this barrier of entry might open the way to a variety of analysis tools that could then be directly applied to datasets of practically arbitrary size. While the range of tools that can currently be applied to large datasets is limited to computationally relatively simple methods like regression, scaling the computation power using clusters can extend this to more complex machine learning and graph mining algorithms, including methods without closed form solution that need to be solved iteratively. Another research area where computationally intensive methods might prove useful is the investigation of reliability and reproducibility of neuroimaging methods as reviewed by Zuo and Xing (2014), who also note that easing the computational demand by aggregation, e.g., averaging the signal from multiple voxels based on anatomical structure, leads to difficulties in the reliability and interpretation of derived results, and strongly encourage voxel-wise analysis for the evaluation of the functional architecture of the brain. The Consortium for Reliability and Reproducibility in particular has gathered a large dataset of over 5000 resting-state fMRI measurements to this end (Zuo et al., 2014), and proposes a number of computational tools for use on this database, yet these do not currently include big data tools.

2. TESTING PLATFORM AND DATA

2.1. Computing Environment

The computation of the connectivity matrices based on Pearson correlation were performed on a server running Ubuntu Linux (version 12.04) equipped with 192 GB random access memory (RAM), two Intel Xeon X5690 processors and four Nvidia Tesla C2070 GPUs. As an aside, it should be noted that while these GPUs are somewhat dated, they already have full support for double precision computations; while modern GPUs no longer have issues with double precision computation, older generations (with compute capability <2.0, corresponding approximately to models developed before 2011) might be slow or unable to perform anything but single precision computations. The linear algebra operations on the GPUs were accessed using CUDA 6.0 and integrated in R (Boubela et al., 2012; R Core Team, 2014). Spark was used via the Scala shell and API for the practical reasons discussed above. OpenBLAS version 0.2.14 was compiled and installed for the Apache Spark compute nodes to enable these machine optimized libraries to be used by Spark's linear algebra functionality in MLlib. Further, R uses the OpenBLAS implementation of the singular value decomposition (SVD) for performance comparison purposes.

The cluster running Apache Spark 1.5.1 consists of ten Sun Fire X2270 servers using Ubuntu Linux (version 14.04) with

48 GB RAM and two Intel Xeon X5550 processors. Additionally, each server uses three 500 GB hard disk drives (HDD) as local disk space for the Apache Spark framework. Beside a standard 1 GBit ethernet connection, the cluster nodes are connected via the IP over Infiniband protocol on QDR Infiniband hardware.

2.2. Subjects

To test the methods described, 200 sample datasets from the Human Connectome Project (Van Essen et al., 2013) were downloaded from the project repository and used for example analyses. In this study, only the resting-state fMRI data were used, though the methods are not limited to this type of data.

2.3. Source Code

All code presented in this work can be found in the github repository <https://github.com/rboubela/biananes>.

3. HUMAN CONNECTOME PROJECT DATA ANALYSIS

3.1. NIfTI File Input for fMRI

One of the most basic obstacles to using Apache Spark for fMRI datasets is the lack of an efficient file input function able to process any file formats usually used in this field like NIfTI-1. Of course, file readers in Java, python or R exist which could be used when using Spark from their respective API, and the Java file reader could be used in Scala (and thus also in the Scala shell), but none of those file readers is suited for the distributed environment. For this, a distributed file reader for fMRI data was implemented in Scala and C which reads 4D NIfTI files in parallel on multiple nodes, with each node reading a different set of the image's volumes, and gathers the results into an RDD in the Spark environment. To avoid unnecessary overhead, a brain mask can be used to restrict reading to in-brain voxels; the brain mask must also be available as a NIfTI file and will be applied to all volumes of the 4D NIfTI file to be read. Files can be read from local harddisks on the nodes or via the network file system (NFS) protocol from a centralized storage accessible to the compute nodes (see **Figure 1**). While in principle, the former method is faster than reading the files over the network, reading the input data is rarely the computational bottleneck in fMRI data analysis, and thus reading the input files even from the same common network storage device is efficient enough while typically being much more convenient. Nonetheless, for situations where fast file access over the network is not available, or if local storage is preferred for other reasons, the reader also allows for reading NIfTI input from local harddisks, in which case the NIfTI input file(s) must be available on all nodes under the same path.

In Spark, the voxelwise timeseries data is stored in the columns of a *RowMatrix* object. This type of object is the most commonly used in the interface of the Apache Spark machine learning library *MLlib* for the distributed handling of large numerical matrices. For example, SVD or principal component analysis (PCA) can be applied directly on this *RowMatrix*, which in turn can be the basis for various further statistical analyses like independent component analysis (ICA). Column similarities

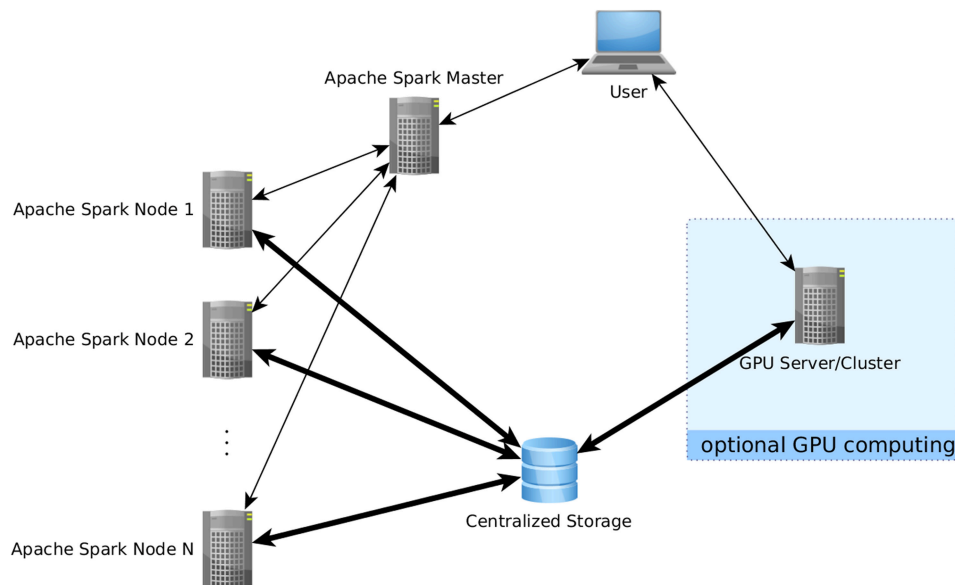


FIGURE 1 | Data flow using the proposed analysis methods. Bold arrows represent intensive data flow, the other arrows represent communication of control commands. In this example, parts of the computations have been performed on a separate GPU computing server which was not part of the Apache Spark cluster. The use of a centralized data storage facilitates the integration of all steps into a comprehensive pipeline, as the fMRI data is loaded directly from there onto the GPU server, who then writes the results as edgelist back on the storage to be directly readable by Apache Spark. Note that storing data directly on the compute nodes is also possible as an alternative if issues related to data transfer speed are encountered.

based on the cosine similarity can also be computed efficiently on a *RowMatrix* in Spark (Zadeh and Goel, 2012; Zadeh and Carlsson, 2013). Examples for using the data input function are shown in code listings 1 and 2 for single-subject and group data import, respectively, and the runtime of the NIfTI reader is shown in **Figure 2**. To exemplify possible linear algebra computations, a call for the SVD computation from *MLlib* is shown at the end of code listing 2. It should be noted that while this toy example demonstrates that using *MLlib* functions is very straightforward and easy, it would not make much sense from a computational point of view in this particular case: on four nodes, the computation of 10 singular values and vectors takes 604 s, and the computation of 100 singular values and vectors takes 2700 s; the same values can be computed on a single one of those nodes using *svd* in R with OpenBLAS as linear algebra backend in 118 and 126 s, respectively. Using Spark for linear algebra computations seems only sensible if the size of input data precludes the use of standard optimized linear algebra packages like OpenBLAS. The examples that follow will thus focus on more data-intensive problems like graph mining, where even single-subject analysis can involve the handling of very large datasets.

Listing 1: Reading a single-subject fMRI dataset

```
import org.biananes.io.NiftiTools

val hcp_root = sys.env("HCP_ROOT")
val func = "/MNINonLinear/Results/rfMRI_REST1_RL/
rfMRI_REST1_RL.nii.gz"
val mask = "/MNINonLinear/Results/rfMRI_REST1_RL/
brainmask_fs.2.nii.gz"
```

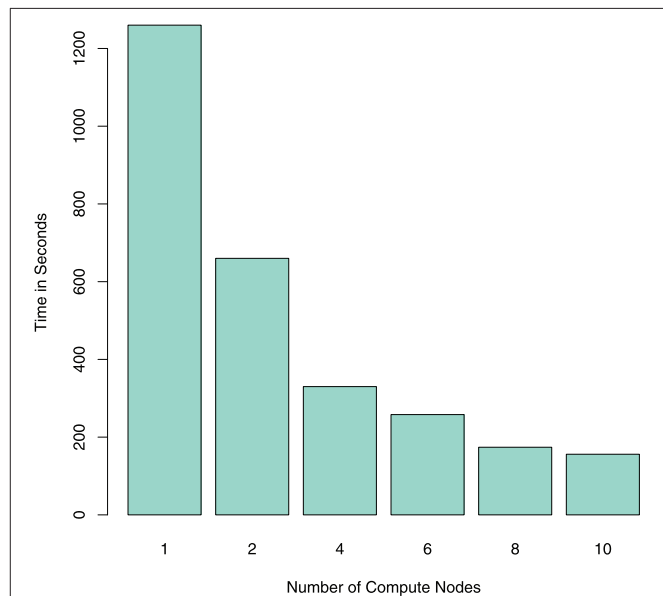


FIGURE 2 | Comparing the runtime (in seconds) for reading one resting-state fMRI dataset using *NiftiImageReadMasked* on an Apache Spark cluster with different numbers of compute nodes. Note that the reduction in computation time scales almost with $1/n$, n being the number of nodes in the cluster.

```
val template = "/usr/share/data/fsl-mni152-templates/
MNI152lin_T1_2mm.nii.gz"

val img_file = hcp_root + "167743" + func_file
```

```
val mask_file = hcp_root + "167743" + mask

val mat = NiftiImageReadMasked(img_file, mask_file, sc
)
```

Listing 2: Reading a group of subjects storing the data in one big group matrix and compute SVD on this matrix

```
val subjects = sc.textFile(hcp_root + "subjectIDs.txt"
)

val input_files = subjects.map{ subject =>
  new Tuple2(new String(hcp_root + subject + func
    ,
    template) ).collect

val group_matrix = input_files.map{
  f => NiftiImageReadMasked(f._1, f._2, sc) }.reduce((
    m1, m2) => new RowMatrix(m1.rows.union(m2.rows))
)

val svd_result = group_matrix.computeSVD(1000)
```

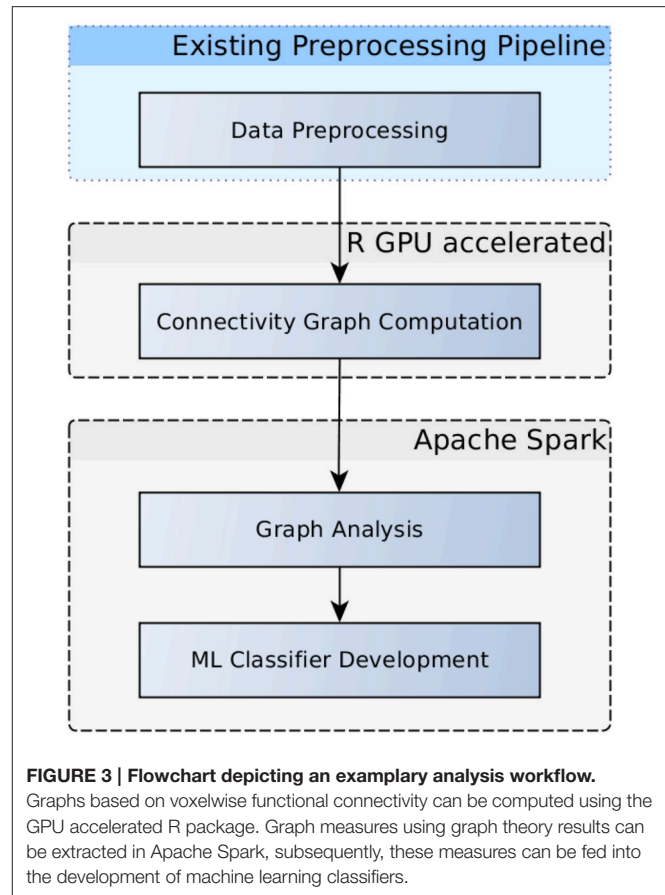
3.2. GPU Connectivity Matrix

A more common similarity measure that can be used to compare voxel time series is the Pearson correlation coefficient, which is often used as functional connectivity measure in fMRI. Beside visualization of these connectivity patterns themselves, this measure can also be used in further analyses including machine learning (Eickhoff et al., 2016) or graph analyses (Craddock et al., 2015; Kalcher et al., 2015), as illustrated in the workflow diagram in **Figure 3**. In contrast to the above mentioned cosine similarity, Pearson correlation coefficients are simple linear algebra computations that can be computed by the arithmetic units on GPUs in a highly parallelized way, making it a viable application for GPU acceleration. Larger matrices might exceed the memory available on a GPU, however, but this problem can be addressed by tiling the input matrices in a way to separately compute submatrices of the result and subsequently concatenating the parts to form the complete matrix. In the case of the Human Connectome Project data, the voxelwise correlation matrix in the original resolution of all in-brain voxels (228200 ± 2589 voxels) for one subject takes up ~ 390 GB, which is divided into 91 tiles of 4.2 GB each (the rest of the GPU RAM is used up by the input needed to compute the respective tile).

The resulting correlation/connectivity matrix can be thresholded to obtain an adjacency matrix for a graph, with different options being available for the choice of a correlation threshold. For the estimation of the runtime for multiple subjects as shown in **Figure 4**, the matrix was thresholded at absolute values of 0.6 of the correlation coefficients. Subsequently, these sparse matrices were saved to RData files for further usage. (Note that since different fMRI datasets can be rather heterogeneous, it is in general more advisable to use an automated selection of a correlation threshold to achieve a certain edge density in the graph, for example defined by the value of $S = \log E / \log K$, with E being the number of edges and K the average node degree.)

3.3. Graph Analysis in Apache Spark

The Apache Spark framework contains the *GraphX* library for the efficient development of distributed and scalable graph



algorithms. A graph object from this library can be constructed from a variety of different inputs, including cosine similarities computed from the *RowMatrix* object or by directly reading a comma separated value (CSV) file containing a list of edges. Graphs defined using this library are represented in the Spark environment as two RDDs, one containing the vertices and the other the edges, in order to allow for distributed computations on the graph. Code listing 3 shows an example of importing multiple graphs from individual subject graph edge lists, and computing and saving connected components in each of the graphs. The corresponding computation times are illustrated in **Figure 5**, and exemplary results from graph analyses are shown in **Figure 6**.

Listing 3: Reading connectivity graphs from text files; computing connected components and storing results on disk

```
import org.apache.spark.graphx._

val edgeListFiles = sc.textFile(hcp_root + "
  hcp_edgelist_files.txt").collect
val graphs = edgeListFiles.map { edgefile => new
  Tuple2(edgefile, GraphLoader.edgeListFile(sc,
    edgefile, false)) }

// compute the connected components for all graphs
val allConnectedComponents = graphs.map { g => new
  Tuple2(g._1, g._2.connectedComponents().vertices)
}

// saving the results
```

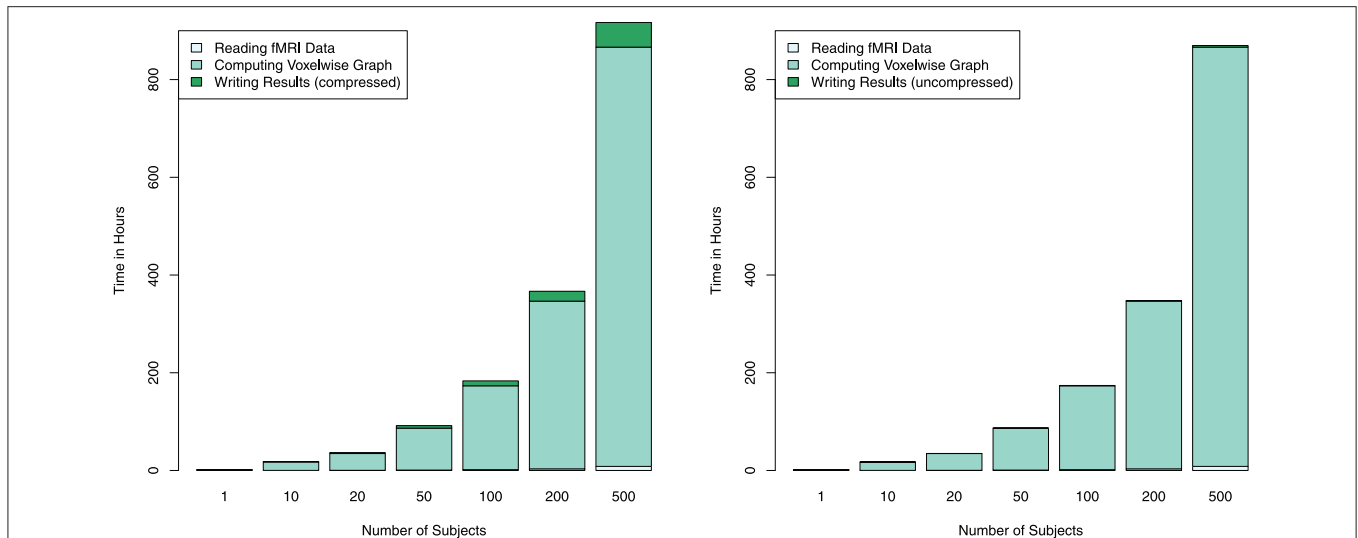


FIGURE 4 | Estimated combined computation times for reading the fMRI data, computing the connectivity graph and writing the thresholded (and thus sparse) connectivity matrix to compressed (left) and uncompressed (right) RData files are depicted for different numbers of subjects on a single GPU. Since computation time depends linearly on the number of subjects, computation time for larger numbers of subjects are estimated using average per-subject computing times measured from 200 subjects. By employing multiple GPUs the runtime can be reduced linearly; for example, using four GPUs instead of one for computing 500 subject's connectivity graph would reduce the computation time from about 36 days down to 9.

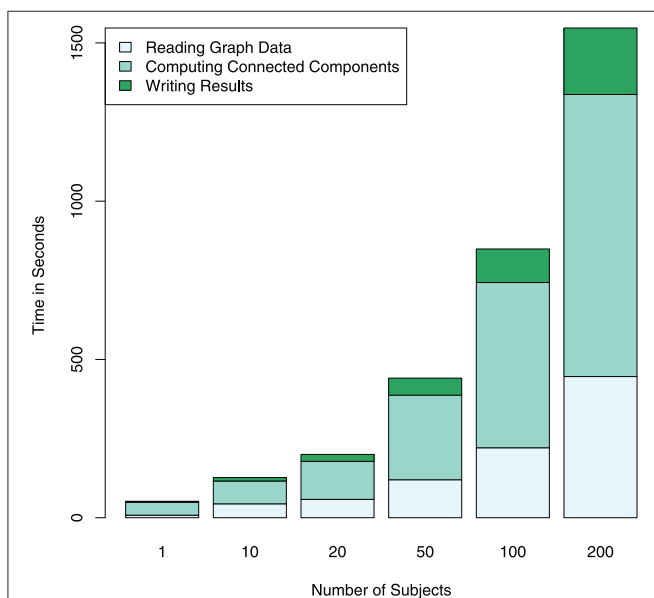


FIGURE 5 | Computation times for reading and writing the graph data in addition to computing connected components for a different number of subjects is shown performed on an Apache Spark cluster using four compute nodes. The largest part of the computation time is spent on the graph computations themselves. Note that the computational complexity of the search for connected components is relatively low ($O(n)$), in the case of more complex computations, the proportion of the total computation time spent with data I/O further decreases.

```
val resfiles = allConnectedComponents.map{ cc => {
  val file = cc._1.substring(0, 106) + "
    connected_components"
  cc._2.coalesce(1, true).saveAsTextFile(file)
}
```

```
file }
}
```

One of the main advantages of using *GraphX* for graph analyses in fMRI is that computations can be distributed very easily to allow for pooled analysis of large groups of subjects. The example in code listing 4 demonstrates this using the example of the computation of voxelwise local clustering coefficients for all single-subject graphs read in the previous code listing. Note how the parallelized computation for all subjects is achieved with only a single line of code, without the need for explicit commands for the parallelization.

Listing 4: Computing the local clustering coefficient for each voxel for all graphs

```
val allClusteringCoeff = graphs.map { g => new Tuple2(
  g._1, g._2.degrees.join(g._2.triangleCount.
    vertices).map{ case (id, (d, tc)) => (id, tc / (d
  * (d - 1) / 2.0))})
}
```

4. DISCUSSION

Big Data technologies are not yet often employed in the analysis of neuroimaging data, though the emergence of large collaborative repositories especially in the field of fMRI provides an ideal environment for their application. One of the main reasons for the currently limited interest in these technologies by researchers in neuroimaging seems to be a comparatively high effort for a first entry into this domain, in particular in view of the lack of interfaces to the data formats typically used in the field. Here, we present a distributed NIfTI file reader written in Scala for Apache Spark and show applications that become possible with this framework, including graph analyses using *GraphX*.

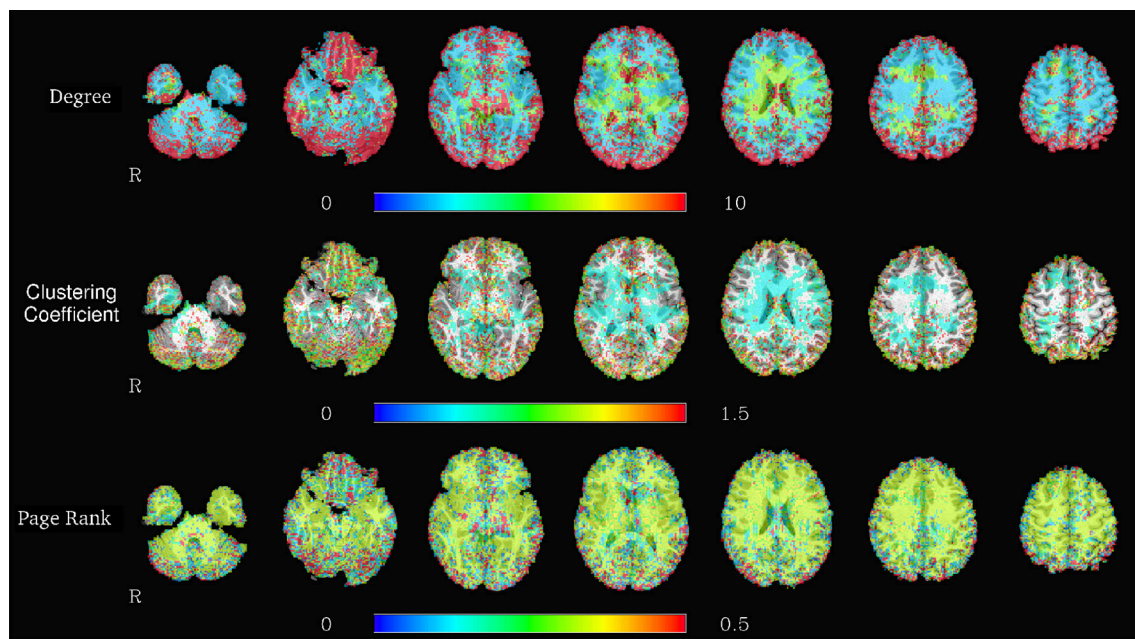


FIGURE 6 | Spatial distribution of node degrees (top), local clustering coefficients (center), and PageRank (bottom) at a voxelwise level for one representative subject, using the graph based on the correlation map thresholded at 0.6.

In addition, the computation of correlation matrices from fMRI time series was implemented to run on GPUs and optimized for the 4D structure of time series fMRI data.

Most Spark code was written in Scala, which is the preferred language for development in this framework at the moment. However, interfaces to different languages are available at various stages of maturity, including python and R, which are both commonly used for fMRI data analysis. Though using Spark via the API from one of those languages does not currently provide access to the full range of analysis tools available in the Scala API, adding wrappers for these languages into our package would be a valuable addition.

Transferring fMRI computations into a Big Data analysis framework like Spark offers the advantage of the direct availability of a range of tools optimized for particular problems. Two of the most notable applications here are machine learning and graph data analysis, provided by the the Spark libraries *MLlib* and *GraphX*, respectively. Both machine learning and graph analysis are rapidly growing subfields in the fMRI community (Bullmore and Sporns, 2009; Craddock et al., 2015), but the applications of these methods is often limited by the computational means available for tackling the comparatively complex calculations involved.

Apart from efficiency in the sense of computation speed, a second type of efficiency is just as important in practical research software development: efficiency in terms of development time. While parallelization tools are available in multiple programming languages at different levels, one of the advantages of Spark in this respect is the relative ease with which it allows for distributing computations in cluster environments even in an interactive shell. As shown in code listing 4, the details of the distribution of computations is hidden from the developer, allowing for easier programming compared to other tools requiring explicit

parallization. Furthermore, ease of access could be further improved by convergence with open data pipelines as developed in the context of data sharing initiatives (Zuo et al., 2014; Xu et al., 2015), as the inclusion of big data tools into published analysis pipelines could help spread such tools to a wider community of researchers who might otherwise not investigate these opportunities.

Another important aspect of using a scalable platform is the ability to avoid buying and operating on premise computing equipment, but instead move data analysis and computation tasks to cloud service providers. As Freeman et al. (2014) have shown in their work, using large amount of quickly available cloud computing resources can conveniently be leveraged using the Spark Framework. For example, in addition to running the Spark Framework, the Amazon web services (AWS) cloud (as used by Freeman et al., 2014) also provides compute nodes with GPUs (<https://aws.amazon.com/ec2/instance-types/>), and therefore, could also be employed for the GPU accelerated computation of connectivity graphs as proposed herein.

It is probably the difficulty of climbing the first steep learning curve that is responsible for the limited application of big data tools in neuroimaging research, with only two published papers so far, one using Hadoop (Wang et al., 2013), the other using Spark (Freeman et al., 2014). The more tools are published to make the first steps with these technologies easier, of which the distributed NIfTI file reader provides a starting point, the more researchers will be able to use these tools, thus incentivizing further developments in this area. Compared to the software packages typically used by researchers in the field, Spark offers much simpler parallelization and scaling of analyses to arbitrarily large data sizes, but lacks most of the practical tools essential for convenient setup of analysis pipelines as they exist in more commonly used languages (i.e., python or R). Stronger links

between these two worlds could allow for the development of analysis pipelines powerful enough to handle large datasets, yet as simple as any of the standard data applications.

ACKNOWLEDGMENTS

This study was financially supported by the Austrian Science Fund (P22813, P23533) and the Federal Ministry of Science,

Research and Economy via the *Hochschulraum-Strukturmittel* project. Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

REFERENCES

- ADHD-200 Consortium (2012). The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Front. Syst. Neurosci.* 6:62. doi: 10.3389/fnsys.2012.00062
- Akil, H., Martone, M. E., and Van Essen, D. C. (2011). Challenges and opportunities in mining neuroscience data. *Science* 331, 708–712. doi: 10.1126/science.1199305
- Assaf, Y., Alexander, D. C., Jones, D. K., Bizzi, A., Behrens, T. E. J., Clark, C. A., et al. (2013). The connect project: combining macro- and micro-structure. *Neuroimage* 80, 273–282. doi: 10.1016/j.neuroimage.2013.05.055
- Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734–4739. doi: 10.1073/pnas.0911855107
- Boubela, R. N., Huf, W., Kalcher, K., Sladky, R., Filzmoser, P., Pezawas, L., et al. (2012). A highly parallelized framework for computationally intensive MR data analysis. *MAGMA* 25, 313–320. doi: 10.1007/s10334-011-0290-7
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Craddock, R. C., Tungaraza, R. L., and Milham, M. P. (2015). Connectomics and new approaches for analyzing human brain functional connectivity. *Gigascience* 4, 13. doi: 10.1186/s13742-015-0045-x
- Dean, J., and Ghemawat, S. (2004). “Mapreduce: simplified data processing on large clusters,” in *Proceedings of the 6th Conference on Symposium on Operating Systems Design & Implementation - Volume 6, OSDI'04* (Berkeley, CA: USENIX Association). Available online at: <http://dl.acm.org/citation.cfm?id=1251254.1251264>
- Eickhoff, S. B., Thirion, B., Varoquaux, G., and Bzdok, D. (2016). Connectivity-based parcellation: critique & implications. *Hum. Brain Mapp.* doi: 10.1002/hbm.22933. [Epub ahead of print].
- Eklund, A., Andersson, M., and Knutsson, H. (2012). fMRI analysis on the GPU: possibilities and challenges. *Comput. Methods Prog. Biomed.* 105, 145–161. doi: 10.1016/j.cmpb.2011.07.007
- Eklund, A., Dufort, P., Forsberg, D., and LaConte, S. M. (2013). Medical image processing on the GPU - past, present and future. *Med. Image Anal.* 17, 1073–1094. doi: 10.1016/j.media.2013.05.008
- Eklund, A., Dufort, P., Villani, M., and LaConte, S. (2014). BROCCOLI: software for fast fMRI analysis on many-core CPUs and GPUs. *Front. Neuroinform.* 8:24. doi: 10.3389/fninf.2014.00024
- Freeman, J., Vladimirov, N., Kawashima, T., Mu, Y., Sofroniew, N. J., Bennett, D. V., et al. (2014). Mapping brain activity at scale with cluster computing. *Nat. Methods* 11, 941–950. doi: 10.1038/nmeth.3041
- Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *Neuroimage* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015
- Jiang, T. (2013). Brainnetome: a new -ome to understand the brain and its disorders. *Neuroimage* 80, 263–272. doi: 10.1016/j.neuroimage.2013.04.002
- Kalcher, K., Boubela, R. N., Huf, W., Nasel, C., and Moser, E. (2015). Identification of voxels confounded by venous signals using resting-state fMRI functional connectivity graph clustering. *Front. Neurosci.* 9:472. doi: 10.3389/fnins.2015.00472
- Marcus, D. S., Harwell, J., Olsen, T., Hodge, M., Glasser, M. F., Prior, F., et al. (2011). Informatics and data mining tools and strategies for the human connectome project. *Front. Neuroinform.* 5:4. doi: 10.3389/fninf.2011.00004
- Mennes, M., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2013). Making data sharing work: the FCP/INDI experience. *Neuroimage* 82, 683–691. doi: 10.1016/j.neuroimage.2012.10.064
- Nooner, K. B., Colcombe, S. J., Tobe, R. H., Mennes, M., Benedict, M. M., Moreno, A. L., et al. (2012). The NKI-rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front. Neurosci.* 6:152. doi: 10.3389/fnins.2012.00152
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org/>
- Satterthwaite, T. D., Connolly, J. J., Ruparel, K., Calkins, M. E., Jackson, C., Elliott, M. A., et al. (2016). The philadelphia neurodevelopmental cohort: a publicly available resource for the study of normal and abnormal brain development in youth. *Neuroimage* 124 (Pt B), 1115–1119. doi: 10.1016/j.neuroimage.2015.03.056
- The Apache Software Foundation (2015). *Apache Spark*. Wilmington, DE. Available online at: <https://spark.apache.org/>
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., et al. (2013). The WU-Minn human connectome project: an overview. *Neuroimage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041
- Van Horn, J. D., and Toga, A. W. (2014). Human neuroimaging as a “Big Data” science. *Brain Imaging Behav.* 8, 323–331. doi: 10.1007/s11682-013-9255-y
- Wang, Y., Goh, W., Wong, L., Montana, G., and the Alzheimer’s Disease Neuroimaging Initiative (2013). Random forests on Hadoop for genome-wide association studies of multivariate neuroimaging phenotypes. *BMC Bioinformatics* 14(Suppl. 16):S6. doi: 10.1186/1471-2105-14-S16-S6
- Xu, T., Yang, Z., Jiang, L., Xing, X.-X., and Zuo, X.-N. (2015). A connectome computation system for discovery science of brain. *Sci. Bull.* 60, 86–95. doi: 10.1007/s11434-014-0698-3
- Zadeh, R. B., and Carlsson, G. (2013). Dimension independent matrix square using mapreduce. Preprint arXiv:1304.1467.
- Zadeh, R. B., and Goel, A. (2012). Dimension independent similarity computation. *J. Mac. Learn. Res.* 14, 1605–1626.
- Zaharia, M., Chowdhury, M., Das, T., Dave, A., Ma, J., McCauley, M., et al. (2012). “Resilient distributed datasets: a fault-tolerant abstraction for in-memory cluster computing,” in *Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation* (Berkeley, CA: USENIX Association).
- Zuo, X.-N., Anderson, J. S., Bellec, P., Birn, R. M., Biswal, B. B., Blautzik, J., et al. (2014). An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci. Data* 1:140049. doi: 10.1038/sdata.2014.49
- Zuo, X.-N., Ehmke, R., Mennes, M., Imperati, D., Castellanos, F. X., Sporns, O., et al. (2012). Network centrality in the human functional connectome. *Cereb. Cortex* 22, 1862–1875. doi: 10.1093/cercor/bhr26
- Zuo, X.-N., and Xing, X.-X. (2014). Test-retest reliabilities of resting-state fmri measurements in human brain functional connectomics: a systems neuroscience perspective. *Neurosci. Biobehav. Rev.* 45, 100–118. doi: 10.1016/j.neubiorev.2014.05.009

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Boubela, Kalcher, Huf, Nasel and Moser. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Statistical image analysis of longitudinal RAVENS images

Seonjoo Lee^{1,2}, Vadim Zipunnikov³, Daniel S. Reich^{3,4} and Dzung L. Pham^{5*}

¹ Department of Psychiatry and Biostatistics, Columbia University, New York, NY, USA, ² New York State Psychiatric Institute, New York, NY, USA, ³ Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA, ⁴ Division of Neuroimmunology and Neurovirology, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA, ⁵ Center for Neuroscience and Regenerative Medicine, The Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, USA

Regional analysis of volumes examined in normalized space (RAVENS) are transformation images used in the study of brain morphometry. In this paper, RAVENS images are analyzed using a longitudinal variant of voxel-based morphometry (VBM) and longitudinal functional principal component analysis (LFPCA) for high-dimensional images. We demonstrate that the latter overcomes the limitations of standard longitudinal VBM analyses, which does not separate registration errors from other longitudinal changes and baseline patterns. This is especially important in contexts where longitudinal changes are only a small fraction of the overall observed variability, which is typical in normal aging and many chronic diseases. Our simulation study shows that LFPCA effectively separates registration error from baseline and longitudinal signals of interest by decomposing RAVENS images measured at multiple visits into three components: a subject-specific imaging random intercept that quantifies the cross-sectional variability, a subject-specific imaging slope that quantifies the irreversible changes over multiple visits, and a subject-visit specific imaging deviation. We describe strategies to identify baseline/longitudinal variation and registration errors combined with covariates of interest. Our analysis suggests that specific regional brain atrophy and ventricular enlargement are associated with multiple sclerosis (MS) disease progression.

Keywords: longitudinal functional principal component analysis, regional analysis of volumes examined in normalized space, voxel-based morphometry, multiple sclerosis, brain volume measurement

OPEN ACCESS

Edited by:

Jian Kang,
Emory University, USA

Reviewed by:

Tingting Zhang,
University of Virginia, USA
Linglong Kong,
University of Alberta, Canada

*Correspondence:

Dzung L. Pham
dzung.pham@nih.gov

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 16 May 2015

Accepted: 23 September 2015

Published: 20 October 2015

Citation:

Lee S, Zipunnikov V, Reich DS and
Pham DL (2015) Statistical image
analysis of longitudinal RAVENS
images. *Front. Neurosci.* 9:368.
doi: 10.3389/fnins.2015.00368

1. INTRODUCTION

Magnetic resonance imaging (MRI) is commonly used in the study of brain structure. Many studies are based on measurements of tissue volumes within a number of predefined regions of interest (ROIs); for example, see Bartzokis et al. (2001) and Bermel et al. (2003). Although ROI analysis can directly quantify the volume of structures and reduce the dimensionality of images, the ROIs have to be defined before the analysis is conducted. In disease studies, this can be difficult without sufficient prior knowledge about what and how various regions will be affected. Moreover, ROI based measurements can be time-consuming and labor-intensive. The results of the analysis will depend on the quality of the ROI delineation and thus depend upon the experience of the operator and accuracy of segmentation algorithms.

Voxel-based morphometry (VBM) is a complementary technique that measures local brain volumes in a normalized space and thus does not suffer from these limitations (Ashburner and Friston, 2000, 2001). In this work, we consider Regional Analysis of Volumes Examined in Normalized Space (RAVENS), which registers each subject brain to a template of anatomy so that the intensities of the RAVENS image represent regional volumes relative to those of template (Shen and Davatzikos, 2002). In voxel-based morphometry methods such as RAVENS, segmentations of structures such as the ventricles, are mapped to a template brain. If a subject's ventricles are larger than the template brain's ventricles, each voxel in the ventricles need to be shrunk to be mapped to the template. This in turn increases the intensity of the RAVENS map at each voxel, implying a larger volume was present in the subject at each voxel. **Figure 1** displays examples of ventricular RAVENS images in the template space. The first subject has much larger ventricles than the second subject (and template). Its RAVENS image of ventricles is displayed underneath the associated T1 image with red and blue colors representing higher and lower intensities, respectively. Subject 1 has larger ventricles, depicted by red in the RAVENS image. Similarly, the second brain, having a smaller ventricle than that of the first subject, has lower intensities in its RAVENS image, depicted by yellow and cyan in RAVENS image. By applying statistical VBM analysis of RAVENS images (RAVENS-VBM) to the resulting spatial distributions of gray matter (GM), white matter (WM), and ventricular cerebrospinal fluid (CSF), local atrophy or enlargement can be detected if the intensities significantly change across subjects.

In many disease studies, *longitudinal patterns* of brain structure between and within control and patient groups are of interest. Such studies are often based on ROI volume measurements followed by statistical analysis, such as a

linear mixed model. Several neuroimaging software platforms, including: FSL (Smith et al., 2004), the SPM-VBM toolbox (available at <http://dbm.neuro.uni-jena.de/vbm>) and SurfStat (Worsley et al., 2009), support flexible longitudinal models. Statistical inference of the contrast between two different time points is the most commonly used approach (Bendfeldt et al., 2011). Numerous other approaches for longitudinal imaging data have been proposed for prediction. The methods include support vector machine classifiers (Chen and Bowman, 2011) and Bayesian spatial models (Derado et al., 2013).

In practice, there are frequently cases that VBM does not find significant longitudinal trend. Possible causes are (1) the chosen statistical method is not sophisticated enough to extract longitudinal information; (2) a substantial amount of visit-to-visit variation to longitudinal signals exists; (3) heterogeneous longitudinal patterns exist within the diseases population.

The obvious solution to overcome such limitations is to combine the VBM analysis with more sophisticated statistical methods such as linear mixed models. However, for the first two cases, hypothesis driven VBM analyses cannot further exploit the data. In that case, figuring out the underlying structures of variation in the longitudinal data would be of interest. Further, we want to quantify the longitudinal and cross-sectional variability, and the association between each subject and their spatial patterns.

Thus, our main goal is to introduce a new statistical framework for longitudinal VBM analysis. To achieve the goal, we consider a data-driven analysis to provide a more complete statistical framework to analyze high-dimensional longitudinal brain images. A framework to allow for this conceptual partition of variability is longitudinal functional principal component analysis (LFPCA; Greven et al., 2011). This method was originally proposed for low to moderate dimensional functional data and

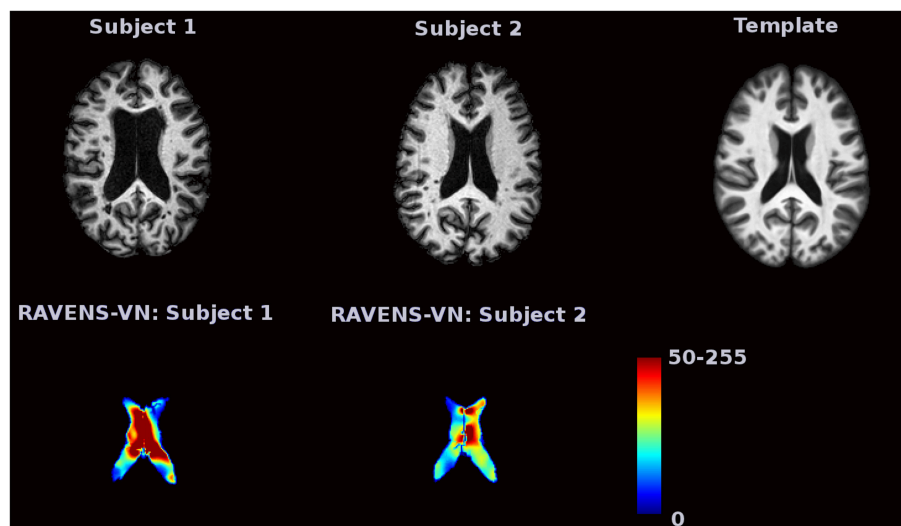


FIGURE 1 | The image intensities of the RAVENS image represent regional volumes relative to those of the template. Red color represents high intensity and blue color represents lower intensity. The first brain, having a larger ventricle than the template brain, has brighter intensities in the RAVENS image. The second brain, which has smaller ventricles, has lower intensities in the associated RAVENS image.

was extended to high dimensional data by Zipunnikov et al. (2014). The main idea of high-dimensional inference is based on projecting onto the intrinsic low dimensional space spanned by high-dimensional vectors (Di et al., 2009; Zipunnikov et al., 2011b). More precisely, we start by modeling the observed data with high-dimensional longitudinal functional principal component analysis (HD-LFPCA). Each RAVENS ventricular image is unfolded into a $p \times 1$ dimensional vector, where $p \approx 80,000$ is the number of voxels in the RAVENS ventricular image. These vectors are decomposed in their baseline, longitudinal and visit-to-visit components; each component is then estimated from the data. The method takes only a few minutes on a standard PC.

In this paper we focus on LFPCA as a useful tool for longitudinal voxel-based analyses, particularly to quantify cross-sectional and longitudinal variability in the data. The simulation study illustrates the application of LFPCA to a simplified imaging setting. It demonstrates that LFPCA effectively separates longitudinal, cross sectional, and other variations. Notably, the simulation study shows that LFPCA can separate registration errors from baseline and longitudinal components of interest.

2. MATERIALS AND METHODS

2.1. Participants

Forty eight MS patients (aged 42 ± 12 years at baseline) were enrolled in a longitudinal study of brain volume change. The study population included 33 female and 16 male patients; 28 patients with relapsing-remitting MS (RRMS), 13 patients with secondary progressive MS (SPMS), 5 patients with primary progressive MS (PPMS) and 2 patients with clinically isolated syndrome (CIS). One hundred forty eight T1 images have been acquired, with three images per subject for 44 subjects and 4 images per subject for 3 subjects. The average time interval between scans was 368 days (± 27). All images were spatially normalized via registration of T1 maps into the mean template, generated using Advanced Normalization Tools (Avants et al., 2010, 2011) from 30 randomly chosen MS patients among those with more than three visits. Ethical approval for the study was granted by IRB-2 and Johns Hopkins Medicine Institutional Review Board. All participants signed their fully informed consent.

2.2. MRI Protocol and Image Analysis

High resolution 3D magnetization-prepared rapid acquisition of gradient echoes (MPRAGE; acquired resolution: $1.1 \times 1.1 \times 1.1$ mm; TR: ~ 10 ms; TE: 6 ms; TI = 835 ms; flip angle: 8° ; SENSE factor: 2; averages: 1) were acquired on a 3.0 T MRI scanner (Intera, Philips Medical Systems).

In the processing, the follow-up images are affinely registered to their baselines via FMRIB's Linear Image Registration Tool (Jenkinson et al., 2002). All T1 images were segmented into GM, WM, VN, and lesions with Lesion-TOADS (Shiee et al., 2010) that was specifically designed for tissue and MS lesion segmentation. In general, as MS progresses, multifocal lesions in the white matter develop, and newly developed lesions can cause inaccuracies in the registration and RAVENS map computation.

Thus, we masked those lesions in the registration using the Lesions-TOADS software. After segmentation, the final tissue maps of GM, WM, and VN were normalized using HAMMER-SUITE (Shen and Davatzikos, 2002) to generate RAVENS images. Finally, the RAVENS maps were separately smoothed with 4 mm FWHM using SPM8.

2.3. Longitudinal Functional Principal Component Analysis

In this section, we provide a description of the original LFPCA approach developed by Greven et al. (2011) and its extension for high-dimensional data analysis (Zipunnikov et al., 2014). Throughout this section, we refer to both as LFPCA.

2.3.1. Random Intercept and Random Slope Model

Consider a longitudinal brain imaging study with subjects labeled by index i with each visit indexed by j and scan time by variable t_{ij} for $j = 1, \dots, J_i$. Each image is unfolded into a p -dimensional column vector $y_{ij}(v)$; the index v of each entry corresponds to a particular location in the brain for each subject and visit in normalized space. A random slope and random intercept model is commonly used to analyze longitudinal data, and it has been extended to functional (Greven et al., 2011) and imaging (Zipunnikov et al., 2014) studies as follows:

$$y_{ij}(v) = \eta(v, t_{ij}) + x_{i,0}(v) + x_{i,1}(v)t_{ij} + W_{ij}(v), \quad (2.1)$$

where $y_{ij}(v)$ denotes the image intensity at voxel v , $\eta(v, t_{ij})$ is a fixed main effect, and $x_{i,0}(v)$ and $x_{i,1}(v)$ denote the random intercept and random slope for subject i , respectively. The term $W_{ij}(v)$ is a random subject-visit specific imaging deviation, which is assumed to be a zero mean, second-order stationary random process uncorrelated with $\mathbf{X}_i(v) = (x_{i,0}(v), x_{i,1}(v))^T$. The covariance operators of $\mathbf{X}_i(v)$ and $W_{ij}(v)$ are denoted as $\mathbf{K}^X(v_1, v_2)$ and $\mathbf{K}^W(v_1, v_2)$, respectively.

While this is a natural and relatively simple model for longitudinally observed data, the scale of the problem requires aggressive dimensionality reduction. LFPCA reduces dimensionality by projecting onto the subspaces which explain principal directions of variation in the data. In model (2.1), there are two sources of variation: subject-to-subject, captured by \mathbf{X}_i , and visit-to-visit within a subject, captured by \mathbf{W}_{ij} and the model assumption on \mathbf{X}_i and \mathbf{W}_{ij} in (2.1) allows us to partition the variation of the data and LFPCA models latent processes \mathbf{X}_i and \mathbf{W}_{ij} using a Karhunen-Loeve (K-L) expansion (Karhunen, 1947; Loève, 1948).

The K-L expansion decomposes the two latent processes as $\mathbf{X}_i(v) = \sum_{k=1}^{\infty} \xi_{ik} \phi_k^X(v)$ and $W_{ij}(v) = \sum_{l=1}^{\infty} \zeta_{ijl} \phi_l^W(v)$, where $\phi_k^X = (\phi_k^{X,0}, \phi_k^{X,1})$ and ϕ_l^W are the eigenfunctions of $\mathbf{K}^X(v_1, v_2)$ and $\mathbf{K}^W(v_1, v_2)$, respectively, such that

$$\begin{aligned} \mathbf{K}^X(v_1, v_2) &= \begin{pmatrix} \mathbf{K}_{00}^X(v_1, v_2) & \mathbf{K}_{10}^X(v_1, v_2) \\ \mathbf{K}_{01}^X(v_1, v_2) & \mathbf{K}_{11}^X(v_1, v_2) \end{pmatrix} \\ &= \sum_{k=1}^{N_X} \lambda_k^X \phi_k^X(v_1) \{\phi_k^X(v_2)\}^T. \end{aligned}$$

LFPCA truncates K-L representations and represents observed data through a linear mixed-effects model:

$$y_{ij}(v) = \eta(v, t_{ij}) + \sum_{k=1}^{N_X} \xi_{ik} \phi_k^{X,0}(v) + t_{ij} \sum_{k=1}^{N_X} \xi_{ik} \phi_k^{X,1}(v) + \sum_{l=1}^{N_W} \zeta_{ijl} \phi_l^W(v),$$

$$(\xi_{ik_1}, \xi_{ik_2}) \sim (0, 0, \lambda_{k_1}^X, \lambda_{k_2}^X, 0); (\zeta_{il_1}, \zeta_{il_2}) \sim (0, 0, \lambda_{l_1}^W, \lambda_{l_2}^W, 0), \quad (2.2)$$

where “ $\sim (\mu_1, \mu_2; \sigma_1^2; \sigma_2^2; \rho)$ ” denotes that a pair of variables has a distribution with mean (μ_1, μ_2) , variance σ_1^2, σ_2^2 , and correlation ρ . We assume that $\lambda_{k_1} \geq \lambda_{k_2}$ if $k_1 \leq k_2$. Since $\mathbf{X}_i(v)$ and $W_{ij}(v)$ are uncorrelated, the scores $\{\xi_{ik}\}_{k=1}^\infty$ and $\{\zeta_{ijl}\}_{l=1}^\infty$ are also uncorrelated. A very important characteristic of model (2.2) is that both N_X and N_W are expected to be small in most applications.

For the unfolded vector, (2.2) can be rewritten as $\mathbf{y}_{ij} = \boldsymbol{\eta}(t_{ij}) + \boldsymbol{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \boldsymbol{\Phi}_X^1 \boldsymbol{\xi}_i + \boldsymbol{\Phi}_W \boldsymbol{\zeta}_{ij}$, where $\mathbf{y}_{ij} = (y_{ij}(v_1), \dots, y_{ij}(v_p))^\top$ is a $p \times 1$ dimensional vector; $\boldsymbol{\Phi}_X^0, \boldsymbol{\Phi}_X^1$, and $\boldsymbol{\Phi}_W$ are $p \times 1$ eigenvectors; $\boldsymbol{\Phi}_X^s = (\phi_1^{X,s}, \dots, \phi_{N_X}^{X,s})$ for $s = 0, 1$; $\boldsymbol{\Phi}_W = (\phi_1^W, \dots, \phi_{N_W}^W)$; $\boldsymbol{\xi}_i = (\xi_{i1}, \dots, \xi_{iN_X})^\top$; $\boldsymbol{\zeta}_i = (\zeta_{i1}, \dots, \zeta_{iN_W})^\top$.

In brain imaging data analysis, LFPCA can separate biological signals from non-biological artifacts. For example, registration errors due to structural differences between subjects can be captured by baseline subject-specific components $\boldsymbol{\Phi}_X^0$ and scanner variability can be captured by visit-to-visit components $\boldsymbol{\Phi}_W$. This will be illustrated via an extensive simulation experiment in Section 3.1.

The fixed effect $\eta(v, t_{ij})$ can be estimated in a number of ways (Greven et al., 2011). The analyses in the later sections simply use the sample mean across all the image observations. Once $\eta(v, t_{ij})$ is estimated by the sample mean $\bar{\eta}(v, t_{ij})$, the longitudinal eigenanalysis is applied to the residual images $\tilde{y}_{ij}(v) = y_{ij}(v) - \bar{\eta}(v, t_{ij})$ that are modeled as follows:

$$\tilde{y}_{ij} = \boldsymbol{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \boldsymbol{\Phi}_X^1 \boldsymbol{\xi}_i + \boldsymbol{\Phi}_W \boldsymbol{\zeta}_{ij}. \quad (2.3)$$

2.3.2. LFPCA Estimation

Zipunnikov et al. (2014) modified the original approach of Greven et al. (2011) and developed a method of moments estimator based on quadratics of \tilde{y}_{ij} . The $p \times p$ -covariance of \tilde{y}_{ij_1} and \tilde{y}_{ij_2} is given by

$$\mathbb{E} \left\{ \tilde{y}_{ij_1} \tilde{y}_{ij_2}^\top \right\} = \mathbf{K}_{00}^X + t_{ij_1} \mathbf{K}_{10}^X + t_{ij_2} \mathbf{K}_{10}^X + t_{ij_1} t_{ij_2} \mathbf{K}_{11}^X + \delta_{j_1, j_2} \mathbf{K}^W, \quad j_1, j_2 = 1, \dots, J_i, \quad (2.4)$$

where $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ otherwise. Model (2.4) can be rewritten in terms of unfolded vectors $\mathbf{K}^v = \{\text{vec} \mathbf{K}_{00}, \text{vec} \mathbf{K}_{01}, \text{vec} \mathbf{K}_{10}, \text{vec} \mathbf{K}_{11}, \text{vec} \mathbf{K}^W\}$ and $\mathbf{f}_{ij_1 j_2} = (1, t_{ij_2}, t_{ij_1}, t_{ij_1} t_{ij_2}, \delta_{j_1, j_2})^\top$ such that $\text{Evec} \tilde{y}_{ij_1} \tilde{y}_{ij_2}^\top = \mathbf{K}^v \mathbf{f}_{ij_1 j_2}$. By concatenating all vectors across all subjects and visits we obtain a

moment matrix identity for the $p^2 \times m$ matrix \mathbf{Y} : $\mathbf{EY} = \mathbf{K}^v \mathbf{F}$, where $m = \sum_{i=1}^N J_i^2$. Then covariance parameters \mathbf{K}^v can be unbiasedly estimated by using ordinary least squares (OLS): $\hat{\mathbf{K}}^v = \mathbf{YF}^\top (\mathbf{FF}^\top)^{-1}$.

The covariance operators \mathbf{K}^X and \mathbf{K}^W are $2p \times 2p$ and $p \times p$ dimensional, respectively. For high-dimensional functional data, storing or diagonalizing these matrices is not feasible. Zipunnikov et al. (2014) proposed HD-LFPCA, a novel estimation approach that takes advantage of an intrinsically small dimension of the space spanned by high-dimensional data vectors. First we form a $p \times J_i$ dimensional matrix $\tilde{\mathbf{y}}_i$, where column j corresponds to a demeaned-RAVENS image obtained for subject i at visit j . The $p \times J$ dimensional data matrix $\tilde{\mathbf{y}} = (\tilde{\mathbf{y}}_1; \dots; \tilde{\mathbf{y}}_n)$ is formed by column-binding the blocks of data corresponding to each subject, where $J = \sum_{i=1}^N J_i$. The data matrix can be decomposed as $\tilde{\mathbf{y}} = \mathbf{VSU}^\top$ using a singular value decomposition (SVD) approach. In the RAVENS image application, $J = 148$. Equation (2.3) can be rewritten as

$$\tilde{\mathbf{y}}_{ij} = \mathbf{VSU}_{ij} = \boldsymbol{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \boldsymbol{\Phi}_X^1 \boldsymbol{\xi}_i + \boldsymbol{\Phi}_W \boldsymbol{\zeta}_{ij}. \quad (2.5)$$

By multiplying with \mathbf{V}^\top to the left, we have

$$\mathbf{SU}_{ij} = \mathbf{V}^\top \boldsymbol{\Phi}_X^0 \boldsymbol{\xi}_i + t_{ij} \mathbf{V}^\top \boldsymbol{\Phi}_X^1 \boldsymbol{\xi}_i + \mathbf{V}^\top \boldsymbol{\Phi}_W \boldsymbol{\zeta}_{ij} = \mathbf{A}_X^0 \boldsymbol{\xi}_i + \mathbf{A}_X^1 \boldsymbol{\xi}_i + \mathbf{A}_W \boldsymbol{\zeta}_{ij}. \quad (2.6)$$

We estimate $\hat{\mathbf{A}}_X^0, \hat{\mathbf{A}}_X^1$, and $\hat{\mathbf{A}}_W$ as described earlier, and estimate $\hat{\boldsymbol{\Phi}}_X^0 = \mathbf{V} \hat{\mathbf{A}}_X^0$, $\hat{\boldsymbol{\Phi}}_X^1 = \mathbf{V} \hat{\mathbf{A}}_X^1$, and $\hat{\boldsymbol{\Phi}}_W = \mathbf{V} \hat{\mathbf{A}}_W$. Note that multiplying by \mathbf{V}^\top in Equation (2.5) reduces the model to its low-dimensional form (2.6), without losing the original correlation structure of the data. Once inference is conducted in model (2.6), then quantities of interest from model (2.5) can be estimated by pre-multiplying Equation (2.6) by \mathbf{V} .

Principal scores ξ_i and ζ_{ij} are estimated via Best Linear Unbiased Predictions (BLUPs) as follows. The stacked vector of i th subject data, $\text{vec} \tilde{\mathbf{y}}_i = (\tilde{y}_{i1}^\top, \dots, \tilde{y}_{ij_i}^\top)^\top$, can be rewritten as $\text{vec} \tilde{\mathbf{y}}_i = \mathbf{B}_i \boldsymbol{\omega}_i$, where $\mathbf{B}_i = (\mathbf{B}_i^X; \mathbf{B}_i^W)$, $\mathbf{B}_i^X = \mathbf{1}_{J_i} \otimes \boldsymbol{\Phi}_X^0 + \mathbf{T}_i \otimes \boldsymbol{\Phi}_X^1$, $\mathbf{B}_i^W = \mathbf{I}_{J_i} \otimes \boldsymbol{\Phi}_W$, where $\mathbf{T}_i = (t_{i1}, \dots, t_{ij_i})^\top$, $\boldsymbol{\omega}_i = (\boldsymbol{\xi}_i^\top, \boldsymbol{\zeta}_i^\top)^\top$, the subject level principal scores $\boldsymbol{\zeta}_i = (\zeta_{i1}^\top, \dots, \zeta_{ij_i}^\top)^\top$, and $\mathbf{1}_{J_i}$ is a $J_i \times 1$ vector of ones. Then the scores can be estimated as $\hat{\boldsymbol{\omega}}_i = (\hat{\mathbf{B}}_i^\top \hat{\mathbf{B}}_i)^{-1} \hat{\mathbf{B}}_i^\top \text{vec} \tilde{\mathbf{y}}_i$. Due to linearity the estimated scores are the same in both models (2.5) and (2.6). Details of the matrix calculation and additional theoretical results of HD-LFPCA can be found in Zipunnikov et al. (2014).

The computed subject-specific principal component scores ξ_i are the derived composite scores computed for each linear trajectories based on the eigenvectors for subject-specific PCs. These scores can be used as predictors or outcomes in subsequent regression analyses to evaluate relationships between high-dimensional longitudinal trajectories and other variables of interest. Also, we can apply cluster analysis on the scores to uncover latent structure in the sample.

2.4. Classical VBM Analysis using Linear Mixed Model

First, we applied traditional VBM analysis using a linear mixed model to find a longitudinal trend. Many previous longitudinal studies have applied pairwise comparisons between two time points (Driemeyer et al., 2008). This study attempts to discover constant longitudinal trends over the time, i.e., focusing on the atrophy or enlargement rates. This may elucidate disease progression patterns of the patients. For the i th subject j th visit, the RAVENS image at voxel v follows the model:

$$\begin{aligned} y_{ij}(v) &= \beta^0(v) + \beta^1(v)t_{ij} + b_i^0 + b_i^1 t_{ij} + \epsilon_{ij}(v), \\ b_i^0 &\sim N(0, \sigma_0^2(v)), \quad b_i^1(v) \sim N(0, \sigma_1^2(v)), \\ \text{Cov}(b_i^0(v), b_i^1(v)) &= \sigma_{12}, \quad \epsilon_{ij}(v) \sim N(0, \sigma_\epsilon^2(v)), \end{aligned} \quad (2.7)$$

where $\beta^0(v)$ and $\beta^1(v)$ are the fixed-effect coefficients, b_i^0 and $b_i^1(v)$ are the random-effect coefficients for subject i , $\epsilon_{ij}(v)$ is the error. The parameters are estimated based on maximum likelihood estimation and the p -values of the fixed effect parameters are computed controlling for false discovery rate using (Benjamini and Yekutieli, 2001). We perform the statistical analysis in R (version 2.15.1).

3. RESULTS

3.1. Simulated Images

In this section, we present a simulation study to test the performance of LFPCA in RAVENS-VBM analysis. We

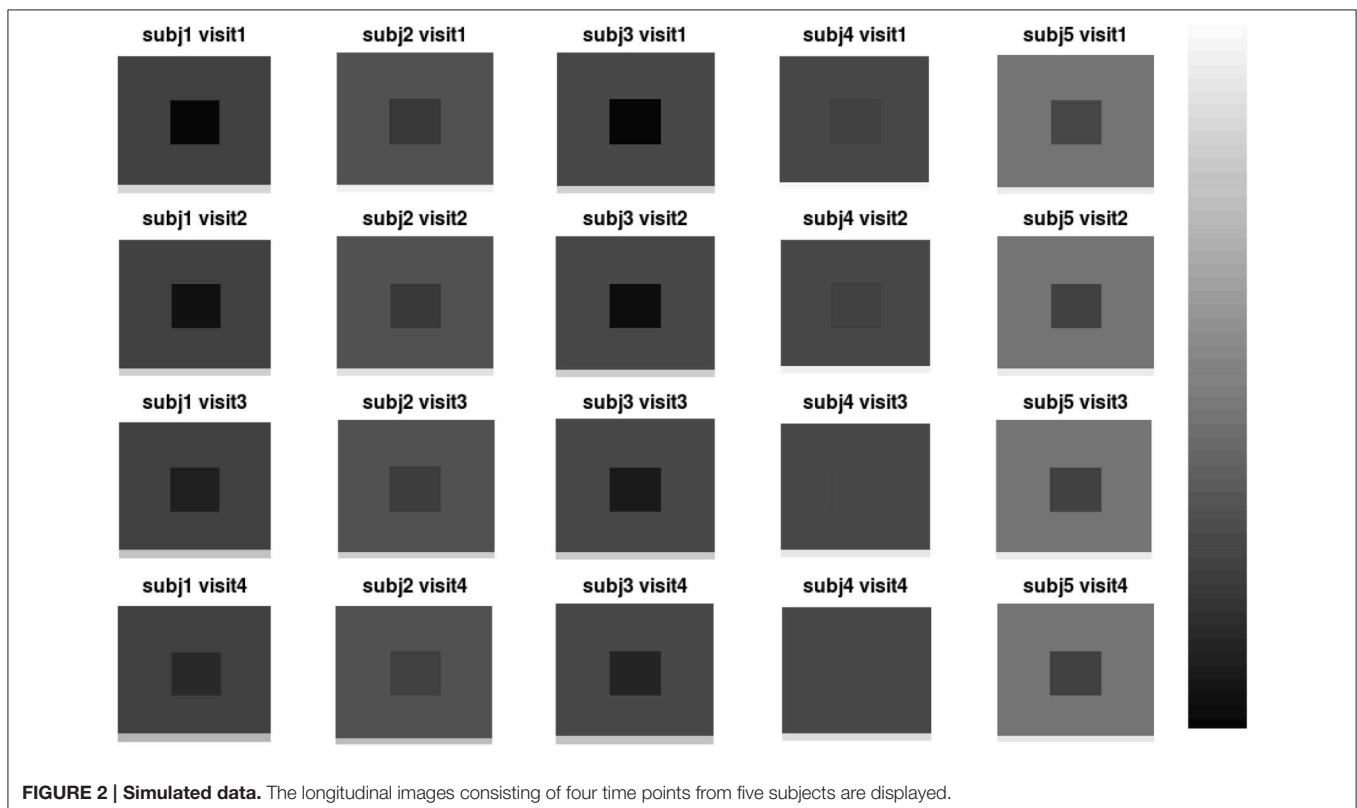
investigate if LFPCA can identify subject-specific signals from noise, particularly registration errors, which often dominate signals in VBM analyses. Also, we identify cross-sectional and longitudinal variation when they exist.

We design a simulation study to mimic longitudinal analysis of RAVENS images. For the purpose of illustration, we use 2D images with $200 \times 200 = 40,000$ pixels. We generate images from 50 subjects ($N = 50$) with three follow-ups. To replicate RAVENS processing routine, we assume that all images are registered to a template space. **Figure 2** displays simulated RAVENS images from 5 randomly chosen subjects. Each column represents four longitudinally collected images of the same subject.

Each image mimics four canonical brain structures: background (B), white matter (W), ventricles (V), and gray matter (G). Those four components are simplified and shown as a background, a big square, a small square inside the big square, and a rectangle at the bottom, respectively. Registration errors are introduced via random rigid shifts of simulated structures as described below.

In **Figure 2**, the images from the first subject, which are displayed at the first column, show the longitudinal patterns. In the images, the color of V changes from darker gray to brighter gray, which represents longitudinal enlargements of V. Similarly, the colors of W and G changed to darker colors, which represent longitudinal atrophy.

Figure 3 shows the first five pairs of subject-specific components (Φ_X). The baseline components (Φ_X^0) are displayed in the top row and their corresponding longitudinal components (Φ_X^1) are displayed in the second row. Each image is colored with



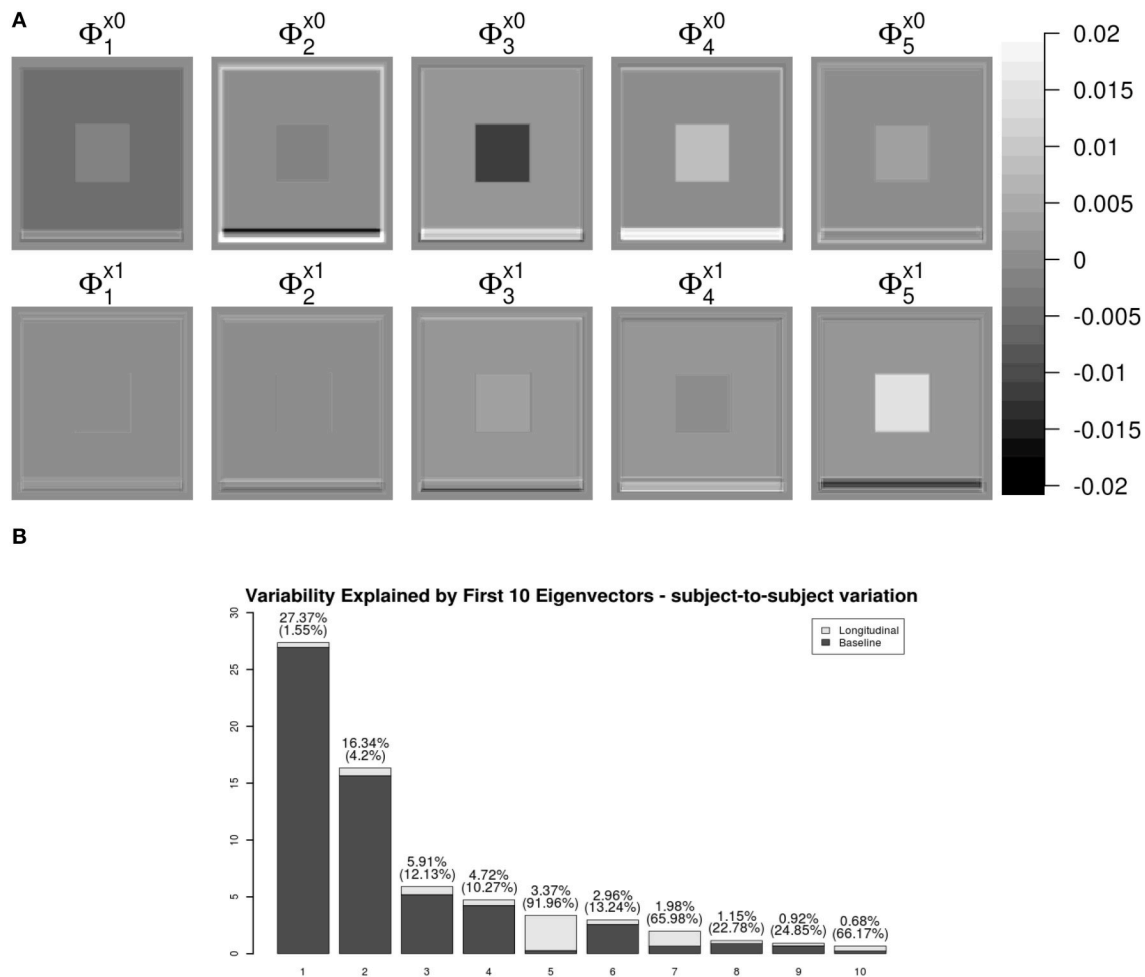


FIGURE 3 | LFPCA subject-specific components Φ_X . (A) Baseline (Φ_X^0) and Longitudinal (Φ_X^1) Components. (B) Variation explained by subject-specific components with ratios of longitudinal and baseline components.

a black(negative)-gray(0)-white(positive) color scheme. The first subject-specific component (Figure 3A first column) represents cross-sectional variations of the intensities of the W. The second component captures subject-specific registration errors, which only depend on cross-sectional variation. The third and fourth components represent the size of V and G. The fifth component shows longitudinal patterns of V and G. For a subject with positive score, the area V enlarges over the time and that of G shrinks, matching the truth used in simulation.

One useful feature of LFPCA is that contributions of the longitudinal and baseline components within each subject-specific component can be quantified on a [0, 1] scale. A subject-specific eigenvector is the stacked vector of baseline and longitudinal components: $\Phi_{X,k} = \left\{ \Phi_{X,k}^0, \Phi_{X,k}^1 \right\}^\top$, such that $\|\Phi_{X,k}\|^2 = \|\Phi_{X,k}^0\|^2 + \|\Phi_{X,k}^1\|^2 = 1$. For each component, the variation or the contribution of the longitudinal component can be calculated as $\frac{\|\Phi_{X,k}^1\|^2}{\|\Phi_{X,k}^0\|^2 + \|\Phi_{X,k}^1\|^2}$. Combined with the contribution of each subject-specific component to

the total variation, Figure 3B displays variations explained by the first 10 subject-specific components with the proportion of the longitudinal components within each subject. Each bar plot intensity represents the amount of variation explained by each subject-specific component and is comprised of variations explained by the longitudinal component (dark) and the baseline component (bright). The top of each bar displays numerical values of the variation explained by the subject-specific component with the variation explained by the longitudinal component within the subject-specific component in parenthesis. Note that the fifth principal component has the highest longitudinal-baseline ratio among all 10 components. This provides a strong indication that the fifth component should be essentially treated as a longitudinal component. Using both visual and quantitative methods, we can conclude that the first four components represent baseline variation and registration error and the fifth component reveals longitudinal variation. In the data set, the longitudinal variation and baseline variations are independent, which agrees with the simulation setting.

An advantage of LFPCA is its ability to couple baseline and longitudinal variation. The longitudinal component is added to the baseline with the time used as a multiplicative weight. **Figure 4** illustrates the temporal trajectory of principal component loadings. We display only the first component, which does not appear to change over time. This pattern is replicated in the first four components. This indicates that the first four components mostly represent baseline variation. The fifth component loading does appear to change over time, while the baseline loading has relatively lower intensities compared to the longitudinal loading.

To summarize, our simulation studies convincingly demonstrate the power and flexibility of LFPCA to address some of key challenges of brain imaging. In particular, LFPCA managed to estimate and separate longitudinal and cross-sectional variation in a complex imaging simulation design with registration errors. The main part of the analysis can be

automated and performed robustly with no operator input. We also applied a classical VBM-linear mixed effect model for the simulated data. As we expected, the linear mixed effect model could identify linear trend in the ventricular area (V), but it did not find significant trend in other areas (W and G) due to low longitudinal changes in signal and high visit-to-visit variation.

3.2. Classical VBM Analysis using Linear Mixed Model

In this section, we apply a standard VBM analysis to the MS cohort described in Section 2. This analysis focuses on the population mean of the longitudinal trend $\beta_1^l(v)$. After an FDR correction (Benjamini and Yekutieli, 2001) combined with cluster level thresholding, there are significant clusters with spatial extent more than 20 voxels. **Table 1** shows information about the significant clusters, including cluster size, maximum or minimum t -values within each

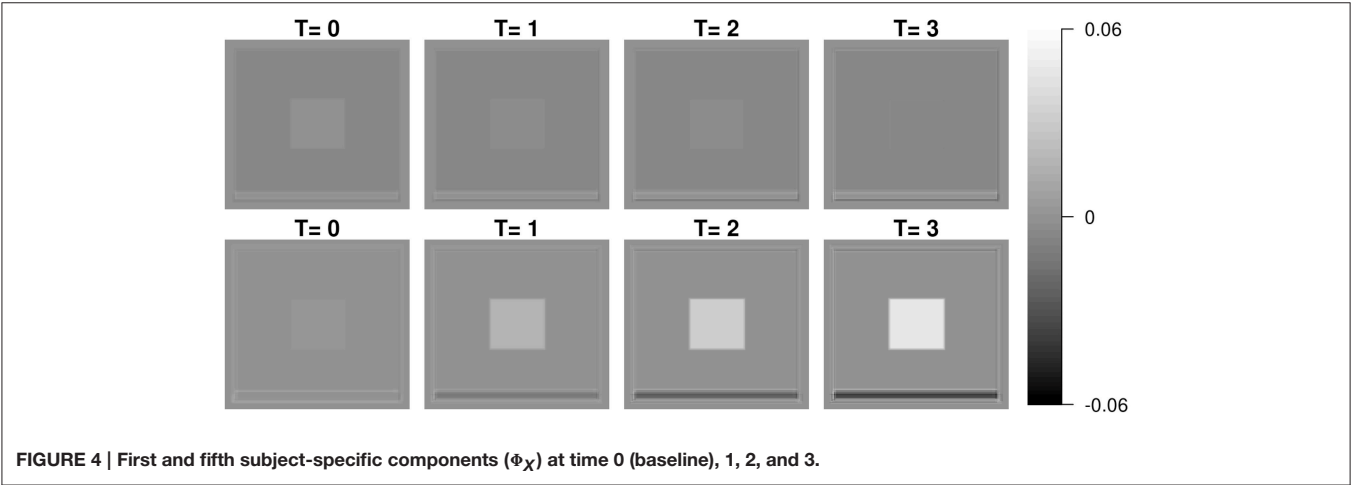


TABLE 1 | Significant clusters of GM/WM/VN VBM results.

		Cluster size	t-MAX ^a (or MIN)	t-MAX(or MIN) ^b			t-COG ^c		
				X	Y	Z	X	Y	Z
GM	Atrophy	112	−6.34	139	86	67	139	90	68.9
		79	−6.36	118	158	35	116	158	35.0
		40	−5.84	122	166	41	121	166	41.0
		28	−6.13	136	148	80	136	149	79.8
		15	−5.36	143	167	89	143	166	89.3
VN	Enlargement	111	5.59	161	151	70	157	152	72.4
WM	Enlargement	154	6.30	118	150	76	117	150	75.6
		100	5.75	127	111	85	128	111	84.8
	Atrophy	210	−5.97	126	83	59	124	83.7	57.6
		157	−5.77	106	98	83	107	98	83.5
		31	−5.46	118	154	37	118	154	36.7

All clusters are small (< 250) and spatially scatter.
^aMaximum t -value for positive values and minimum t -value for negative values.
^bLocation of maximum (minimum) Z -value (Z -MAX(MIN)).
^cCentre of gravity (COG) of the cluster.

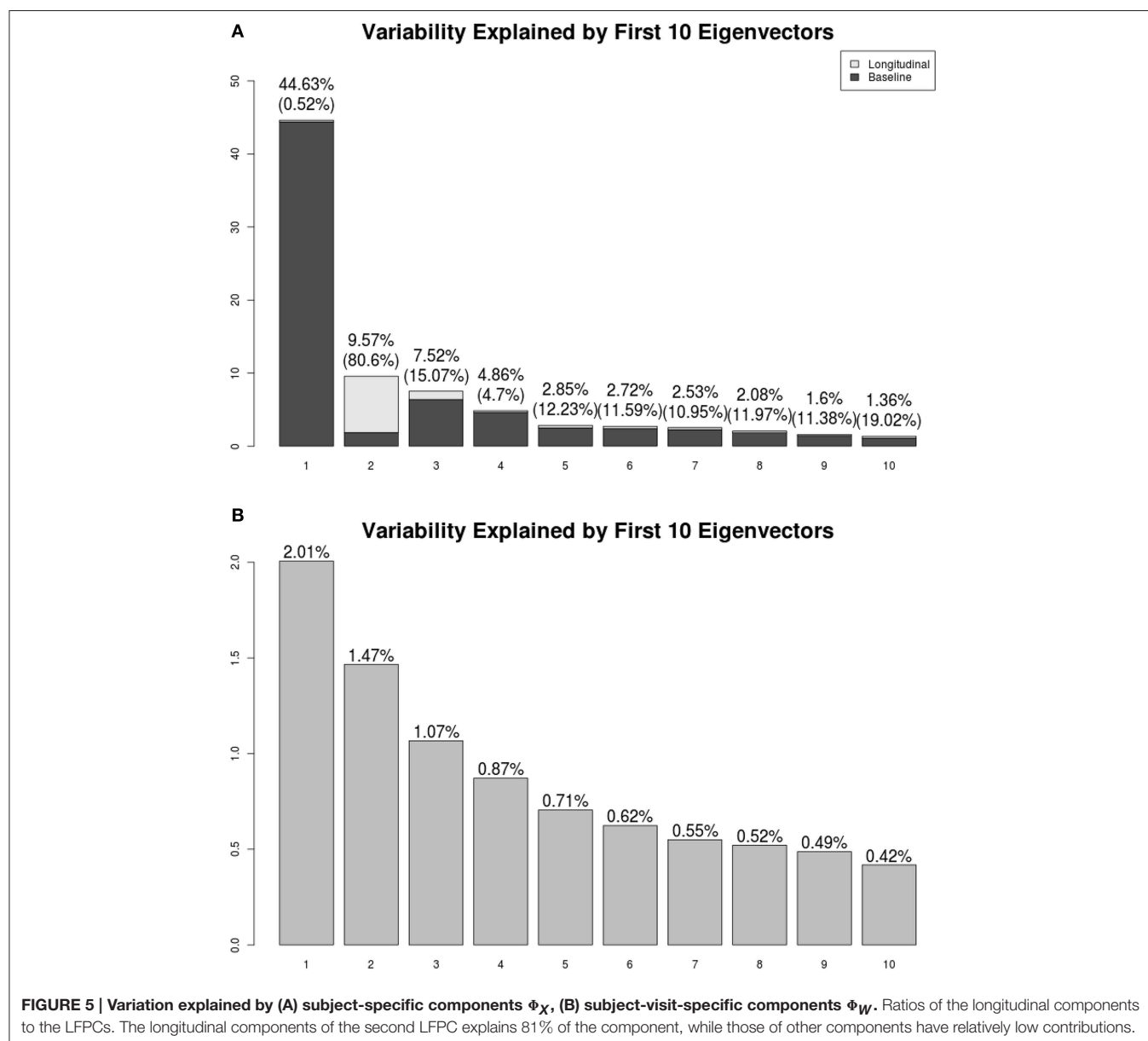
cluster and its location, and center of gravity of the clusters. We do not include an image of VBM results, since the clusters are very sparse having small cluster sizes. The results show that the longitudinal patterns do not appear to be significant for the most of brain regions. We suspect that it is because of the large variation within and between images due to real anatomic variation as well as registration error. We apply HD-LFPCA to uncover more subtle underlying variation.

3.3. HD-LFPCA Results

We present the LFPCA results for ventricular RAVENS images in **Figure 5**. **Figures 5A,B** display the amount of variation explained by subject-specific components and subject and visit-specific components to the total variation, respectively. **Figure 5A** displays variation explained by the first 10 subject-specific

components along the proportion of longitudinal variation represented within each subject-specific component. Each bar's height represents the percent of variation explained by each subject-specific component. It is color coded by the proportion of the variation explained by the longitudinal component and the baseline component. The top of each bar displays the variation explained by the subject-specific component; the fraction of that variation that is explained by the longitudinal component is given in parentheses.

The first subject-specific LFPC explains 45% of the overall variation, almost completely due to the cross-sectional part. The longitudinal part explains 81% of the variation within the second subject-specific LFPC. **Figure 5B** displays variation explained by the first 10 subject-visit-specific components to the total variation. The remaining LFPCs explain less than 1% of the total variation.



Most of the subject-specific LFPCs are driven by cross-sectional variation, which possibly include registration errors. The longitudinal changes are mainly captured by the second LFPC, which explains about 8% of the total variation. This provides an explanation as to why traditional VBM using linear mixed models did not find meaningful longitudinal patterns.

Figures 6, 7 shows the first two pairs of LFPCs of ventricles. **Figures 6A,B** show the baseline and longitudinal components of the first LFPC. The LFPC loadings are color-coded with a red-yellow color scheme for positive values and blue-cyan for negative values. The first components reveal baseline ventricular morphometric variation, while the longitudinal component has relatively lower intensities. To investigate the characteristics of the first component, we fit the linear regression with covariates of interest and volumes of 6 ROIs obtained by the Lesion-TOADS segmentation algorithm. **Figure 7C** displays scatter plots of the LFPC scores with covariates, baseline age, baseline Expanded Disability Status Scale (EDSS) score, and volumes of 6 ROIs (thalamus, ventricle, cortical gray matter volume, caudate, sulcal CSF, putamen). The dashed lines overlaid on the scatter plots are the linear regression lines and are colored as red when the linear trend is significant and colored as green otherwise.

The significant correlation between the first subject-specific score and baseline VN volume (first row, fourth column) confirms that the first component represents baseline variation ($R^2: 0.9684$), i.e., a subject with a positive score has larger ventricles at the baseline. The scores are significantly correlated to the subject's baseline age ($R^2: 0.1402$) and three gray matter ROIs (thalamus, caudate, and putamen).

Figures 7A,B display the baseline and longitudinal components of the second subject-specific component, respectively. A subject with a positive score tends to have a larger regional volume at the baseline (yellow-red colored voxels in **Figure 7A**) and longitudinal enlargement. The second subject-specific scores have significant correlations with the baseline age, EDSS, thalamus volume, VN volumes, sulcal CSF volume and putamen volume. The second component scores have higher correlation with baseline age ($R^2: 0.2371$) and EDSS ($R^2: 0.2053$) than the first component scores that represent cross-sectional variation. This indicates that the spatial patterns of longitudinal enlargement in ventricles are superior for modeling disease progression and age compared to simple ventricular volume measures.

We have applied a similar analysis to gray matter and white matter RAVENS images. **Figure 8** shows variation explained by first 10 subject-specific LFPCs in gray matter and white matter RAVENS images. In gray matter, around 20% variation comes from the longitudinal part across all subject-specific LFPCs. Lower proportions of variation, around 15%, are explained by the longitudinal part in white matter. Unlike the ventricles, any subject-specific component of gray and white matter is not dominated by the longitudinal part. We speculate it is due to spatial heterogeneity of longitudinal brain atrophy, which may depend on subject-specific disease progression. In the correlation analyses with age and EDSS scores, the first LFPC component of the gray matter was significantly associated with age ($r = -0.48$, $p < 0.001$) and EDSS ($r = -0.57$, $p < 0.001$) indicating gray

matter thinning is highly associated with age and MS progression, while other components were not significantly associated with age or EDSS. For white matter, LFPC1 was not significantly associated with age ($r = -0.07$, $p = 0.63$) but with EDSS ($r = -0.32$, $p = 0.03$). LFPC2 was significantly associated with both age ($r = -0.36$, $p = 0.01$) and EDSS ($r = -0.34$, $p = 0.02$). Although those LFPC components did not contain substantial longitudinal changes, the results indicate that local atrophy in the white matter can inform disease progression.

As described above, LFPCA is a useful dimension reduction tool for high-dimensional longitudinal data. In this section, we illustrated how the LFPC scores can be used in the correlation analyses. Further, LFPCA scores can be used as predictors or outcomes in regression analyses, classification or cluster analysis.

4. DISCUSSION

In this manuscript, we described and evaluated a coherent methodology for the study of longitudinal RAVENS—or other methodological—volumetric imaging studies. Our simulation studies demonstrate that LFPCA tightly links the analysis methodology with the morphometric image processing stream. We demonstrated that LFPCA can uncover interesting, yet subtle, directions of longitudinal variation in a case where independent voxel-level investigations fail. Of note, this study represents the first application of the high dimensional variation of LFPCA to voxel-based morphometric analysis. Related work includes Zipunnikov et al. (2014), who investigated DTI imaging data and Zipunnikov et al. (2011a) and Zipunnikov et al. (2011b), who studied RAVENS images with cross-sectional and clustered (but not longitudinal) settings. Moreover, this manuscript represents the first application of LFPCA to morphometric imaging in multiple sclerosis.

A key insight from the simulation studies is the ability of LFPCA to uncover interesting directions of variation in the presence of errors from registration to a template. Previously, registration errors were handled via either extremely aggressive smoothing during post-registration processing or by improved normalization algorithms. While improved algorithms are certainly a desirable goal, all normalization algorithms must be tuned and suffer from trade offs (such as bias and variance). Our results suggest the possibility of employing less aggressive normalization.

The performance of LFPCA depends on the number of subjects, the number of time points, and time span over which data is collected. In designing imaging studies for LFPCA, having both a large number of subjects and a large number of visits may be challenging to obtain. Simulation studies we have conducted during the process of examining LFPCA showed that LFPCA performed well as long as we have either many subjects with smaller number of visits or smaller number of subjects with many visits. It is recommended to make the time span over which data is collected roughly similar across subjects, and long enough to observe longitudinal changes.

In our study of MS, we found that the majority of variation is focused in cross-sectional components. This will likely be true in any study of adults, as variation in head size, brain

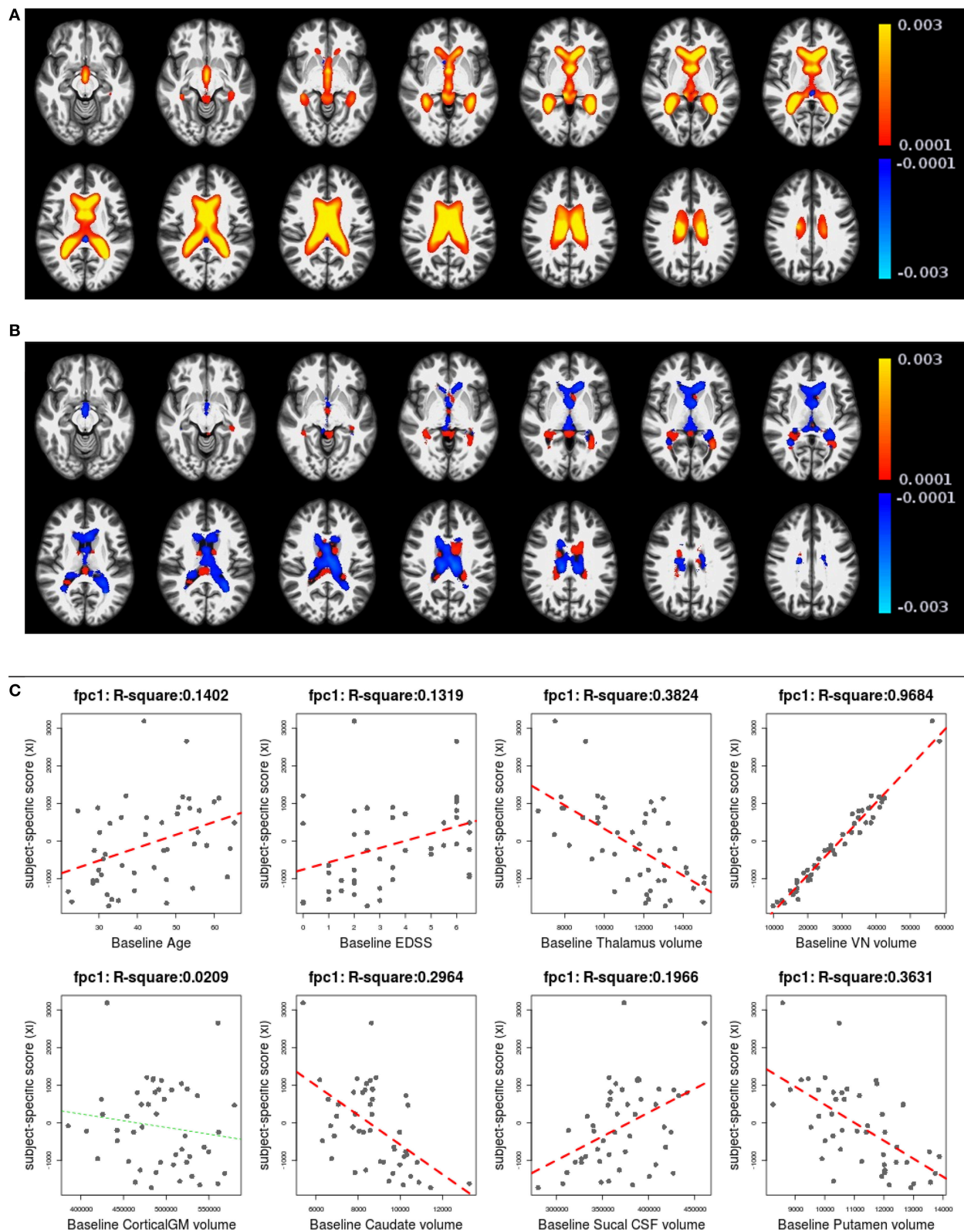


FIGURE 6 | The first subject-specific LFPC and scores. (A) The baseline map, **(B)** the longitudinal map, **(C)** the first subject-specific LFPC scores and covariates of interest. The baseline map represents cross-sectional variation while the longitudinal map has lower loadings. The LFPC scores have higher correlations with baseline age, EDSS, volumes of GM substructures (thalamus, caudate, putamen), ventricle, and sulcal CSF. The correlation with ventricular volume is very high, which confirms that the first LFPC represents cross-sectional ventricle size.

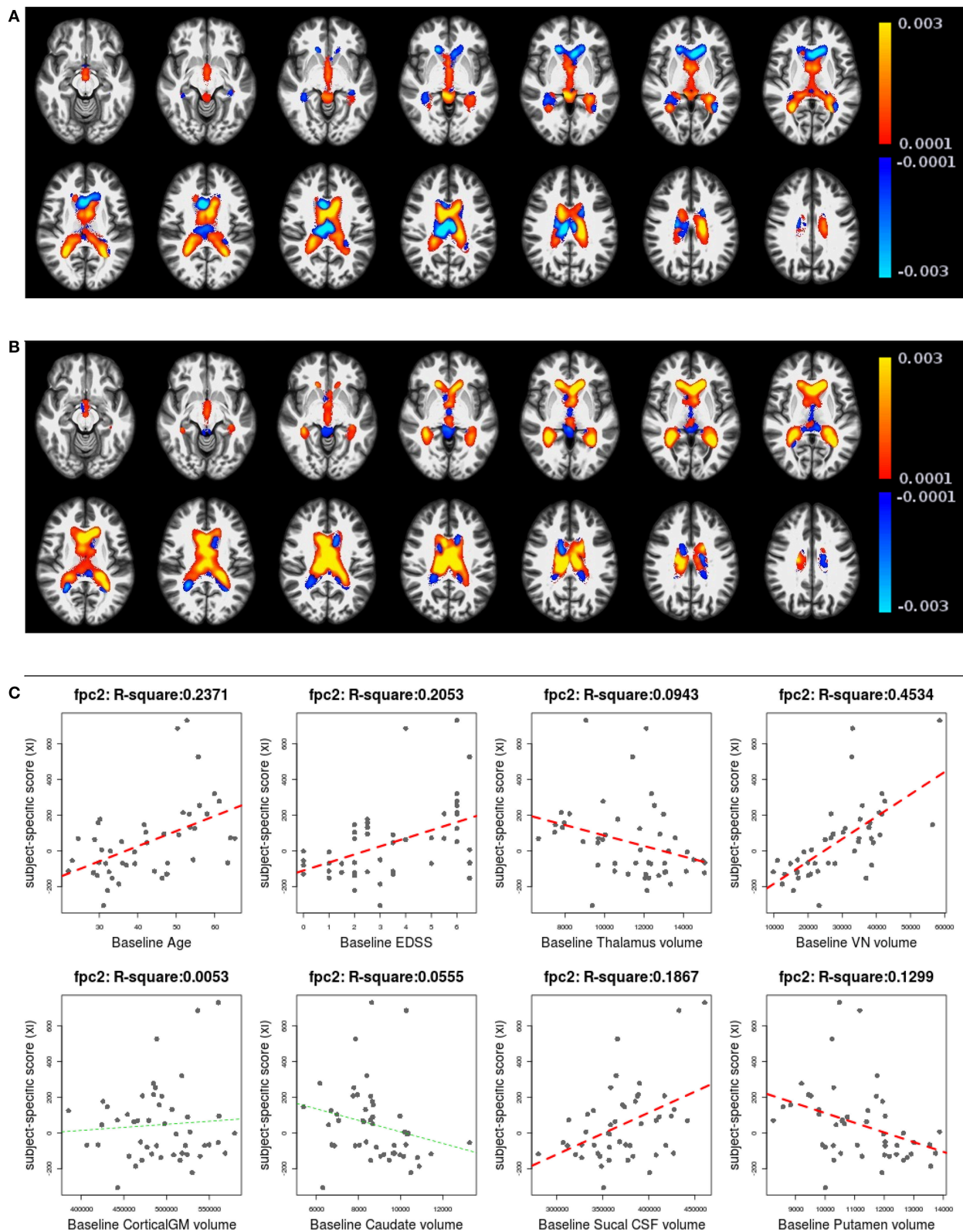
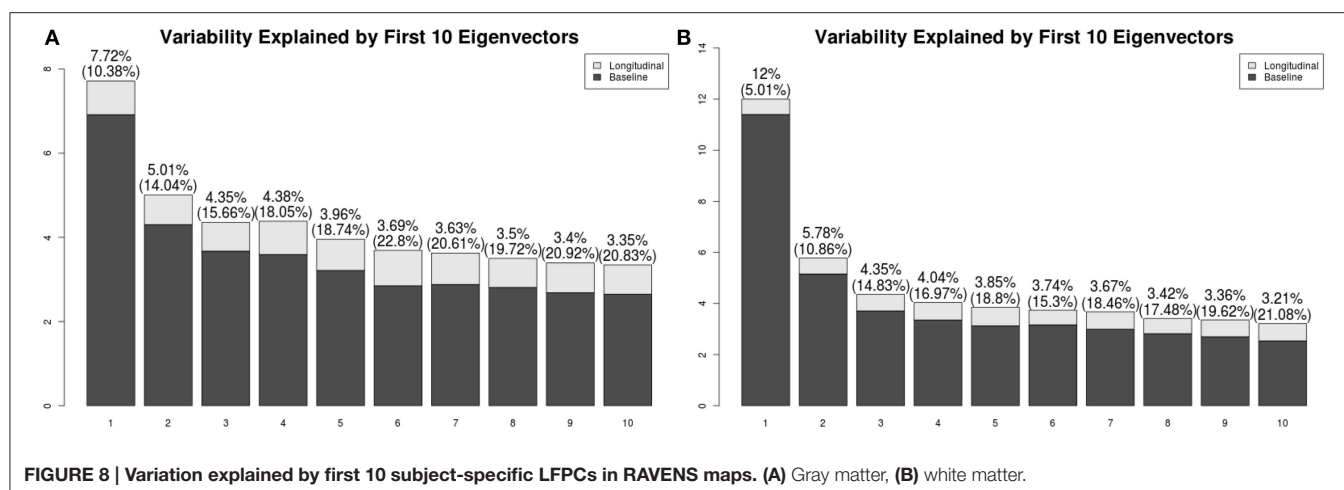


FIGURE 7 | The second subject-specific LFPCA and scores. (A) The baseline map, **(B)** the longitudinal map, **(C)** the second subject-specific LFPCA scores and covariates of interest. The baseline map has relatively lower loadings while the longitudinal map shows an enlargement pattern in the ventricles. The LFPCA scores have higher correlations with baseline age, EDSS, volumes of gray matter substructures (thalamus, putamen), ventricle, and sulcal CSF. The correlations with baseline age and EDSS scores are higher than those of the first LFPCA scores.



size and intracranial volume will vary far more substantially than longitudinal decline, not unlike if one were to study adult cross-sectional and longitudinal trends in heights. It would be of interest to apply LFPCA to developmental populations or populations with severe progressive brain disorders significantly after disease onset.

The correlation between subject-specific LFPC scores of ventricles and EDSS indicates that EDSS is better associated with longitudinal ventricular enlargement than baseline ventricular size. This implies ventricular enlargement is a sensitive measurement of disease progression. Some cross-sectional MS patient studies have reported that brain atrophy is related to irreversible clinical disability in (MS) and ventricular enlargement may be a sensitive marker of this tissue loss that is seen at all stages of MS (Turner et al., 2001; Benedict et al., 2005; Hildebrandt et al., 2006; Tekok-Kilic et al., 2007). In existing longitudinal studies, longitudinal ventricular enlargement and gray matter atrophy have been detected in both ROI and VBM with paired *t*-test or factor models (Simon et al., 1999; Dalton et al., 2002, 2006; Kalkers et al., 2002; Sepulcre et al., 2006; Lukas et al., 2010; Bendfeldt et al., 2009), which agrees with our finding. Unlike other methods, LFPCA is able to show spatially heterogeneous patterns of longitudinal enlargement, which ROI based methods cannot provide.

In the manuscript, we employed a registration strategy similar to Ashburner and Ridgway (2012). Recent developments in longitudinal registration algorithms (e.g., 4DHammer: Shen et al., 2003, CLASSIC: Shen et al., 2005, GLIRT: Wu et al., 2010, Quarc: Holland et al., 2011) are potentially capable of providing higher accuracy in tracking within subject anatomical changes. Improvements in registration would likely increase the sensitivity of LFPCA to subject-specific signals. However, visit-to-visit variation caused by image acquisition inconsistencies or large anatomical differences between subjects often cannot be corrected by registration. An advantage of the proposed method is that it can simultaneously quantify and characterize both cross-sectional and longitudinal signals of

interest in the presence of potentially large amounts of visit-to-visit variation.

As demonstrated previously for longitudinal diffusion imaging analysis, and here for longitudinal voxel-based morphometry, LFPCA is a compelling alternative to linear mixed model analysis for exploring spatial patterns of anatomical variation within and across subjects. We emphasize that this approach is not limited to a specific brain modality. Besides neuroimaging, we look forward to seeing this method is applied to many other exciting studies including epigenetics. For example, genome-wide DNA methylation data collected at multiple time points could be analyzed to study mechanisms of epigenetic changes related to certain diseases (Martino et al., 2014) or environmental exposure (Martino et al., 2013). Another potential domain of application is for analyzing dynamic imaging data, such as functional MRI or motion imaging. Such data often possess much larger numbers of time points, which would be needed to model the more complex variations in signal.

The LFPCA method described here is designed to model a linear trajectory over time. Given a relatively small number of visits (e.g., three visits on average) it is not feasible to model non-linear trends. However, if the data are collected over greater than 5 time points, the modeling of non-linear trajectories is possible. Currently, we are under a preliminary development of a method to extend LFPCA for non-linear trends modeled using spline functions. Further investigation on the numerical stability and performance will be conducted in the near future.

FUNDING

Research reported in this work was supported by National Institute of Health under award numbers R01NS070906, Z01NS003119, K01AG051348, R01HL12407 and R01NS060910. Support for this work included funding from the Department of Defense in the Center for Neuroscience and Regenerative Medicine.

REFERENCES

- Ashburner, J., and Friston, K. (2000). Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821. doi: 10.1006/nimg.2000.0582
- Ashburner, J., and Friston, K. (2001). Why voxel-based morphometry should be used. *Neuroimage* 14, 1238–1243. doi: 10.1006/nimg.2001.0961
- Ashburner, J., and Ridgway, G. R. (2012). Symmetric diffeomorphic modeling of longitudinal structural mri. *Front. Neurosci.* 6:197. doi: 10.3389/fnins.2012.00197
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., and Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* 54, 2033–2044. doi: 10.1016/j.neuroimage.2010.09.025
- Avants, B. B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., et al. (2010). The optimal template effect in hippocampus studies of diseased populations. *Neuroimage* 49, 2457–2466. doi: 10.1016/j.neuroimage.2009.09.062
- Bartzikos, G., Beckson, M., Lu, P. H., Nuechterlein, K. H., Edwards, N., and Mintz, J. (2001). Age-related changes in frontal and temporal lobe volumes in men: a magnetic resonance imaging study. *Arch. Gen. Psychiat.* 58:461. doi: 10.1001/archpsyc.58.5.461
- Bendfeldt, K., Hofstetter, L., Kuster, P., Traud, S., Mueller-Lenke, N., Naegelin, Y., et al. (2011). Longitudinal gray matter changes in multiple sclerosis: differential scanner and overall disease-related effects. *Hum. Brain Mapp.* 33, 1225–1245. doi: 10.1002/hbm.21279
- Bendfeldt, K., Kuster, P., Traud, S., Egger, H., Winkhofer, S., Mueller-Lenke, N., et al. (2009). Association of regional gray matter volume loss and progression of white matter lesions in multiple sclerosis—a longitudinal voxel-based morphometry study. *Neuroimage* 45, 60–67. doi: 10.1016/j.neuroimage.2008.10.006
- Benedict, R. H., Zivadinov, R., Carone, D. A., Weinstock-Guttman, B., Gaines, J., Maggiore, C., et al. (2005). Regional lobar atrophy predicts memory impairment in multiple sclerosis. *Am. J. Neuroradiol.* 26, 1824–1831.
- Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188.
- Bermel, R. H., Innus, M. D., Tjoa, C. W., and Bakshi, R. (2003). Selective caudate atrophy in multiple sclerosis: a 3D MRI parcellation study. *Neuroreport* 14, 335–339. doi: 10.1097/00001756-200303030-00008
- Chen, S., and Bowman, F. (2011). A novel support vector classifier for longitudinal high-dimensional data and its application to neuroimaging data. *Stat. Anal. Data Mining* 4, 604–611. doi: 10.1002/sam.10141
- Dalton, C. M., Brex, P. A., Jenkins, R., Fox, N. C., Miszkil, K. A., Crum, W. R., et al. (2002). Progressive ventricular enlargement in patients with clinically isolated syndromes is associated with the early development of multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* 73, 141–147. doi: 10.1136/jnnp.73.2.141
- Dalton, C. M., Miszkil, K. A., O'Connor, P. W., Plant, G. T., Rice, G. P., and Miller, D. H. (2006). Ventricular enlargement in ms. *Neurology* 66, 693–698. doi: 10.1212/01.wnl.0000201183.87175.9f
- Derado, G., Bowman, F. D., Zhang, L., and Alzheimer's Disease Neuroimaging Initiative. (2013). Predicting brain activity using a Bayesian spatial model. *Stat. Methods Med. Res.* 22, 382–397. doi: 10.1177/0962280212448972
- Di, C., Crainiceanu, C., Caffo, B., and Punjabi, N. (2009). Multilevel functional principal component analysis. *Ann. Appl. Stat.* 3, 458–488. doi: 10.1214/08-AOAS206
- Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., and May, A. (2008). Changes in gray matter induced by learning—revisited. *PLoS ONE* 3:E2669. doi: 10.1371/journal.pone.0002669
- Greven, S., Crainiceanu, C., Caffo, B., and Reich, D. (2011). “Longitudinal functional principal component analysis,” in *Recent Advances in Functional Data Analysis and Related Topics*, ed F. Ferraty (Berlin; Heidelberg: Physica-Verlag HD), 149–154. doi: 10.1007/978-3-7908-2736-1_23
- Hildebrandt, H., Hahn, H. K., Kraus, J. A., Schulte-Herbrüggen, A., Schwarze, B., and Schwendemann, G. (2006). Memory performance in multiple sclerosis patients correlates with central brain atrophy. *Mult. Sci.* 12, 428–436. doi: 10.1191/1352458506ms1286oa
- Holland, D., Dale, A. M., and Alzheimer's Disease Neuroimaging Initiative. (2011). Nonlinear registration of longitudinal images and measurement of change in regions of interest. *Med. Image Anal.* 15, 489–497. doi: 10.1016/j.media.2011.02.005
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. doi: 10.1006/nimg.2002.1132
- Kalkers, N. F., Ameiziane, N., Bot, J. C., Minneboo, A., Polman, C. H., and Barkhof, F. (2002). Longitudinal brain volume measurement in multiple sclerosis: rate of brain atrophy is independent of the disease subtype. *Arch. Neurol.* 59:1572. doi: 10.1001/archneur.59.10.1572
- Karhunen, K. (1947). Über lineare methoden in der wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae. A Math. Phys.* 1947:79.
- Loève, M. (1948). *Fonctions Aléatoires du Second Ordre*. Supplement to P. Lévy, Processus Stochastiques et Mouvement Brownien. Paris: Gauthier-Villars.
- Lukas, C., Minneboo, A., de Groot, V., Moraal, B., Knol, D. L., Polman, C. H., et al. (2010). Early central atrophy rate predicts 5 year clinical outcome in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* 81, 1351–1356. doi: 10.1136/jnnp.2009.199968
- Martino, D., Joo, J. E., Sexton-Oates, A., Dang, T., Allen, K., Saffery, R., et al. (2014). Epigenome-wide association study reveals longitudinally stable dna methylation differences in cd4+ t cells from children with ige-mediated food allergy. *Epigenetics* 9, 998–1006. doi: 10.4161/epi.28945
- Martino, D., Loke, Y. J., Gordon, L., Ollikainen, M., Cruickshank, M. N., Saffery, R., et al. (2013). Longitudinal, genome-scale analysis of dna methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biol.* 14:R42. doi: 10.1186/gb-2013-14-5-r42
- Sepulcre, J., Sastre-Garriga, J., Cercignani, M., Ingle, G. T., Miller, D. H., and Thompson, A. J. (2006). Regional gray matter atrophy in early primary progressive multiple sclerosis: a voxel-based morphometry study. *Arch. Neurol.* 63:1175. doi: 10.1001/archneur.63.8.1175
- Shen, D., and Davatzikos, C. (2002). HAMMER: hierarchical attribute matching mechanism for elastic registration. *Med. Imaging IEEE Trans.* 21, 1421–1439. doi: 10.1109/TMI.2002.803111
- Shen, D., Resnick, S. M., and Davatzikos, C. (2003). “4D HAMMER image registration method for longitudinal study of brain changes,” in *Proceedings of the Human Brain Mapping* (New York, NY).
- Shen, D., Sundar, H., Xue, Z., Fan, Y., and Litt, H. (2005). “Consistent estimation of cardiac motions by 4D image registration,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005*, eds J. S. Duncan and G. Gerig (Berlin; Heidelberg: Springer), 902–910. doi: 10.1007/11566489_111
- Shiee, N., Bazin, P. L., Ozturk, A., Reich, D. S., Calabresi, P. A., and Pham, D. L. (2010). A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. *Neuroimage* 49, 1524–1535. doi: 10.1016/j.neuroimage.2009.09.005
- Simon, J. H., Jacobs, L. D., Campion, M. K., Rudick, R. A., Cookfair, D. L., Herndon, R. M., et al. (1999). A longitudinal study of brain atrophy in relapsing multiple sclerosis. *Neurology* 53:139. doi: 10.1212/WNL.53.1.139
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al. (2004). Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage* 23, S208–S219. doi: 10.1016/j.neuroimage.2004.07.051
- Tekok-Kilic, A., Benedict, R. H., Weinstock-Guttman, B., Dwyer, M. G., Carone, D., Srinivasaraghavan, B., et al. (2007). Independent contributions of cortical gray matter atrophy and ventricle enlargement for predicting neuropsychological impairment in multiple sclerosis. *Neuroimage* 36, 1294–1300. doi: 10.1016/j.neuroimage.2007.04.017
- Turner, B., Ramli, N., Blumhardt, L. D., and Jaspan, T. (2001). Ventricular enlargement in multiple sclerosis: a comparison of three-dimensional and linear mri estimates. *Neuroradiology* 43, 608–614. doi: 10.1007/s00234000457
- Worsley, K., Taylor, J., Carbonell, F., Chung, M., Duerden, E., Bernhardt, B., et al. (2009). Surfstat: a matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models

- and random field theory. *Neuroimage* 47:S102. doi: 10.1016/S1053-8119(09)70882-1
- Wu, G., Wang, Q., Jia, H., and Shen, D. (2010). "Registration of longitudinal image sequences with implicit template and spatial-temporal heuristics," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2010*, Vol. 6362, eds T. Jiang, N. Navab, J. P. W. Pluim, and M. A. Viergever (Berlin; Heidelberg: Springer), 618–625. doi: 10.1007/978-3-642-15745-5_76
- Zipunnikov, V., Caffo, B., Yousem, D., Davatzikos, C., Schwartz, B., and Crainiceanu, C. (2011a). Functional principal components model for high-dimensional brain imaging. *Neuroimage* 58, 772–784. doi: 10.1016/j.neuroimage.2011.05.085
- Zipunnikov, V., Caffo, B., Yousem, D. M., Davatzikos, C., Schwartz, B. S., and Crainiceanu, C. (2011b). Multilevel functional principal component analysis for high-dimensional data. *J. Comput. Graph. Stat.* 20, 852–873. doi: 10.1198/jcgs.2011.10122
- Zipunnikov, V., Greven, S., Shou, H., Caffo, B., Reich, D. S., Crainiceanu, C., et al. (2014). Longitudinal high-dimensional principal components analysis with application to diffusion tensor imaging of multiple sclerosis. *Ann. Appl. Stat.* 8, 2175–2202. doi: 10.1214/14-AOAS748
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Lee, Zipunnikov, Reich and Pham. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Reconstruction of human brain spontaneous activity based on frequency-pattern analysis of magnetoencephalography data

Rodolfo R. Llinás^{1*}, Mikhail N. Ustinin², Stanislav D. Rykunov², Anna I. Boyko², Vyacheslav V. Sychev², Kerry D. Walton¹, Guilherme M. Rabello¹ and John Garcia¹

¹ Department of Neuroscience and Physiology, New York University School of Medicine, New York, NY, USA, ² Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Russia

OPEN ACCESS

Edited by:

Jian Kang,
Emory University, USA

Reviewed by:

Seonjoo Lee,
Columbia University and New York
State Psychiatric Institute, USA
Jing Zhang,
Georgia State University, USA

*Correspondence:

Rodolfo R. Llinás
rodolfo.llinas@nyumc.org

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 09 June 2015

Accepted: 28 September 2015

Published: 16 October 2015

Citation:

Llinás RR, Ustinin MN, Rykunov SD,
Boyko AI, Sychev VV, Walton KD,
Rabello GM and Garcia J (2015)
Reconstruction of human brain
spontaneous activity based on
frequency-pattern analysis of
magnetoencephalography data.
Front. Neurosci. 9:373.
doi: 10.3389/fnins.2015.00373

A new method for the analysis and localization of brain activity has been developed, based on multichannel magnetic field recordings, over minutes, superimposed on the MRI of the individual. Here, a high resolution Fourier Transform is obtained over the entire recording period, leading to a detailed multi-frequency spectrum. Further analysis implements a total decomposition of the frequency components into functionally invariant entities, each having an invariant field pattern localizable in recording space. The method, addressed as functional tomography, makes it possible to find the distribution of magnetic field sources in space. Here, the method is applied to the analysis of simulated data, to oscillating signals activating a physical current dipoles phantom, and to recordings of spontaneous brain activity in 10 healthy adults. In the analysis of simulated data, 61 dipoles are localized with 0.7 mm precision. Concerning the physical phantom the method is able to localize three simultaneously activated current dipoles with 1 mm precision. Spatial resolution 3 mm was attained when localizing spontaneous alpha rhythm activity in 10 healthy adults, where the alpha peak was specified for each subject individually. Co-registration of the functional tomograms with each subject's head MRI localized alpha range activity to the occipital and/or posterior parietal brain region. This is the first application of this new functional tomography to human brain activity. The method successfully provides an overall view of brain electrical activity, a detailed spectral description and, combined with MRI, the localization of sources in anatomical brain space.

Keywords: magnetic encephalography, frequency-pattern analysis, functional tomography, phantom data, alpha rhythm, inverse problem solution

INTRODUCTION

Magnetoencephalography (MEG) has become one of the foremost biological technologies addressing detailed analysis of human brain function and recently an open archive has been established (Niso et al., 2015). Thus, recorded magnetic fields with a high sampling rate, and hundreds of recording channels, can provide a functional image of unprecedented precision, comprising cortical as well as deep brain structures. Due to its methodological character, this

approach can analyze large data sets affording the comprehensive analysis of functional detail. Concerning the human brain, two main parameters have challenged global analysis of function. One is the simultaneous nature of brain neuronal activity where, at any given instant, millions of neuronal functional events co-exist. The other is the great variety in neuronal morphologies that, upon activation, lead to the generation of different electromagnetic field profiles (Llinás, 1988). Historically, the most common approach to such a conundrum has been to address the brain activity that follows a given stimuli (evoked or induced potentials), or to address the characteristics of spontaneous (resting or ongoing) electromagnetic activity related to large events such as various sleep and waking states (Llinás and Pare, 1991).

Historically, the former approaches, i.e., the analysis of sensory evoked potentials, or those obtained from abnormal brain function relating to synchronous activation of vast number of neurons (e.g., epilepsy, Ossenblok et al., 2007) has been the most commonly addressed. However, the final results, even under such favorable conditions as the analysis of repeated simple stimuli that may be averaged, have not yielded the imaging required to address dynamic brain functional activity and this remains a field of active research (David et al., 2006a,b; Klimesch et al., 2007; Sauseng et al., 2007; Ros et al., 2015). Under those conditions, the content of moment-to-moment brain function is lost and only those aspects that relate to the common features of the given repeated stimuli are addressable.

In an attempt to move away from the evoked activity approach and toward the analysis of ongoing brain activity a new method has been developed to represent global brain activity as a set of elementary coherent oscillations (Llinás and Ustinin, 2014a,b). The core of the proposed technology lies in the performing of the precise detailed Fourier transform of the long multichannel time series and in the analysis of the frequency components obtained. Theorems were proved, stating that if phases are equal in all channels of some elementary oscillation (characterized by distinct frequency), then the normalized pattern of this oscillation is constant through the period. Mathematically it means separation of variables (time course and spatial structure are simply multiplied). It means, that such elementary coherent oscillation ideally cleans the spatial structure of the field at this particular frequency. This approach was applied to 19 experimental MEG data sets of human spontaneous activity, and it was found, that many elementary oscillations reveal high coherence and hence are representing static structures, generating corresponding frequencies. The next step was to further divide those oscillations, which are not coherent, but still look rather simple, because of the detailed frequency representation. As a result, the multichannel signal is decomposed into the set of elementary coherent oscillations. Note, that this decomposition is obtained by the direct nonparametric transformation of the initial data, it is precise and totally reversible. The solution of the inverse problem for each elementary oscillation provides the spatial structure of the source, oscillating as a whole with time course, extracted earlier. When inverse problems will be solved for all oscillations, the system under study will be represented as a sum of stable

sources (functional entities), each of them oscillating as a whole. Many methods of inverse problem solution were developed (e.g., Hämäläinen and Ilmoniemi, 1994; Sekihara and Nagarajan, 2008; Kozunov and Ossadtchi, 2015). Some of these methods, especially those devised for simple source models, can be effectively used to reconstruct functional entities, extracted by the proposed technology. The fact, that proposed technology splits MEG into elementary oscillations with relatively simple patterns, can revive few-channel measurements, including those combined with MRI (Zotев et al., 2008; Cottureau et al., 2015; Fukushima et al., 2015).

Here we assume, that the considerable part of the MEG signal in the alpha rhythm frequency band can be represented as a sum of equivalent current dipoles, while each coherent oscillation is described by one dipole. In order to check this assumption, the following experiments were performed. Computer simulation (61 dipoles plus noise) and physical modeling (3 dipoles) were used as a benchmark, estimating initial data with good precision. Then the method was blindly applied to study the alpha rhythm in 10 human subjects, localizing ~2000 dipoles for each person. The alpha rhythm is often used as a benchmark for different methods (Pascual-Marqui et al., 2014), because of the relatively well-known nature of this phenomenon. The blind application of the method means here, that no anatomical information about the brain is used to solve the inverse problem. When initial MEG is split into the set of elementary coherent oscillations, inverse problem is solved for every oscillation pattern in one current dipole model and the energy of oscillation is attributed to the spatial position of the dipole. Repeating this procedure for all oscillations, one can obtain the Functional Tomogram (FT)—spatial distribution of the sources, generating the initial MEG. The allowable localization space is a $25 \times 25 \times 25 \text{ cm}^3$ cube, and it is the geometrical size of the FT.

In order to better evaluate the method, functional tomogram is compared with individual brain anatomy only after the end of computations. It can be schematically shown as:

MEG registration → Calculation of the FT → Representation of the FT with MRI ← MRI registration

Here both MRI and MEG data are obtained from the same subject, using fiducial markers.

Biologically interpretable results of localization, obtained under this condition, point at the possible applicability of the proposed method in the studies of the brain ongoing activity.

METHODS

Computer Simulation

The MEG-data were simulated using 61 current dipoles, randomly distributed in space $8 \times 8 \times 8 \text{ cm}^3$. The forward problem for the dipole in spherical conductor was solved, generating sinusoidal signal. The time of simulation was equal to 1 min with sampling frequency 1200 Hz. Frequencies changed from 9.5 to 10.5 Hz with the step 0.0167 Hz. Amplitudes were randomly distributed from 10 to 100 fT, corresponding to experimental results for humans in this frequency band. Parameters of the gradiometer for simulation were taken from experimental “noise” collection data. This was obtained by

making a 1-min recording under the same conditions as during a MEG recording from a subject, in the absence of the subject (sampling frequency 1200 Hz) in order to estimate the level of noise. The sum of 61 model MEGs was calculated, and the estimated noise MEG was added to account for the noise. Resulting magnetoencephalogram and its multichannel spectrum qualitatively correspond to experimental data for humans in the alpha frequency band.

Phantom

A current dipole phantom (CTF Systems) was used. This phantom consists of a spherical saline-filled vessel, 13 cm inner diameter, providing an appropriate current flow conductor. Inside this vessel, several current dipoles were installed. Each dipole comprises two gold spheres 2 mm in diameter with a 9 mm center-to-center separation.

Subjects and Data Acquisition

MEG recordings were acquired from 10 healthy adults (5 men and 5 women) aged 28–76 years of age (mean 41.8 ± 5.4 years; median age 33.5 years). This study was carried out with the approval of the New York University School of Medicine Institutional Review Board. All subjects gave written informed consent in accordance with the Declaration of Helsinki. Participants were recruited from the New York University Medical Center and the local community. MEG recordings were implemented at the New York University School of Medicine Center for Neuromagnetism (CNM) located at the Bellevue Hospital Center. The subjects were asked to relax but stay awake during each 7-min recording period in 42 10-s trials. Recordings were made during both “eyes closed” (EC) and “eyes open” (EO) conditions. Three fiducial markers were applied (left and right preauricular points, and the nasion) to localize the head during the MEG recording.

MEG recordings were carried out in a mu-metal magnetically shielded room using a 275-channel instrument (CTF Systems) while the subject sat upright (sample rate 600 or 1200 Hz). Recordings were from 275 channels. Artifacts and distant noise were reduced using a 3rd order gradientometer (McCubbin et al., 2004). The activity of the instrument and distant noise were recorded before each session.

Data Analysis

The MEG instrumentation supports simultaneous multichannel recordings of magnetic fields from brain activity generated at discrete time points, thus providing sets of discrete experimental vectors $\{b_k\}$. Instantaneous field value $b_k(i)$ is registered at the time moment τ_i , $i = 1, \dots, L$, $\tau_1 = 0$. The first step in our methodology is the interpolation of the experimental data in every channel (Boyd, 2001):

$$\tilde{b}_k(t) = \frac{(t - \tau_{i+1})}{(\tau_i - \tau_{i+1})} b_k(i) + \frac{(t - \tau_i)}{(\tau_{i+1} - \tau_i)} b_k(i+1), \quad t \in [\tau_i, \tau_{i+1}],$$

$$i = 1, \dots, L-1, k = 1, \dots, K. \quad (1)$$

Interpolation provides the continuous function $\tilde{b}_k(t)$, $t \in [0, T]$, $T = \tau_L - \tau_1$, where T is the time of measurement, k is the channel number.

The multichannel Fourier transform calculates a set of spectra for interpolated functions $\{\tilde{b}_k(t)\}$:

$$a_{0k} = \frac{2}{T} \int_0^T \tilde{b}_k(t) dt, \quad a_{nk} = \frac{2}{T} \int_0^T \tilde{b}_k \cos(2\pi \nu_n t) dt,$$

$$b_{nk} = \frac{2}{T} \int_0^T \tilde{b}_k \sin(2\pi \nu_n t) dt, \quad (2)$$

Where a_{0k} , a_{nk} , b_{nk} are Fourier coefficients for the frequency ν_n in the channel number k , and $\nu_n = \frac{n}{T}$, $n = 1, \dots, N$, $N = \nu_{\max} T$, where ν_{\max} is the highest desirable frequency. The coefficient a_{0k} will not be considered hereafter, because the constant field component has no meaning in superconducting quantum interference device (SQUID) measurements.

All spectra are calculated for the whole registration time T , which is sufficient to reveal the detailed frequency structure of the system. The step in frequency is equal to $\Delta \nu = \nu_n - \nu_{n-1} = \frac{1}{T}$, thus frequency resolution is determined by the recording time. Gaussian quadrature formulas are used to calculate integrals on any interval $[0, T]$, so the interpolation (1) makes it possible to optimize frequency grid, changing T (Llinás and Ustinin, 2014a,b). If the optimization is not necessary, and the time array τ provides quadrature nodes to calculate integrals with sufficient precision, then the data are used without interpolation. In this study integrals were calculated without interpolation.

Given a precise multichannel spectrum, it is possible to perform the inverse Fourier transform:

$$B_k(t) = \sum_{n=1}^N \rho_{nk} \sin(2\pi \nu_n t + \varphi_{nk}), \quad \nu_n = \frac{n}{T}, \quad N = \nu_{\max} T, \quad (3)$$

Where $\rho_{nk} = \sqrt{a_{nk}^2 + b_{nk}^2}$, $\varphi_{nk} = \text{atan2}(a_{nk}, b_{nk})$, and a_{nk} , b_{nk} are Fourier coefficients, found in (2).

Precision of the direct and inverse Fourier transforms, used in our approach, can be illustrated by the fact, that initial MEG is restored from (3) with relative error less than 10^{-20} .

In order to study the detailed frequency structure of the brain, we restore multichannel signal at every frequency and analyze the functions obtained. Multichannel signal is restored at frequency ν_n in all channels:

$$B_{nk}(t) = \rho_{nk} \sin(2\pi \nu_n t + \varphi_{nk}), \quad (4)$$

where $t \in [0, T_{\nu_n}]$, $T_{\nu_n} = \frac{1}{\nu_n}$ is the period of this frequency.

If $\varphi_{nk} = \varphi_n$, then formula (4) describes a coherent multichannel oscillation and can be written as:

$$B_{nk}(t) = \rho_{nk} \sin(2\pi \nu_n t + \varphi_n) = \hat{\rho}_{nk} \rho_n \sin(2\pi \nu_n t + \varphi_n), \quad (5)$$

where $\rho_n = \sqrt{\sum_{k=1}^K \rho_{nk}^2}$ is the amplitude, and $\hat{\rho}_{nk} = \frac{\rho_{nk}}{\rho_n}$ is the normalized pattern of oscillation.

In multichannel measurements space is determined by positions of channels. If time course does not depend on k , we have separation of time and space variables.

The normalized pattern makes it possible to determine the spatial structure of the source from the inverse problem solution,

and this structure is constant throughout the entire period of the oscillation. The time course of the field is determined by the function $\rho_n \sin(2\pi \nu_n t + \varphi_n)$, which is common for all channels, i.e., this source is oscillating as a whole at the frequency ν_n .

The theoretical foundations for the reconstruction of static functional entities (neural circuits, or sources) have been developed (Llinás and Ustinin, 2014a,b). This reconstruction is based on detailed frequency analysis and extraction of the frequencies, having high coherence and similar patterns.

The algorithm of mass precise frequency-pattern analysis was formulated as:

- (1) Precise Fourier Transform of the multichannel signal.
- (2) Inverse Fourier Transform—restoration of the signal at each frequency.
- (3) If the coherence at the particular frequency is close to 1, then use the pattern and frequency as elementary coherent oscillation, see Equation (5).
- (4) If the restored signal consists of several phase-shifted coherent oscillations, then extract those oscillations:
 - (a) Apply second order blind identification (SOBI) algorithm (Belouchrani et al., 1997) to restored time-series in Equation (4);
 - (b) Select nonzero components;
 - (c) Apply direct Fourier transform to each selected component and calculate amplitude, normalized pattern and phase using Equation (5).

After the fourth step of this analysis, the initial multichannel signal is represented as a sum of elementary coherent oscillations:

$$B_k(t) \cong \sum_{n=1}^N \sum_{m=1}^M D_{mn} \hat{\rho}_{mnk} \sin(2\pi \nu_n t + \varphi_{mn}),$$

$$\nu_n = \frac{n}{T}, N = \nu_{\max} T, m = 1, \dots, M, \quad (6)$$

where M is maximal number of coherent oscillations, extracted at the frequency ν_n .

Each elementary oscillation is characterized by frequency ν_n , phase φ_{mn} , amplitude D_{mn} , normalized pattern $\hat{\rho}_{mnk}$ and is produced by the functional entity having a constant spatial structure.

The method of functional tomography reconstructs the structure of the system from the analysis of the set of normalized patterns $\{\hat{\rho}_{mn}\}$.

The functional tomogram displays a 3-dimensional map of the energy produced by all the sources located at a given point. In order to build a functional tomogram, the space under study is divided into $N_x \times N_y \times N_z$ elementary cubicles with centers in \mathbf{r}_{ijs} . The edge of the cubicle is selected in accordance with desirable precision and/or computational facilities; in this study, it was 1.0 mm for simulated data, 1.5 mm for phantom data, and 3.0 mm for human data. To calculate the energy produced by all the sources located in the center of the cubicle, the set of L trial dipoles \mathbf{Q}_{ijsl} is build. The magnetic induction, produced by the trial dipole \mathbf{Q}_{ijsl} located in \mathbf{r}_{ijs} , is registered by the probe number k with position \mathbf{r}_k and direction \mathbf{n}_k . The k -th component ρ_{ijslk}^{tr}

of the trial pattern $ijsl$ is calculated from the model of a current dipole in a spherical conductor (Sarvas, 1987):

$$\rho_{ijslk}^{tr} = \frac{\mu_0}{4\pi F^2} (((\mathbf{Q}_{ijsl} \times \mathbf{r}_{ijs}) F - (\mathbf{Q}_{ijsl} \times \mathbf{r}_{ijs}, \mathbf{r}_k) \nabla F), \mathbf{n}_k), \quad (7)$$

where $F = a(ar_k + r_k^2 - (\mathbf{r}_{ijs}, \mathbf{r}_k))$, $\nabla F = (a^2 r_k^{-1} + a^{-1}(\mathbf{a}, \mathbf{r}_k) + 2a + 2r_k)\mathbf{r}_k - (a + 2r_k + a^{-1}(\mathbf{a}, \mathbf{r}_k))\mathbf{r}_{ijs}$, $\mathbf{a} = \mathbf{r}_k - \mathbf{r}_{ijs}$, $a = |\mathbf{a}|$, $r_k = |\mathbf{r}_k|$, $|\mathbf{n}_k| = 1$, $\mu_0 = 4\pi \cdot 10^{-7}$. Full set of ρ_{ijslk}^{tr} provides lead field matrix for the particular device (Hämäläinen and Ilmoniemi, 1994; Sekihara and Nagarajan, 2008).

The normalized trial pattern is calculated as:

$$\hat{\rho}_{ijslk}^{tr} = \frac{\rho_{ijslk}^{tr}}{|\rho_{ijsl}^{tr}|}, \text{ where } |\rho_{ijsl}^{tr}| = \sqrt{\sum_{k=1}^K (\rho_{ijslk}^{tr})^2}. \quad (8)$$

All trial dipoles, originating from \mathbf{r}_{ijs} , lie in the same plane, orthogonal to \mathbf{r}_{ijs} , because the vector product $\mathbf{Q}_{ijsl} \times \mathbf{r}_{ijs}$ is nonzero only for those dipoles. Trial dipoles cover the circle in L_{\max} directions with $360/L_{\max}$ degrees step, in this study $L_{\max} = 8$.

The set of normalized trial patterns is then calculated, using (8) for each trial dipole:

$$\{\hat{\rho}_{ijsl}^{tr}\}, i = 1, \dots, N_x; j = 1, \dots, N_y; s = 1, \dots, N_z; l = 1, \dots, L_{\max}. \quad (9)$$

In this study more than 2.5 million trial patterns were used for each person. Those patterns were produced by trial dipoles, uniformly distributed in the localization space.

For each normalized pattern $\hat{\rho}_{mn}$, the following function was calculated, giving the difference between this pattern and one of the trial patterns:

$$\chi(i, j, s, l) = \sum_{k=1}^K (\hat{\rho}_{ijslk}^{tr} - \hat{\rho}_{mnk})^2, \quad (10)$$

where $\hat{\rho}_{ijslk}^{tr}$ is a k -th component of the trial pattern jsl and $\hat{\rho}_{mnk}$ is a k -th component of the normalized pattern mn , k —number of channel.

The position and direction of the source producing the pattern $\hat{\rho}_{mn}$ were determined by numbers (I, J, S, L) , providing the minimum to the function $\chi(i, j, s, l)$ over the variables $i = 1, \dots, N_x; j = 1, \dots, N_y; s = 1, \dots, N_z; l = 1, \dots, L_{\max}$. The minimum of this function was found by the exhaustive search, selecting the smallest value from the whole set of 2.5 millions χ for each $\hat{\rho}_{mn}$. Such procedure determines \mathbf{r}_{IJS} —the inverse problem solution for the pattern $\hat{\rho}_{mn}$, without filtering of channels, or weighting functions.

The energy of this source D_{mn}^2 is added to the energy produced from the cubicle with the center at \mathbf{r}_{IJS} .

Performing this procedure for all normalized patterns: $m = 1, \dots, M; n = 1, \dots, N$, it is possible to distribute in space the energy of all oscillations from formula (6). The result of such distribution is the Functional Tomogram of the brain, reconstructed from MEG.

EXPERIMENTAL RESULTS

Computer Simulation

The simulated MEG was analyzed by the method proposed in Section Data Analysis. The functional tomogram yielded a 3-dimensional map of energy in the frequency band 9.5–10.5 Hz, distributed in a $8 \times 8 \times 8 \text{ cm}^3$ cube (in empty space) with a 1.0 mm resolution. For each frequency, the calculated functional tomogram was compared with the coordinates of simulated current dipole. The average distance between the dipole true position and the center of the elementary

cubicle to which this dipole was localized, was estimated as $0.7 \pm 0.1 \text{ mm}$.

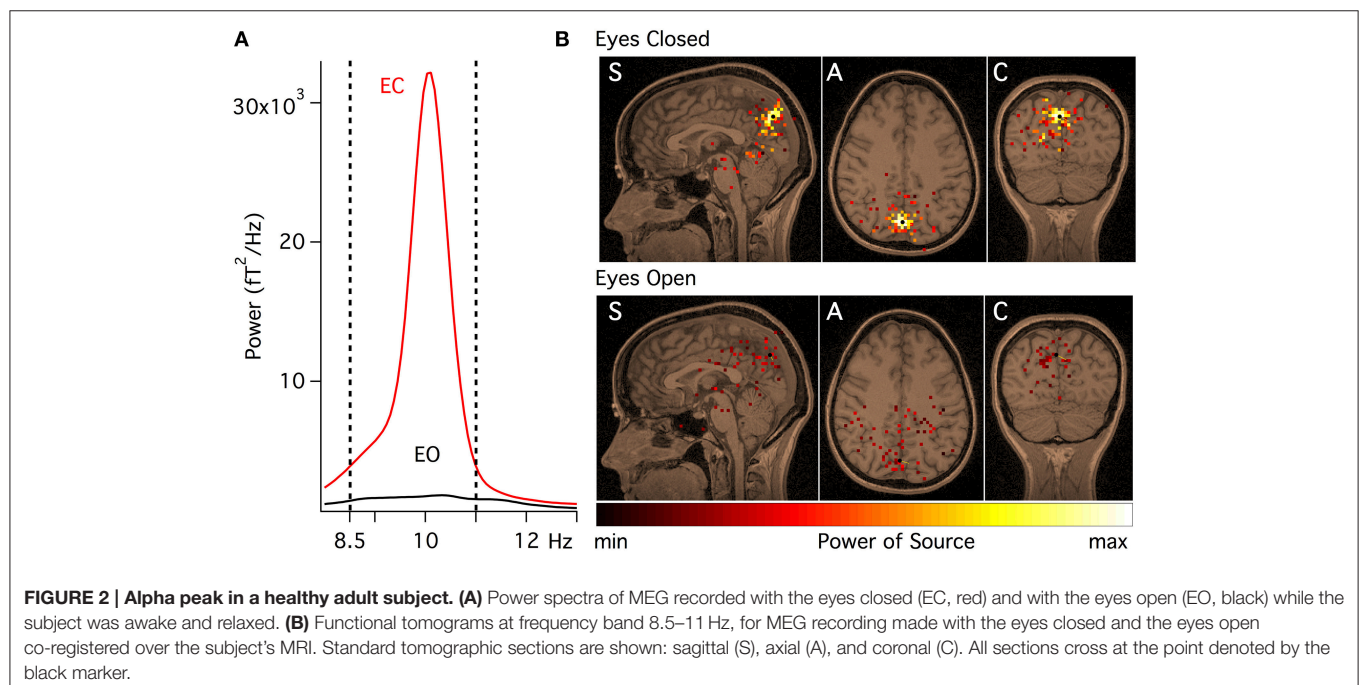
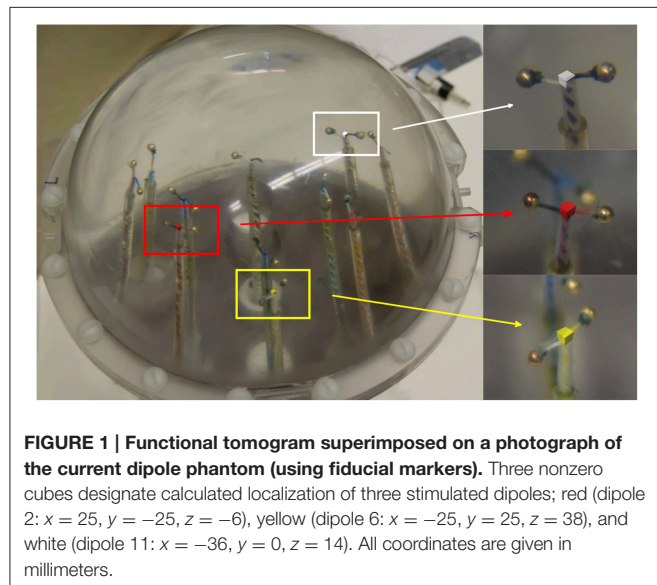
Phantom

The phantom was placed in the center of the MEG recording helmet. Three localization coils were placed on the spherical vessel, corresponding to usual head placement (front, left and right side, separated by 90° on the sphere's equator) thus providing the necessary fiducial markers. Three dipoles were activated simultaneously with alternate current from separate generators, at 7.00, 7.83, and 11.00 Hz. The magnetic field produced by the phantom dipoles was recorded for 100 s.

The functional tomogram, calculated as described in Section Methods, yielded a 3-dimensional map of energy in the frequency band 1–40 Hz, distributed in a $10 \times 10 \times 10 \text{ cm}^3$ cube (in empty space) with a 1.5 mm resolution. Then the calculated functional tomogram was superimposed on a photograph of the phantom (white, red and yellow cubes $1.5 \times 1.5 \times 1.5 \text{ mm}^3$ in **Figure 1**). All cubes were localized to the centers of the phantom dipoles with an error of less than 1 mm.

The Alpha Rhythm

The current method makes it possible to study spontaneous brain resting activity, and to analyze the distribution of sources in the brain. The alpha rhythm was selected for this study since it is the dominant oscillation in healthy adults when the eyes are closed (see Basar, 2012, for a review). In broad terms, the alpha band has been defined as 8–13 Hz brain generated rhythm having, typically, a 9–11 Hz frequency in healthy adults (Nunez et al., 2001). To eliminate differences in the alpha peak across individuals, the individual alpha frequency (IAF) (Klimesch et al., 1999) was determined for each subject.



Let us consider the processing of experimental data set for one subject (#4 in **Figure 3**). Two multichannel spectra were calculated in the frequency band 8–13 Hz, each spectrum contains 2100 frequency peaks in 275 channels:

$$B_k(t) = \sum_{n=n_{\min}}^{n_{\max}} \rho_{nk} \sin(2\pi v_n t + \varphi_{nk}), \quad v_n = \frac{n}{T}, \quad n_{\min} = 3361, \\ n_{\max} = 5460, \quad k = 1, \dots, 275 \quad (11)$$

Figure 2A shows the power spectra (calculated using the Welch method) (Welch, 1967) for subject #4 recorded with the eyes open (EO) and the eyes closed (EC). It can be concluded, that those two states demonstrate different spectral features, namely, the spectrum for the EC condition contains a peak near 10 Hz that decreased when the recording was made with subject's eyes open.

From the analysis of **Figure 2A**, frequency band 8.5–11 Hz was selected for further analysis, as characteristic for EC condition. Source localizations for this band, recorded with the

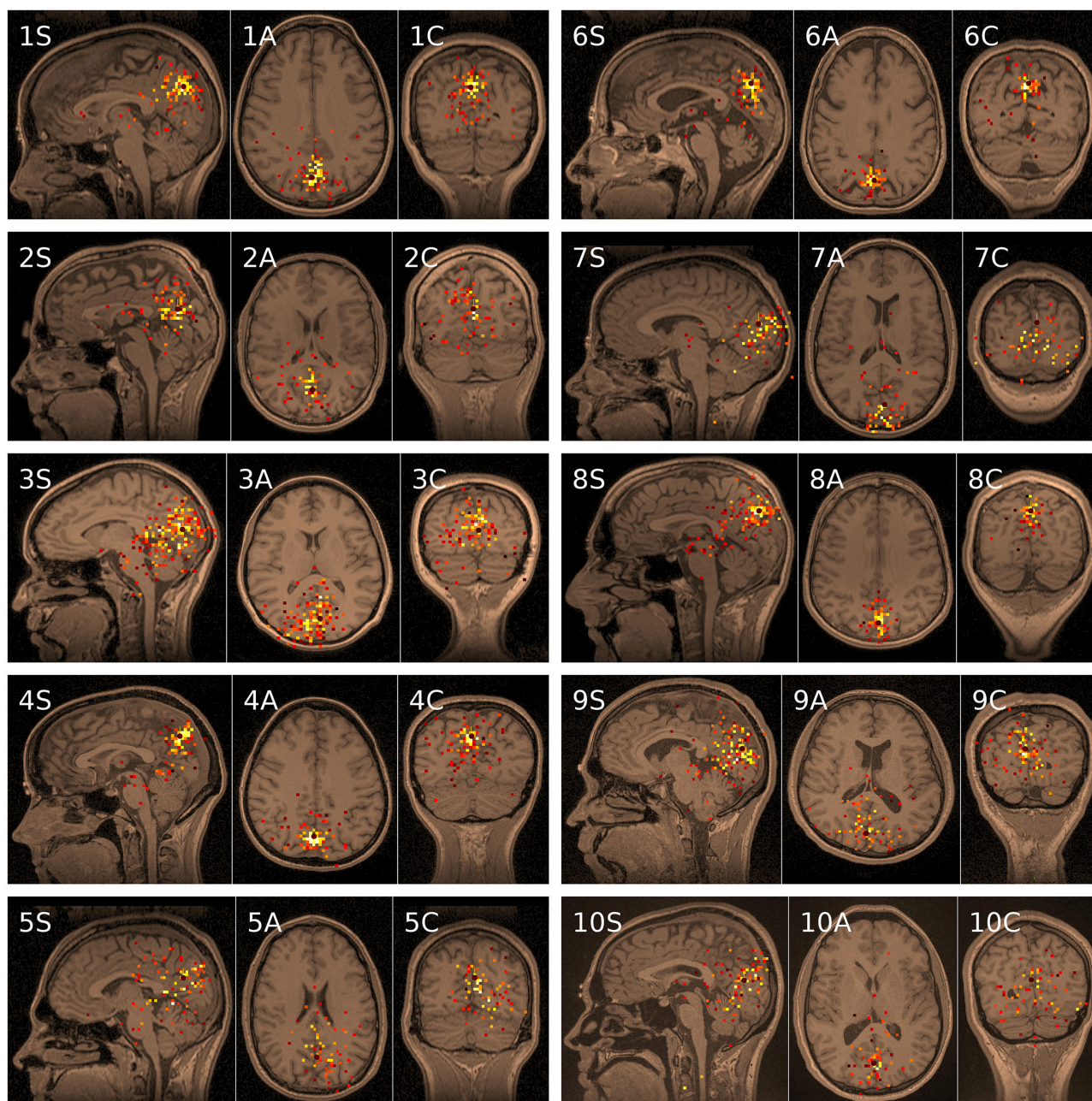


FIGURE 3 | Functional tomograms of alpha band spontaneous activity co-registered with MRIs for 10 subjects recorded with the eyes closed. Each tomogram shows three standard tomographic sections S (sagittal), axial (A), and C (coronal).

EC and the EO, are shown in **Figure 2B** co-registered on the subject's MRI. The total energy in the alpha frequency band recorded with the EC was much stronger and was concentrated in a smaller volume, than the corresponding spectral energy generated when the eyes were open during the recording, as would be expected in healthy adults (Nunez et al., 2001).

The same data processing protocol was applied and similar results were obtained for all 10 subjects. **Figure 3** shows 10 functional tomograms with corresponding MRIs for the EC condition. For each subject three tomographic sections (**Figure 4**, S, sagittal; A, axial; C, coronal) are shown. The sections transect the same point in space (black marker) located in the region of the strongest source. Such sources were denoted by white voxels in the functional tomogram, in accordance with legend for **Figure 2B**. Presentation of the data was performed in the program environment MEGMRIAn (Ustinin et al., 2014). The spatial resolution for the MRI is equal to 1 mm, for the functional tomogram it is 3 mm. Eight directions were used for trial dipoles in every point of the spatial grid, as explained in Section Subjects and Data Acquisition, Equations (7) and (8).

Figure 4 shows the superposition of the 10 functional tomograms shown on the MRI from subject #5 in **Figure 3**. This summation was performed in the head coordinate system, common for all functional tomograms. Note, that regardless of individual variances, the alpha rhythm energy distribution displays general tendency to be located in occipital and posterior parietal lobes.

Resolution of the Method

There are two kinds of resolution in this approach: frequency resolution and space resolution. Frequency resolution $\Delta\nu = \nu_n - \nu_{n-1} = \frac{1}{T}$ is determined by the time of measurement, on condition that Fourier integrals for the full time of measurement (2) are calculated precisely. It is reflecting the fundamental fact: the longer one registers time series, the better one determines frequency structure of the system. In this study of spontaneous activity T was equal to 420 s, thus providing 420 frequencies per one Hz.

Spatial resolution has no theoretical limitations in this approach. Note, that the functional tomograms were calculated with spatial resolution of 1.0 mm for simulated data, 1.5 mm for

phantom data, and 3.0 mm for human data. These differences were determined by computational limitations and followed from the usage of a space of $8 \times 8 \times 8 \text{ cm}^3$ for simulated data, of $10 \times 10 \times 10 \text{ cm}^3$ for phantom and of $25 \times 25 \times 25 \text{ cm}^3$ for human functional tomograms. By increasing computer memory, one can obtain a higher spatial resolution. Precision of localization can be estimated from the known dipoles positions in cases of simulated and physical dipoles. It was found, that precision ≈ 0.7 of resolution.

For each elementary coherent oscillation, found in (6), unique dipolar source is localized by selection of the best trial source from 2.5 million, distributed in the whole space of MRI. It means that no a-priori limitations are used for the location of sources, and their combined representation with MRI may provide new information. Using normalized patterns, one can obtain localization of weak sources, if they are extracted from Fourier analysis, with precision equal to the precision of localization of strong sources. It opens new possibilities to study deep brain sources.

Two and more oscillations can have common position and direction, thus providing the spectrum of the particular source (partial spectrum) (see Llinás and Ustinin, 2014a,b). The inverse Fourier transform gives time series, produced by this source. Selecting two or more such sources, one can study their connectivity, using methods described in Greenblatt et al. (2012).

DISCUSSION

A novel method to implement the analysis of human brain activity addressed as functional tomography is introduced. This novel methodology was used to calculate the spatial distribution of brain activity power sources recorded with an MEG instrument. This method is free of arbitrary parameters, it is computationally stable, and it is free from matrix inversion requirements. Computational demands are reasonable for modern computers. Thus, a functional tomogram may be implemented in 20 min on a computer with 2.4 GHz 4-cores Haswell CPU and 16 GB RAM.

Functional tomograms were obtained for alpha rhythm from multichannel MEG data. These functional tomograms

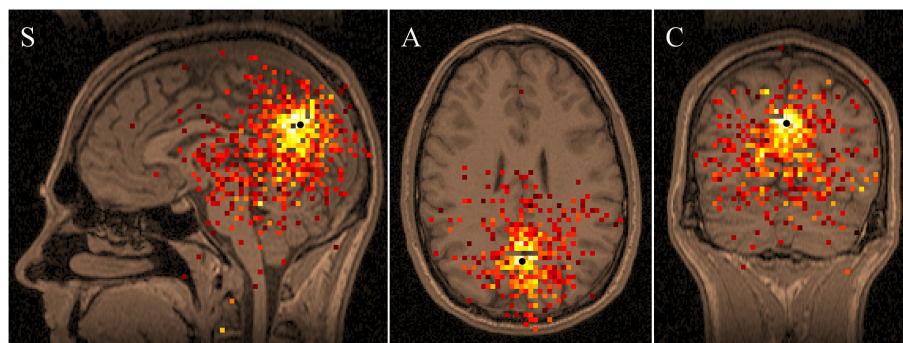


FIGURE 4 | Average “Eyes Closed” functional tomogram of alpha band activity recorded from 10 individuals, plotted over the MRI from subject #5. Sagittal (S), axial (A), and coronal (C) tomographic sections are shown.

demonstrate individual variances of the power spatial distribution, generally corresponding to our present knowledge concerning the alpha rhythm localization in the occipital and posterior parietal lobes (Nunez et al., 2001; Basar, 2012). It can be concluded, therefore, that the functional tomography method, based on magnetic-encephalograms analysis, can determine spontaneous brain activity sources.

A fundamental advantage of this framework lies in the fact, that all recorded data is fully utilized.

Method of functional tomography can be applied to the diagnostics of activity in the whole brain and in broad frequency

band, revealing areas of abnormally high or abnormally low activity.

ACKNOWLEDGMENTS

Authors are grateful to engineer Alex Porras PhD. The study was partly supported by the CRDF Global (USA) (grants CRDF RB1-2027 and RUB-7095-MO-13), by the Russian Foundation for Basic Research (grants 13-07-00162, 13-07-12183, 14-07-00636, 14-07-31309), and by the Program 43 for Fundamental Research of the Russian Academy of Sciences.

REFERENCES

- Basar, E. (2012). A review of alpha activity in integrative brain function: fundamental physiology, sensory coding, cognition and pathology. *Int. J. Psychophysiol.* 86, 1–24. doi: 10.1016/j.ijpsycho.2012.07.002
- Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F., and Moulines, E. (1997). A blind source separation technique using second-order statistics. *IEEE Trans. Signal Processing* 45, 434–444. doi: 10.1109/78.554307
- Boyd, J. P. (2001). *Chebyshev and Fourier Spectram Methods*. New York, NY: Dover.
- Cottareau, B. R., Ales, J. M., and Norgia, A. M. (2015). How to use fMRI functional localizers to improve EEG/MEG source estimation. *J. Neurosci. Methods* 250, 64–73. doi: 10.1016/j.jneumeth.2014.07.015
- David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., and Friston, K. J. (2006a). Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage* 30, 1255–1272. doi: 10.1016/j.neuroimage.2005.10.045
- David, O., Kilner, J. M., and Friston, K. J. (2006b). Mechanisms of evoked and induced responses in MEG/EEG. *Neuroimage* 31, 1580–1591. doi: 10.1016/j.neuroimage.2006.02.034
- Fukushima, M., Yamashita, O., Knösche, T. R., and Sato, M. A. (2015). MEG source reconstruction based on identification of directed source interactions on whole-brain anatomical networks. *Neuroimage* 105, 408–427. doi: 10.1016/j.neuroimage.2014.09.066
- Greenblatt, R. E., Pflieger, M. E., and Ossadtchi, A. E. (2012). Connectivity measures applied to human brain electrophysiological data. *J. Neurosci. Methods* 207, 1–16. doi: 10.1016/j.jneumeth.2012.02.025
- Hämäläinen, M. S., and Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42. doi: 10.1007/BF02512476
- Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., and Winkler, T. (1999). 'Paradoxical' alpha synchronization in a memory task. *Brain Res. Cogn. Brain Res.* 7, 493–501. doi: 10.1016/S0926-6410(98)00056-1
- Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., and Freunberger, R. (2007). Event-related phase reorganization may explain evoked neural dynamics. *Neurosci. Biobehav. Rev.* 31, 1003–1016. doi: 10.1016/j.neubiorev.2007.03.005
- Kozunov, V. V., and Ossadtchi, A. (2015). GALA: group analysis leads to accuracy, a novel approach for solving the inverse problem in exploratory analysis of group MEG recordings. *Front. Neurosci.* 9:107. doi: 10.3389/fnins.2015.00107
- Llinás, R. R. (1988). The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* 242, 1654–1664. doi: 10.1126/science.3059497
- Llinás, R. R., and Paré, D. (1991). Of dreaming and wakefulness. *Neuroscience* 44, 521–535. doi: 10.1016/0306-4522(91)90075-Y
- Llinás, R. R., and Ustinin, M. N. (2014a). Frequency-pattern functional tomography of magnetoencephalography data allows new approach to the study of human brain organization. *Front. Neural Circuits* 8:43. doi: 10.3389/fncir.2014.00043
- Llinás, R. R., and Ustinin, M. N. (2014b). *Precise Frequency-Pattern Analysis to Decompose Complex Systems into Functionally Invariant Entities*. USA patent application US20140107979 A1, New York, NY.
- McCubbin, J., Vrba, J., Spear, P., McKenzie, D., Willis, R., Loewen, R., et al. (2004). Advanced electronics for the CTF MEG system. *Neurol. Clin. Neurophysiol.* 2004:69.
- Niso, G., Rogers, C., Moreau, J. T., Chen, L. Y., Madjar, C., Das, S., et al. (2015). OMEGA: the open MEG archive. *Neuroimage*. doi: 10.1016/j.neuroimage.2015.04.028. [Epub ahead of print].
- Nunez, P. L., Wingeier, B. M., and Silberstein, R. B. (2001). Spatial-temporal structures of human alpha rhythms: theory, microcurrent sources, multiscale measurements, and global binding of local networks. *Hum. Brain Mapp.* 13, 125–164. doi: 10.1002/hbm.1030
- Ossenblok, P., de Munck, J. C., Colon, A., Drolsbach, W., and Boon, P. (2007). Magnetoencephalography is more successful for screening and localizing frontal lobe epilepsy than electroencephalography. *Epilepsia* 48, 2139–2149. doi: 10.1111/j.1528-1167.2007.01223.x
- Pascual-Marqui, R. D., Biscay, R. J., Bosch-Bayard, J., Lehmann, D., Kochi, K., Kinoshita, T., et al. (2014). Assessing direct paths of intracortical causal information flow of oscillatory activity with the isolated effective coherence (iCoh). *Front. Hum. Neurosci.* 8:448. doi: 10.3389/fnhum.2014.00448
- Ros, B. P., Bijma, F., de Gunst, M. C., and de Munck, J. C. (2015). A three domain covariance framework for EEG/MEG data. *Neuroimage* 119, 305–315. doi: 10.1016/j.neuroimage.2015.06.020
- Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.* 32, 11–22. doi: 10.1088/0031-9155/32/1/004
- Sauseng, P., Klimesch, W., Gruber, W. R., Hanslmayr, S., Freunberger, R., and Doppelmayr, M. (2007). Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience* 146, 1435–1444. doi: 10.1016/j.neuroscience.2007.03.014
- Sekihara, K., and Nagarajan, S. S. (2008). *Adaptive Spatial Filters for Electromagnetic Brain Imaging*. Berlin-Heidelberg: Springer-Verlag.
- Ustinin, M. N., Sychev, V. V., Walton, K. D., and Llinas, R. R. (2014). New methodology for the analysis and representation of human brain function: MEGMRIAn. *Math. Biol. Bioinformatics* 9, 464–481. doi: 10.17537/2014.9.464
- Welch, P. D. (1967). Use of fast fourier transform for estimation of power spectra - a method based on time averaging over short modified periodograms. *IEEE Trans. Audio Electroacoust.* 15, 70–73. doi: 10.1109/TAU.1967.1161901
- Zotov, V. S., Matlashov, A. N., Volegov, P. L., Savukov, I. M., Espy, M. A., Mosher, J. C., et al. (2008). Microtesla MRI of the human brain combined with MEG. *J. Magn. Reson.* 194, 115–120. doi: 10.1016/j.jmr.2008.06.007

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Llinás, Ustinin, Rykunov, Boyko, Sychev, Walton, Rabello and Garcia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

An exploratory data analysis of electroencephalograms using the functional boxplots approach

Duy Ngo^{1*}, Ying Sun², Marc G. Genton², Jennifer Wu³, Ramesh Srinivasan⁴, Steven C. Cramer³ and Hernando Ombao^{1*}

¹ Department of Statistics, University of California, Irvine, Irvine, CA, USA, ² Computer, Electrical and Mathematical Sciences & Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, ³ Department of Anatomy & Neurobiology, University of California, Irvine, Irvine, CA, USA, ⁴ Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, USA

OPEN ACCESS

Edited by:

Brian Caffo,
Johns Hopkins University, USA

Reviewed by:

Anand Joshi,
University of Southern California, USA
Haley Hedlin,
Stanford University, USA

*Correspondence:

Duy Ngo and Hernando Ombao,
Department of Statistics, University of
California, Bren Hall 2019, Irvine,
CA 92697-1250, USA
dngo5@uci.edu;
hombao@uci.edu

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 13 May 2015

Accepted: 28 July 2015

Published: 19 August 2015

Citation:

Ngo D, Sun Y, Genton MG, Wu J,
Srinivasan R, Cramer SC and Ombao
H (2015) An exploratory data analysis
of electroencephalograms using the
functional boxplots approach.
Front. Neurosci. 9:282.
doi: 10.3389/fnins.2015.00282

Many model-based methods have been developed over the last several decades for analysis of electroencephalograms (EEGs) in order to understand electrical neural data. In this work, we propose to use the functional boxplot (FBP) to analyze log periodograms of EEG time series data in the spectral domain. The functional boxplot approach produces a median curve—which is not equivalent to connecting medians obtained from frequency-specific boxplots. In addition, this approach identifies a functional median, summarizes variability, and detects potential outliers. By extending FBP's analysis from one-dimensional curves to surfaces, surface boxplots are also used to explore the variation of the spectral power for the alpha (8–12 Hz) and beta (16–32 Hz) frequency bands across the brain cortical surface. By using rank-based nonparametric tests, we also investigate the stationarity of EEG traces across an exam acquired during resting-state by comparing the spectrum during the early vs. late phases of a single resting-state EEG exam.

Keywords: EEGs time series, functional boxplots, surface boxplots, spectral analysis, band depth, exploratory analysis, stationarity

1. Introduction

Electroencephalograms (EEGs) have been used for many decades to study the complex spatio-temporal dynamics of brain processes (Nunez and Srinivasan, 2006). Due to its excellent temporal resolution (sampling rates usually range from 100 to 1000 Hz), EEGs can capture transient changes in brain activity, identify oscillatory behavior and study cross-dependence between EEG components. Since EEGs indirectly measure neuronal electrical activity, they can be used to infer the statistical properties of the underlying brain stochastic process. One such statistical property is the spectrum (or power spectrum) which decomposes the total variability in the EEG according to the contribution of oscillations at different frequencies. Most approaches to analyzing EEGs focus immediately on statistical modeling and spectral estimation. Here, we offer a systematic framework for exploring structures, patterns and features in the signal—prior to formal modeling. We explore the spectral properties only in a single channel using EEG traces from several epochs.

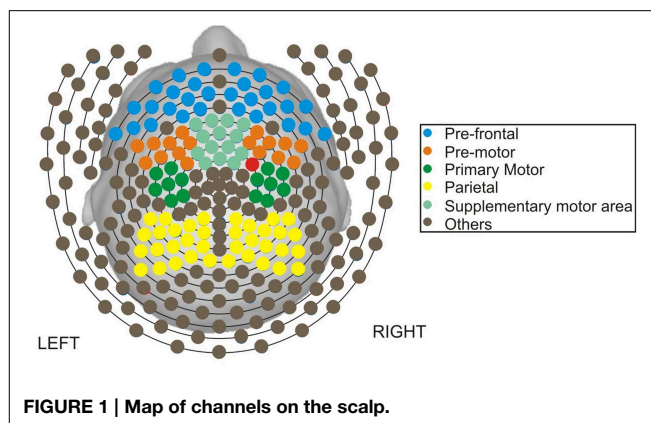
One approach to estimating the spectrum using EEG traces is to fit a parametric time domain model, such as the autoregressive moving average (ARMA) model. Applications of parametric modeling of EEGs have a long history. See (Bohlin, 1973; Isaksson et al., 1981; Krystal et al., 1999; Jain and Deshpande, 2004) among many others. When the spectrum of the EEG evolves over time

(e.g., within an epoch), one could still use the ARMA model but allow the coefficients to vary over time. A key element in ARMA models is the order of the autoregressive (AR) and moving average (MA) components. These can be obtained objectively using an information-theoretic criterion such as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Using these criteria, we obtain an optimal AR and MA order that jointly gives the best fit with the least complexity (as determined by the orders). BIC puts a heavier penalty for complexity compared to AIC and thus often gives a model with lower orders (lower complexity). From the parametric fit, we derive the estimates of the auto-correlation function and the spectrum. The theoretical background for parametric models are developed in Priestley (1981), Shumway and Stoffer (2000), and Brockwell and Davis (2009).

One could also estimate the spectrum without resorting to a parametric model. Under this approach, the EEGs are considered to be superpositions of sines and cosines (Fourier waveforms) with different frequencies and random amplitudes. These random amplitudes (or coefficients) are computed using the fast Fourier transform (FFT). The squared magnitude of these amplitudes, often called the periodograms, are the data-analogs of the spectrum defined on discrete frequencies. The theoretical background on the frequency domain approach to time series is developed in Brillinger (1981) and Percival and Walden (1993). This approach to analyzing EEGs continues to be popular in the cognitive and brain sciences. The following papers cover both methods and applications of spectral analysis to EEGs: Pfurtscheller and Aranibar (1979), Bressler and Freeman (1980), Makeig (1993), and Srinivasan and Deng (2012), to name a few.

The common practice prior to spectral estimation is to pre-process EEGs, often to remove artifacts (Makeig et al., 1996). After artifact rejection and segmentation according to epochs, the spectrum is estimated from each EEG trace. As noted, there is a lack a systematic framework for exploring structures, patterns and features in the signal—prior to formal modeling. Due to the complexity of EEG data, exploratory data analysis (EDA) plays an important role, especially when data are recorded from many epochs or trials during an experiment. For example, it is often expected that brain responses to the same stimulus ought to be relatively uniform, with minimal variation across epochs. In contrast, greater variability across epochs may be expected during neuroimaging studies that examine the brain in resting-state, as cognitive processes can vary within and across sessions for individual subjects and across subjects. An appropriate EDA methods can provide insights into features of EEG, including similarities and variability of the brain responses across epochs to facilitate the statistical model. In this paper, we propose to use the functional boxplot (FBP) method originally developed by Sun and Genton (2011) to address these questions.

The methods presented in this paper are motivated by a motor skill acquisition study at the Neuro-rehabilitation laboratory at the University of California, Irvine (Principal Investigator: Steven C. Cramer). In the previous study, EEG was recorded from 17 subjects both during resting-state prior to motor skill training and during motor skill training using dense-array EEG (256 electrodes) as shown in **Figure 1**. The resting-state EEG exam



was 3 min, and during post-processing, was segmented into 1-s non-overlapping epochs. As demonstrated in Wu et al. (2014), the spectral features of the resting-state EEGs when combined with a partial least squares regression analysis, was predictive of an individual's subsequent ability to acquire a novel motor skill. These may be of clinical importance to the field of rehabilitation, as improved methods for stratifying patients may significantly improve response to treatment and assist allotment of limited resources.

We present an *exploratory spectral analysis* (ESA) of resting-state EEG traces using FBP for one subject. In spectral analysis, the spectrum is an important stochastic property of the signal. It indicates the amount (or proportion) of variance that is explained by each frequency bin. Thus, the spectrum or the log spectrum of the EEG signal can be used to examine relative amounts of variability explained by slow (delta or theta) waves and fast (alpha or beta) waves. Throughout this analysis, we obtain a *sample* spectral curve by smoothing the log periodograms of each 1-s EEG epoch, and treat it as one observation unit in the FBP. By using the FBP, we address three primary objectives. The *first* objective is to identify the median, i.e., the most characteristic spectral curve rather than the pointwise frequency-specific medians. In addition, outliers are demonstrated by their unusual sample log spectral curve, and can be caused by extra-brain artifacts, including eye blinks, eye movements, and muscle movements in the EEG signal. Subsequently, confirmed outliers will be removed from subsequent analyses. The advantage of the FBP approach, over the usual pointwise boxplot method, is that it identifies epochs that have potential outlying spectral curves.

The *second* objective is to compare the median curves and the variability of the spectral curves from multiple phases of the resting state period. To test the stationarity of the EEG signal over the entire recording, we compare the spectral curves and the frequency-specific spatial distribution of spectral power during the early phase (first 60 epochs) vs. the late phase (last 60 epochs). Evidence against stationarity must be taken seriously since this would suggest an evolution of brain processes across the recording (Fiecas and Ombao, under revision). Moreover, the FBP approach is able to provide some characterization of the variation of the sample log spectral curves across EEG recording. In experiments comparing more than one group (e.g., healthy

controls vs. patients with stroke), it would be also interesting to determine whether groups differ with respect to consistency (uniformity) of the EEG signal over time.

The *third* objective is to investigate the spatial variability of spectral power across the brain for a given frequency band using the surface boxplot, which is a generalization of the FBP. Using the surface boxplots approach, it is possible to identify cortical regions (or channels) that, relative to the other channels, exhibit a high proportion of beta power. The beta band is particular interest to neuroscientists, as changes in beta activity have a good association with motor function (Roopun et al., 2006; Joundi et al., 2012).

The remainder of the paper is organized as follows. In Section 2, we present a comprehensive exploratory method which consists of the following: a review of the spectra in Section 2.1, a demonstration of automatic bandwidth selector for periodogram smoothing using the gamma generalized crossvalidation criterion in Section 2.2, some remarks on smoothing the periodogram in Section 2.3, a description of the FBPs in Section 2.4, a description of the surface boxplots in Section 2.5, and a demonstration of testing for differences in mean curves between families of curves in Section 2.6. In Section 3, we examine the finite sample performance of the proposed exploratory method. In Section 4, the resting-state EEG data are analyzed. Finally, in Section 5, conclusions and future work are discussed.

2. Method for Exploratory Spectral Analysis (ESA)

In this section, we review the methods that are needed for ESA of the EEG data. In Section 2.1, we first formally define the spectrum and then discuss a consistent estimator which is obtained by smoothing the periodogram using a bandwidth that is automatically selected by the gamma generalized cross-validation (Gamma-GCV) method described in Section 2.2. Next, we highlight two remarks on smoothing the periodogram in Section 2.3, then we present the FBPs method in Section 2.4 and surface boxplots method in Section 2.5. Finally, we present a rank sum test which tests for differences in median curves or surfaces between families of curves or surfaces in Section 2.6.

2.1. Spectrum

The spectrum of an EEG signal (which is assumed to be stationary) can give the amount of variance contributed by oscillatory components (from delta to beta band activity). Let $X(t)$, $t = \dots, -1, 0, 1, \dots$ be a zero-mean stationary time series with covariance function $\gamma(\tau) = E(X(t)X(t+\tau))$ ($\tau = \dots, -1, 0, 1, \dots$) that is assumed to be absolutely summable, i.e., $\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)| < \infty$. The spectrum, denoted $f(\omega)$, is defined to be

$$f(\omega) = \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-i2\pi\omega\tau}, \quad \omega \in \left[-\frac{1}{2}, \frac{1}{2}\right].$$

The starting point for estimating $f(\omega)$ is the periodogram. Denote $I(\omega_k)$ to be the periodogram computed from a finite sample of the

stationary process $X(0), X(2), \dots, X(T-1)$ at frequency $\omega_k = k/T$ which is defined to be

$$I(\omega_k) = \frac{1}{T} \left| \sum_{t=0}^{T-1} X(t) e^{-i2\pi\omega_k t} \right|^2, \quad k = -\lfloor T/2 \rfloor - 1, \dots, \lfloor T/2 \rfloor,$$

where $\lfloor T/2 \rfloor$ is the quotient of $T/2$.

To characterize the spectra of the EEG signals, we classify the oscillatory patterns of periodograms into four primary frequency bands: delta (0–4 Hz), theta (4–8 Hz), alpha (8–16 Hz), beta (16–32 Hz), and gamma (32–50 Hz) as shown in **Figure 2**. Since each frequency band is defined by a range, we define $\widehat{S}(\Omega)$ to be the estimated spectral power at the Ω band:

$$\widehat{S}(\Omega) = \int_{\omega \in \Omega} I(\omega) d\omega.$$

It is well-known that the periodogram $I(\omega_k)$ is an asymptotically unbiased estimator for $f(\omega_k)$, but it is inconsistent because its variance approaches a positive constant when $T \rightarrow \infty$. Therefore, to reduce the variance, we smoothed the periodogram. A number of nonparametric smoothing methods have been proposed including the kernel smoother (Lee, 1997; Ombao et al., 2001), wavelet (Gao, 1997), smoothing spline (Wahba, 1980; Pawitan and O'sullivan, 1994), or local polynomial (Fan and Kreutzberger, 1998). For kernel smoothing, Ombao et al. (2001) developed an automatic span selector via the generalized crossvalidation criterion for generalized additive models based on the *deviance* which is discussed in Section 2.2.

2.2. Automatic Span Selector Using the Gamma Generalized Crossvalidation Method

From Brillinger (1981) (Theorem 5.2.6), $I(\omega_k)$ follows an asymptotic distribution

$$I(\omega_k) \sim \begin{cases} \text{Gamma}(1, f(\omega_k)) & k = 1, \dots, T/2 - 1 \\ \text{Gamma}(\frac{1}{2}, 2f(\omega_k)) & k = 0, T/2, \end{cases}$$

where $I(\omega_0), \dots, I(\omega_{T/2})$ are independent. As a caveat, we note here that the actual result requires that the number of frequencies is fixed and does not depend on T . However, in most applications, this is often ignored. This result can be equivalently stated as $I(\omega_k)/f(\omega_k) \sim \epsilon_k$ where $\epsilon_k \sim \chi^2(1)$ when $k = 0$ or $T/2$ and $\epsilon_k \sim \frac{1}{2}\chi^2(2)$ when $k = 1, \dots, T/2 - 1$. As noted, we need to smooth the periodogram $I(\omega_k)$ to produce a consistent estimator for $f(\omega_k)$. Let $\widehat{f}_p(\omega_k)$ be a smoothed periodogram estimator of $f(\omega_k)$ which we define to be

$$\widehat{f}_p(\omega_k) = \sum_{j=-p}^p W_{p,j} I(\omega_{k+j}) \quad k = 0, \dots, T/2, \text{ and } j = -p, \dots, p$$

where $2p + 1$ is the smoothing span and $W_{p,j}$ are nonnegative weights that satisfy the following conditions for any fixed p :

$$W_{p,j} = W_{p,-j} (j = 1, \dots, p), \quad \sum_{j=-p}^p W_{p,j} = 1.$$

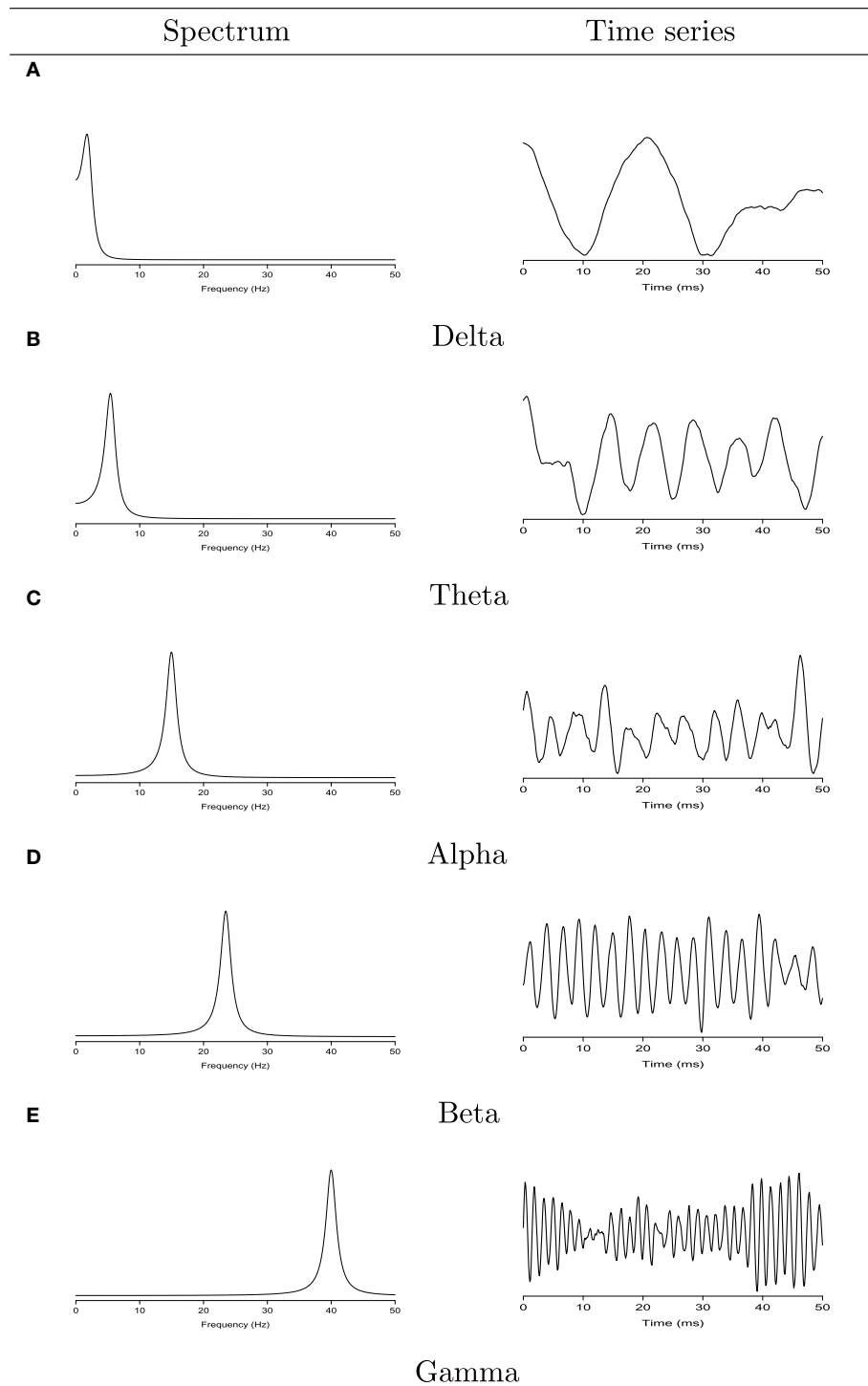


FIGURE 2 | Left: the spectrum of second order auto-regressive processes AR(2) with power concentrated at the delta (0–4 Hz), theta (4–8 Hz), alpha (8–16 Hz), beta (16–32 Hz), and gamma (32–50 Hz) bands. **Right:** realizations from each corresponding AR(2) process. **(A)** Delta band. Left: the spectrum of second order auto-regressive processes AR(2). Right: realizations from each corresponding AR(2) process. **(B)** Theta band. Left: the spectrum of second order auto-regressive processes AR(2). Right:

realizations from each corresponding AR(2) process. **(C)** Alpha band. Left: the spectrum of second order auto-regressive processes AR(2). Right: realizations from each corresponding AR(2) process. **(D)** Beta band. Left: the spectrum of second order auto-regressive processes AR(2). Right: realizations from each corresponding AR(2) process. **(E)** Gamma band. Left: the spectrum of second order auto-regressive processes AR(2). Right: realizations from each corresponding AR(2) process.

Generally, the weights are chosen so that $W_{p,j}$ is a decreasing function of p , but (Priestley, 1981) shows that the choice of the weights $W_{p,j}$ is of secondary importance to the value of the span or bandwidth. Thus, for simplicity, we use the boxcar smoother with weights defined by $W_{p,j} = 1/(2p + 1)$ for all $j = -p, \dots, p$. The gamma generalized crossvalidation method selects p to minimize the generalized crossvalidated deviance function

$$GCV(p) = \frac{M^{-1} \sum_{j=0}^{M-1} D(I(\omega_j), \hat{f}_p(\omega_j))}{(1 - \text{tr}(H_p)/M)^2},$$

where $M = T/2 + 1$. The deviance $D(I(\omega_j), \hat{f}_p(\omega_j))$ can be chosen as $q_j \{-\log(I(\omega_j)/\hat{f}_p(\omega_j)) + (I(\omega_j) - \hat{f}_p(\omega_j))/\hat{f}_p(\omega_j)\}$ (McCullagh and Nelder, 1989). Here, $q_j = 1 - 0.5\mathcal{I}\{j = 0, M - 1\}$, and \mathcal{I} is the indicator function. The H_p is the smoother matrix with smoothing parameter p , and the term $(1 - \text{tr}(H_p)/M)^2$ often referred to as the model degrees of freedom, can be expressed in terms of the weight at the center of the smoothing window: $(1 - W_{p,0})^2$. Then, the generalized crossvalidated deviance function can be written as

$$GCV(p) = M^{-1} \sum_{j=0}^{M-1} q_j \left\{ \frac{-\log(I(\omega_j)/\hat{f}_p(\omega_j)) + (I(\omega_j) - \hat{f}_p(\omega_j))/\hat{f}_p(\omega_j)}{(1 - W_{p,0})^2} \right\}.$$

2.3. Remarks

For frequencies over 100 Hz, the periodogram values are almost negligible because the signals underwent low-pass filtering at 100 Hz. , so for simplicity, we will only show the spectrum over the frequency range of 0–100 Hz. In **Figure 3**, we show the location of channel 197 in right pre-motor region at the resting-state. **Figure 4** gives an illustration of smoothing the periodograms for randomly selected epochs 3, 85, and 160 for a fixed channel 197. It can be seen that the power at these periodograms are dominated by low frequencies, and the values of smoothing span minimizing the generalized crossvalidated deviance function are about 3–5. Also, the smoothing lines reasonably approximate the periodograms and the small bandwidths preserve the peaks. Second, since the distribution of $I(\omega_k)$ is a multiple of the spectral density, its variance [which depends on $f(\omega_k)$] also changes across the frequencies ω_k . To stabilize the variance across frequencies and to standardize comparisons of median curves across two phases (early vs. late phases of the resting-state EEG recording) we will use the log transformed periodograms. It is convenient then, that the variance of the log periodograms at each frequency is constant and takes the approximate value of $\frac{\pi^2}{6}$. Moreover, while the periodogram is approximately unbiased for the spectrum, the log periodogram is no longer (approximately) unbiased for the log spectrum due to Jensen's inequality. This is easily fixed by adding the Euler Mascheroni constant 0.57721 to log transformed periodograms to obtain the log bias-corrected periodograms

(Wahba, 1980). Let $g(\omega_k)$ be the true log spectrum, then $Y_r(\omega_k)$, the log bias of the corrected periodogram at epoch r , is defined as

$$Y_r(\omega_k) = g(\omega_k) + 0.57721, \quad k = 0, 1, \dots, T/2.$$

Figure 5 gives the log bias-corrected periodograms, $Y_r(\omega_k)$, corresponding to **Figure 4**. Throughout this paper, we will apply the gamma crossvalidation method to obtain the optimal smoother of log bias-corrected periodograms.

2.4. Functional Boxplots

The FBP is constructed in a similar manner to the classical (pointwise) boxplot. Each observation will be sorted based on decreasing values of some depth measure, and band depth is one notion. A curve is said to be “deeply situated” within a sample of curves if it is covered by many bands from pairs of curves. This idea is an extension of a pointwise boxplot where the median is also located “deep” in a sample because it is situated in the middle of the boxplot and hence covered by many pairs of points. Here, our observation units are curves (or real-valued functions) which are the log bias-corrected periodograms $Y_r(\omega_k)$, $k = 0, \dots, T/2$ over many epochs r . The notion of a band depth was introduced in López-Pintado and Romo (2009) through a graph-based approach to order all sample curves which we briefly describe. Suppose that a curve $Y(\omega_k)$ is the subset of the plane $G(Y(\omega_k)) = \{(\omega_k, Y(\omega_k)) : \omega_k \in \mathcal{A} = [0, T/2]\}$. The band in \mathbb{R}^2 can be delimited by a number J of curves, and this number is fixed as $J = 2$ in our study. Now, let Y_α, Y_β be two continuous functions, $L_k = \min(Y_\alpha(\omega_k), Y_\beta(\omega_k))$, and $U_k = \max(Y_\alpha(\omega_k), Y_\beta(\omega_k))$. Then the band delimited by Y_α, Y_β is

$$B(Y_\alpha, Y_\beta) = \{(\omega_k, Y'(\omega_k)) : \omega_k \in \mathcal{A}, L_k \leq Y'(\omega_k) \leq U_k\}.$$

Let Y_1, \dots, Y_n be n independent sample curves, then the band depth for a given curve Y_i , $i = 1, \dots, n$ is defined as

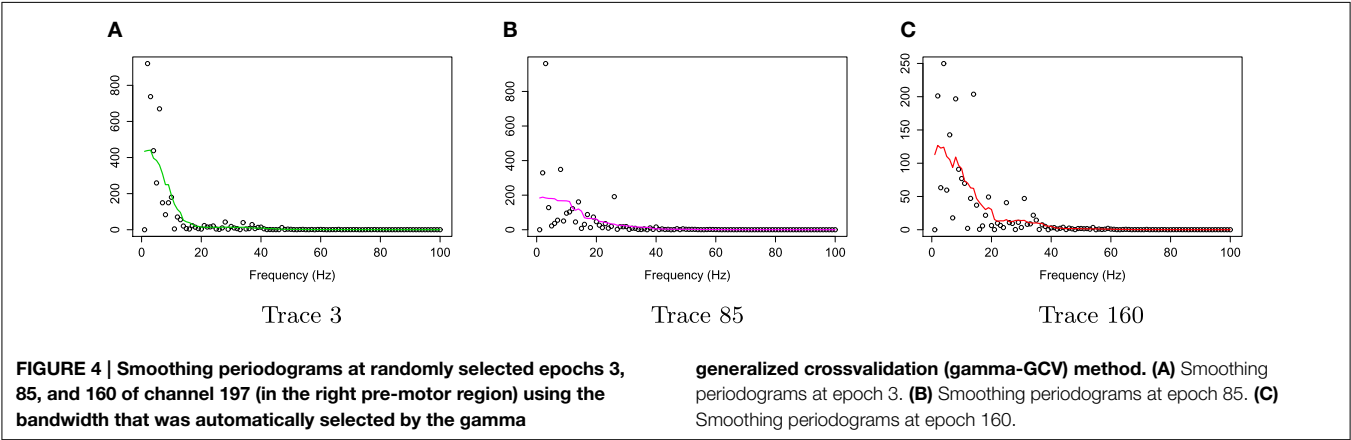
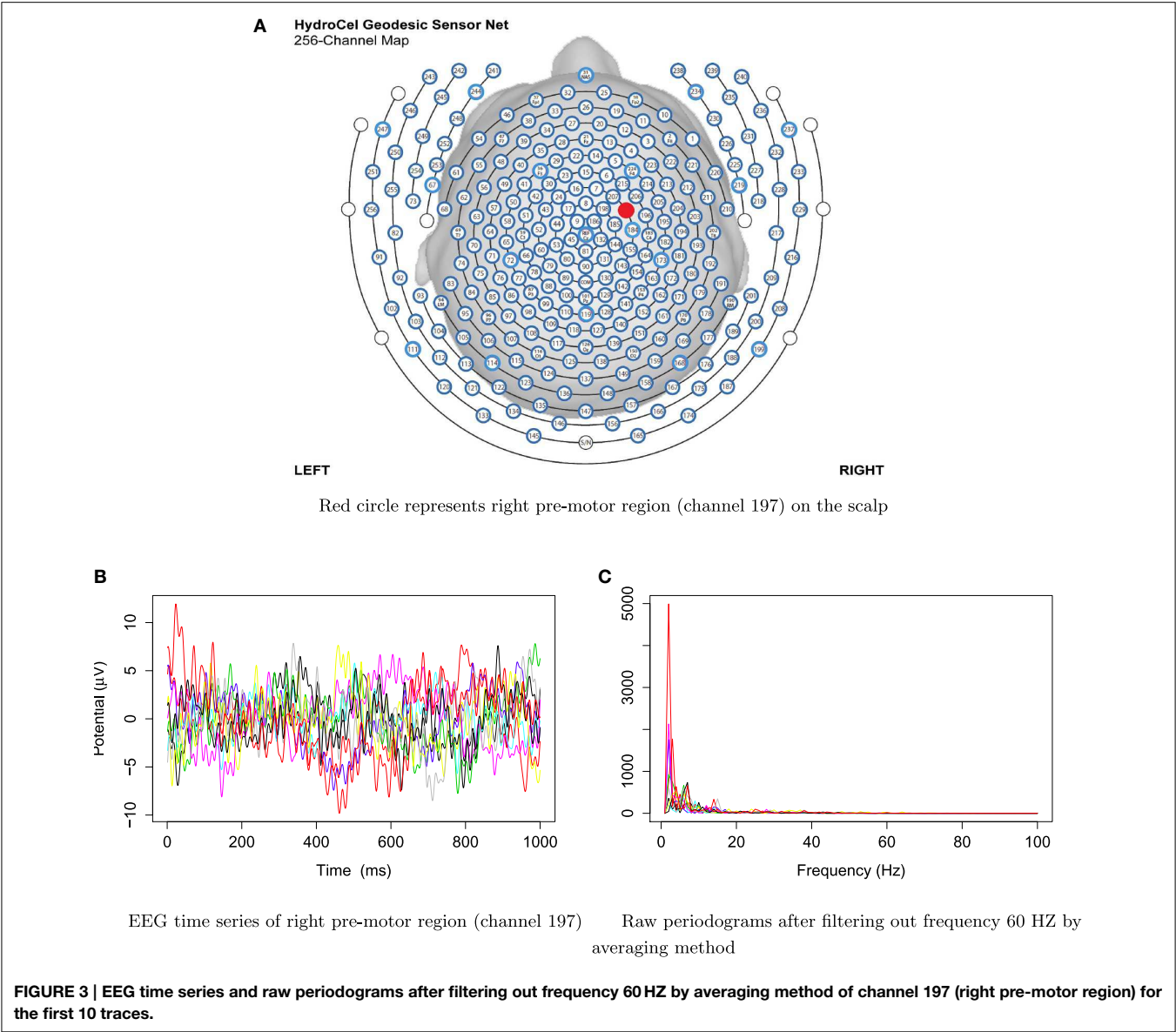
$$BD(Y_i) = \binom{n}{2}^{-1} \sum_{\alpha=1, \dots, n; \beta=1, \dots, n} \mathcal{I}\{G(Y_i) \subseteq B(Y_\alpha, Y_\beta)\}$$

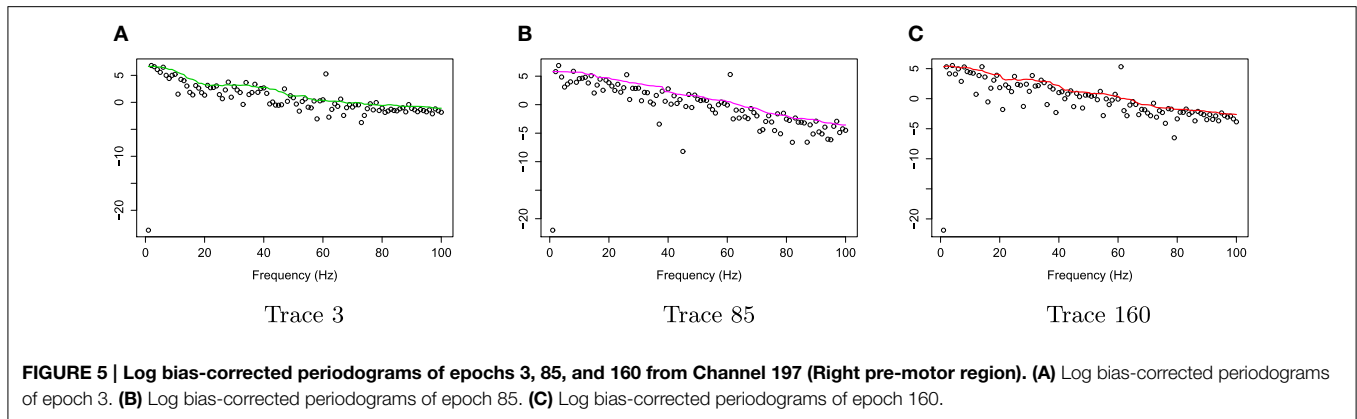
where $\mathcal{I}(\cdot)$ is the indicator function. When $J = 2$, there are $\binom{n}{2}$ possible bands delimited by two curves. The limit of the band depth BD is that it does not measure the proportion of curve inside the band. Thus, López-Pintado and Romo (2009) also proposed a modified band depth method (MBD), which measures the proportion of a curve Y_i that is actually in a band:

$$MBD(Y_i) = \binom{n}{2}^{-1} \sum_{\alpha=1, \dots, n; \beta=1, \dots, n} \lambda\{A(Y_i; Y_\alpha, Y_\beta)\}$$

where $A(Y_i; Y_\alpha, Y_\beta) \equiv \{\omega_k \in \mathcal{A} : L_k \leq Y_i \leq U_k\}$, $\lambda(Y_i) = \lambda(A(Y_i; Y_\alpha, Y_\beta))/\lambda(\mathcal{A})$, and λ is a Lebesgue measure on \mathcal{A} . We notice that the MBD computation will be time-consuming when n is large, so we use an exact fast method from Sun et al. (2012) to compute the MBD for the EEG data.

Based on the ranks of the depths of the curves, the FBPs can provide the descriptive statistics, such as the 50% central





region, the median curve, and the maximum and minimum non-outlying curves. Moreover, the potential outliers can be detected by the 1.5 times inter-quartile range (IQR) empirical rule, which is commonly used for classical boxplots. The boundary region is defined as 1.5 times the height of the 50% central region. Any curves outside this region are considered potential outliers. In contrast with a constant factor 1.5 in classical boxplot, a factor 1.5 in FBP can be modified due to potential spatio-temporal outliers. This is because the curves from different locations will be spatially correlated, and there can be dependence in time/frequency for each curve (Sun and Genton, 2012a).

2.5. Surface Boxplots

Similar to FBPs, one can compute the data depth of all the observations, then order them according to decreasing depth values. Suppose that the observed sample surfaces, $z_1(s), \dots, z_n(s)$, $s \in \mathcal{S}$, where \mathcal{S} is a region in \mathbb{R}^2 . The information unit for such a dataset is the entire surface. To order sample surfaces, we need to generalize univariate order statistics to surfaces. To this end, we generalize the MBD with $J = 2$ to \mathbb{R}^3 through a volume. Genton et al. (2014) define the sample modified volume depth (MVD) to be

$$MVD_n(z) = \left(\binom{n}{2} \right)^{-1} \sum_{1 \leq i_1 \leq i_2 \leq n} \lambda_r A(z; z_{i_1}, z_{i_2}),$$

where $A(z; z_{i_1}, z_{i_2}) \equiv \mathbf{s} \in \mathcal{S} : \min_{r=i_1, i_2} z_r(\mathbf{s}) \leq z(\mathbf{s}) \leq \max_{r=i_1, i_2} z_r(\mathbf{s})$ and $\lambda_r(z) = \frac{\lambda(A(z; z_{i_1}, z_{i_2}))}{\lambda(\mathcal{S})}$, if λ is the Lebesgue measure on \mathbb{R}^3 . A sample median surface is a surface from the sample with the largest sample MVD value, designed by $\arg \max_{z \in z_1, \dots, z_n} MVD_n(z)$. If there are ties, the median will be the average of the surfaces maximizing the sample MVD.

The first step for constructing surface boxplots is the surface ordering. Sample surfaces are ordered from the center outwards based on their MVD values, inducing the order $z_{[1]}, z_{[2]}, \dots, z_{[n]}$. The sample α central region is naturally defined as the volume delimited by the α proportion ($0 < \alpha < 1$) of the deepest surfaces. In particular, the sample 50% central region is

$$C_{0.5} = \{(\mathbf{s}, z(\mathbf{s})) : \min_{r=1, \dots, [n/2]} z_{[r]}(\mathbf{s}) \leq z(\mathbf{s}) \leq \max_{r=1, \dots, [n/2]} z_{[r]}(\mathbf{s})\},$$

where $[n/2]$ is the smallest integer not less than $n/2$. The border of the 50% central region is defined as the inner envelope representing the box in a surface boxplot. This is the surface analog of the first and third quartiles of the classical boxplot. The median surface in the box is the one with the largest depth value. Because the ordering is from the center outwards, the volume of the central region increases as α increases. Hence, the maximum envelope, or the outer envelope, is defined as the border of the maximum non-outlying central region. To determine this region, we propose to identify outlying surfaces by an empirical rule similar to the 1.5 times the 50% central region rule in a FBP. The fences (or the upper and lower surface boundaries for flagging potential outliers) are obtained by inflating the inner envelope (as defined above) by 1.5 times the height of the 50% central region. Any surface crossing the fences are flagged as potential outliers. The factor 1.5 can be also adjusted as in the adjusted FBPs to take into account spatial autocorrelation and possible correlations between surfaces.

2.6. Testing for Differences in Median Between Families of Curves or Surfaces

To compare the median curves from two populations of curves, López-Pintado and Romo (2009) proposed the rank sum test. Let $\tilde{\mu}_Y$ and $\tilde{\mu}_{Y'}$ be the median curves of two populations Y and Y' , respectively. Define the null hypothesis to be

$$H_0 : \tilde{\mu}_Y = \tilde{\mu}_{Y'} \text{ for all } \mu.$$

Suppose that we observe two sets of curves, namely $\{y_1, \dots, y_n\}$ and $\{y'_1, \dots, y'_m\}$. Then define the reference sample to be $\{r_1, \dots, r_k\}$ which is from one of the two observed sets with $k \geq \max(n, m)$. The position of a particular y_i for $i = 1, \dots, n$, or y'_j for $j = 1, \dots, m$ with respect to the reference sample r , is defined as

$$R(y_i) = \frac{1}{n} \sum_{l=1}^n \mathcal{I}\{MVD(z_l) \leq MBD(y_i)\},$$

$$R(y'_j) = \frac{1}{m} \sum_{l=1}^m \mathcal{I}\{MVD(z_l) \leq MBD(y'_j)\},$$

where MBD is the MBD defined in previous section, and \mathcal{I} is the indicator. Then, we can order the values $R(y_i)$ and $R(y'_i)$ from the smallest to the largest, and their ranks are between 1 and $n+m$. The test statistics $T = \sum_{i=1}^m \text{rank } R(y'_i)$, then under the null hypothesis H_0 , the distribution of T is the distribution of the sum of m numbers that are randomly chosen from $1, 2, \dots, n+m$ (Sun and Genton, 2012b).

2.7. Remarks on the Applications of Functional and Surface Boxplots

In this paper, we use functional and surface boxplots to explore the structure of EEGs. However, these methods are general and can be applied to other types of data such as growth data and climate time series (Sun and Genton, 2012b).

3. Simulation Study

The purpose of the simulation study is to examine the performance of the exploratory spectral methods under various experimental settings. In Section 3.1, we demonstrate the performance of the FBP on the smoothed log periodograms of a mixture of two first order AR time series, denoted AR(1). In Section 3.2, we illustrate the rank sum test to compare the functional median from two families of curves.

3.1. Functional Boxplot Simulation Study

For the r^{th} epoch, let $U_{1r}(t)$ be an AR(1) process with its spectra dominated by high frequencies and $U_{2r}(t)$ be another AR(1) with its spectra mostly containing low frequencies. The AR(1) parameters are allowed to vary across epochs. Here, we set $t \in T = \{1, \dots, 1000\}$. We define $X_r(t)$ to be the mixture of $U_{1r}(t)$ and $U_{2r}(t)$, such that

$$X_r(t) = a_{1r}U_{1r}(t) + a_{2r}U_{2r}(t)$$

where $r = 1, \dots, 220$, a_{1r} and a_{2r} are weighted coefficients of $U_{1r}(t)$ and $U_{2r}(t)$, respectively. Then, the model for high and low frequency AR(1) processes are defined as

$$U_{\ell r}(t) = \phi_{\ell r}U_{\ell r}(t-1) + W_{\ell r}(t)$$

where $\ell = 1, 2$ and $W(t)$ is white noise. In this setting, the high and low frequency AR(1) are distinguished by the value of $\phi_{\ell r}$. For example, for high frequency $U_{1r}(t)$, we set $\phi_{1r} = 0.9 + \xi_r$, where ξ_r are independent and identically distributed from $\mathcal{N}(0, 0.001)$. Similarly, for low frequency $U_{2r}(t)$, we set $\phi_{2r} = -0.5 + \eta_r$, and η_r are also independent and identically distributed from $\mathcal{N}(0, 0.001)$. Here, we need the variance of ξ_r and η_r to be small so that it guarantees causality, i.e., $\xi_r \in (-1, 1)$ and $\eta_r \in (-1, 1)$. Next, we split the 220 subjects into two groups, such that the first group will include both high and low frequency series, $U_{1r}(t)$ and $U_{2r}(t)$, while the second group will only have the high frequency series $U_{1r}(t)$. To split $X_r(t)$ into two groups, we set the weight coefficients a_{1r} and a_{2r} as following

$$\begin{aligned} a_{1r} &\sim \mathcal{N}(10, 1) \text{ for } r = 1, \dots, 220 \\ a_{2r} &\sim \mathcal{N}(5, 1), \text{ for } r = 1, \dots, 120, \text{ and} \\ a_{2r} &\sim \mathcal{N}(0, 0.001) \text{ for } r = 121, \dots, 220. \end{aligned}$$

The two groups of $X_r(t)$ are shown in **Figure 6**. Using the gamma generalized crossvalidation method, **Figure 7** displays the log bias-corrected periodograms for each group, and **Figure 8** shows the corresponding FBP. Note that group 1 is dominated by both high (right) and low (left) frequencies while group 2 includes only low frequencies. Thus, the functional median of group 1 should have two peaks, one each in high and low frequency ranges, while the functional median of group 2 has only one peak in the low frequency range. In **Figure 8**, the black curve is the median curve in the center of the FBP. The two median curves from each group have clearly summarized the typical power distribution for each group. The blue curves in the center form the envelope of the 50% central region. The blue curves outside of the 50% central region are the non-outlying minimum and maximum curves. It is worth remarking that the envelope of group 1 is smaller than the envelope of group 2, and therefore, we demonstrate that group 2 has more dispersion than group 1. Moreover, the envelope of group 1 is in the middle of the non-outlying minimum and maximum curves, while the envelope of group 2 tends to move upwards. This indicates that group 2 shows more skewness than group 1. The red dashed curve in **Figure 8** denotes the outliers. We see that the curves from group 1 that are dominated by high frequencies only are detected as outliers while the curves from group 2 that include both high and low frequencies are detected as outliers.

In order to illustrate the usefulness of the FBP compared to the pointwise boxplot, we introduce a simulation study which randomly chooses 10 bias-corrected log periodograms among 160 total periodograms. We simulate an outlying curve by adding additional noise across the 0–100 Hz frequency range, and close to the center for the remaining frequencies. **Figure 9A** shows the simulation data including the 10 random bias-corrected log periodograms (gray curves) and a simulated outlying curve (red curve). In **Figure 9B**, the FBP successfully detects the simulated outlying curve and other outliers. However, **Figure 9C** shows that the pointwise boxplot fails to detect the simulated outlying curve, and provides some disconnected outlying curves across frequencies. We also notice that the non-outlying maximum and minimum curves of pointwise boxplot are actually the outlying curves detected by FBP. **Figure 9D** compares the two median curves from these two methods, and by visual inspection, there is a slight difference between the two median curves at low frequencies. Thus, FBP can be a non-parametric method to obtain the median curve and the variability around it for EEG data compared to pointwise boxplot.

3.2. Rank Sum Test Simulation Study

To investigate the performance of this nonparametric test, we simulated two sets of curves, which are defined as below:

$$Y_{\ell,r}(\omega_k) = f_{\ell}(\omega_k) + a_r g(\omega_k) + h_r(\omega_k),$$

where $r = 50$, $\ell = 1, 2$, $g(\omega_k) = 1$ for all ω_k , and ω_k is defined as $\omega_k = k/100$, where $k = 1, \dots, 100$. In the model, $f_1(\omega_k)$ and $f_2(\omega_k)$ are the mean functions; $a_r \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 5)$ and $h_r(\omega_k) \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 2)$ represent the variation between and within the

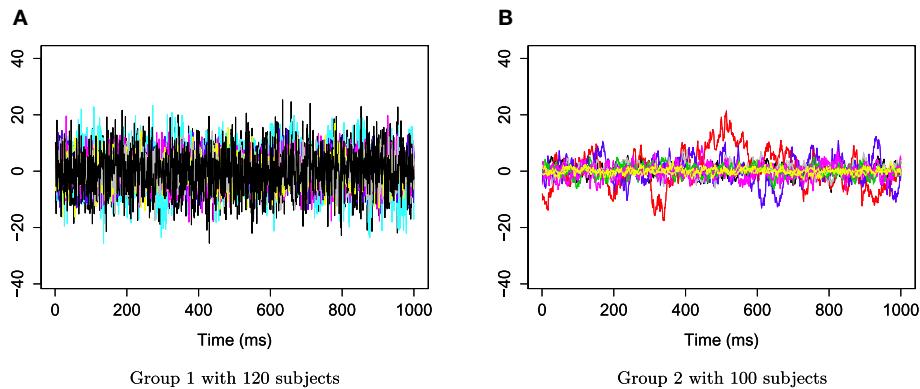


FIGURE 6 | Time series AR(1) for group 1 and group 2. (A) Time series AR(1) for group 1 with 120 subjects. **(B)** Time series AR(1) for group 2 with 100 subjects.

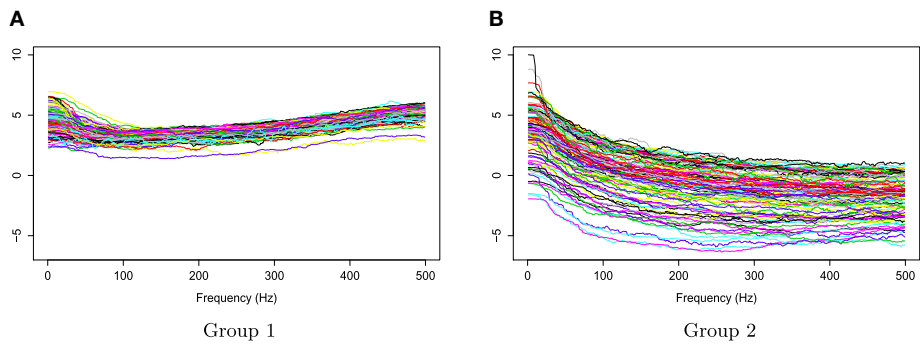


FIGURE 7 | Smoothed log bias-corrected periodograms for Group 1 and Group 2.

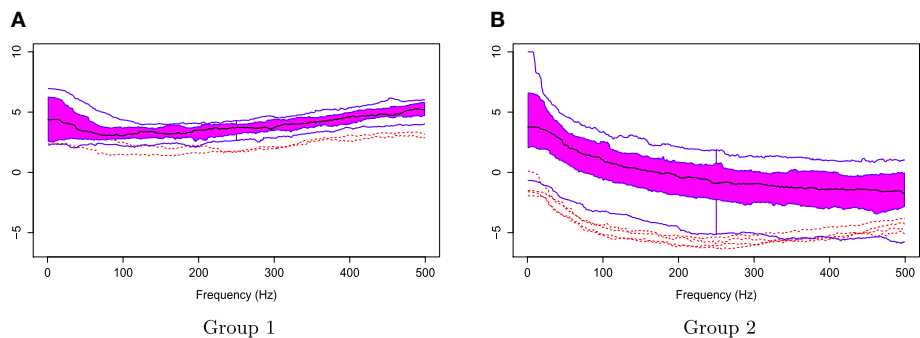


FIGURE 8 | Functional boxplots of Group 1 and Group 2 with a black curve representing the median curve, the pink area denoting the 50% central region, the two inside blue curves indicating the envelopes of 50% central region, the two outside

blue curves representing for two non-outlying extreme curves, and the red dashed curves illustrating the outlier candidates detected by 1.5 times the 50% central region rule. (A) Functional boxplots of Group 1. **(B)** Functional boxplots of Group 2.

curves, respectively. Let the function $f_1(\omega_k)$ be defined as

$$f_1(\omega_k) = 5 \cdot \sqrt{1000 \cdot \omega_k},$$

and consider three different cases:

1. The two means are identical, let $f_2(\omega_k) = f_1(\omega_k)$ for all ω_k .

2. There is a slight deviation between the two means; define $f_2(\omega_k) = 5 \cdot \sqrt{900 \cdot \omega_k}$.

3. There is an appreciable deviation between the two means; let $f_2(\omega_k) = f_1(\omega_k) + 2k/3$.

We applied the kernel average smoother with window size 7 to smooth each curve from these two families. **Figure 10** illustrates

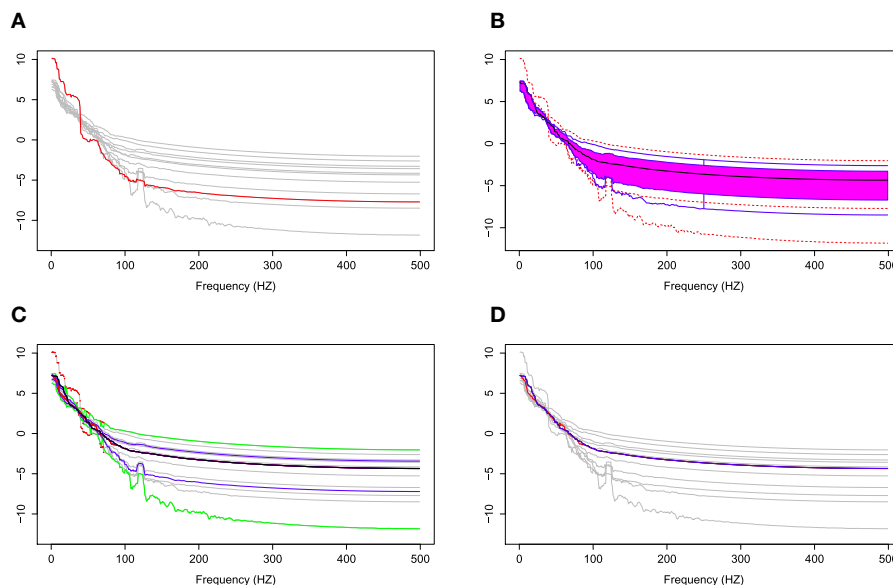


FIGURE 9 | (A) Simulation data with gray curves representing sample curves, a red curve denoting the simulated outlying curve **(B)** functional boxplots, **(C)** pointwise boxplots with black curve representing a mean curve, blue curves for the envelope of the 50% central region, the green

curves for the non-outlying minimum and maximum curves, and the red points for outliers, and **(D)** two median curves obtaining by functional boxplots method (blue) and pointwise boxplot method (red) are shown in the same plot.

the simulated curves (left panel) and the smoothed curves (right panel). In order to investigate the rank sum test performance in each case, we simulated two families of curves and obtain p -values of rank sum test; this procedure was repeated 1000 times. Let the type I error α be 5%, we report the percentage of time that the rank sum test rejects $H_0 : f_1(\omega_k) = f_2(\omega_k)$ for all ω_k in **Table 1**.

Overall, the rank sum test method performed well in each case. When the two families are identical, this method rejected the null hypothesis of equality only 44 times (4.4%) out of 1000 times, which is close to the nominal α . When the two families are nearly identical, this method rejects 605 times (the power is 60.5%), and when the two families are completely different, the power is 100%. Thus, this method demonstrates power and sensitivity to differences.

4. Analysis of Resting-State EEGs Data

4.1. Data Description

In this paper, we analyze EEG data from one participant in a resting-state EEG study approved by the Institutional Review Board of the University of California, Irvine. The overarching aim of this study was to identify a pattern of EEG-derived coherence acquired during rest-state that could predict subsequent response to training on a novel motor skill. During EEG acquisition, subjects sat quietly with both feet flat on the floor, and were instructed to fixate their gaze to the center of a fixation cross. Each recording was 3 min in duration. While the original EEG recording included 256 channels, only 194 were used in subsequent analyses, as extra-brain artifacts, including cheek and neck muscle artifacts, and heart rhythms,

are more likely to contaminate EEG signals recorded from electrodes overlying cheek and neck regions. Following data acquisition, pre-processing steps included: 100 Hz low pass filter; EEG segmentation into 1-s consecutive, non-overlapping epochs; mean detrend; and EEG signal re-reference to mean signal across all 194 channels. In addition, a combination of visual inspection and Infomax Independent Component Analysis decomposition were used to remove extra-brain artifacts, including eye blinks, eye movements, muscle artifact, and heart rhythm artifacts. The final dataset consisted of 160 epochs, with each epoch lasting 1 s, and $T = 1000$ time points for each epoch.

The goals of the present analysis are as follows: In Section 4.2, we closely examined a representative channel in the pre-motor region (specifically channel 197 in this dataset). Since EEGs are not well-localized in space (as opposed to local field potentials), conclusions are constrained to the sensor space. However, electrical activity captured in channel 197 reflects activity roughly around the pre-motor area. Specifically, we estimated the (log) spectrum for each epoch to identify any frequency bin or frequency band that accounts for the majority of the power spectrum. Moreover, using the method of estimating the functional medians, we obtained an estimate of the median curve from the log periodogram curves obtained from several epochs. The median curve is interpreted as a “typical” (log) spectral profile across several epochs. Using this method, we also identified outlier curves which could also be interpreted as epochs with “unusual” EEG activity. In Section 4.3, we investigated the possibility of non-stationarity across the 3 min resting-state EEG recording. Our specific goal was to compare the log spectrum during the early phase (first 60

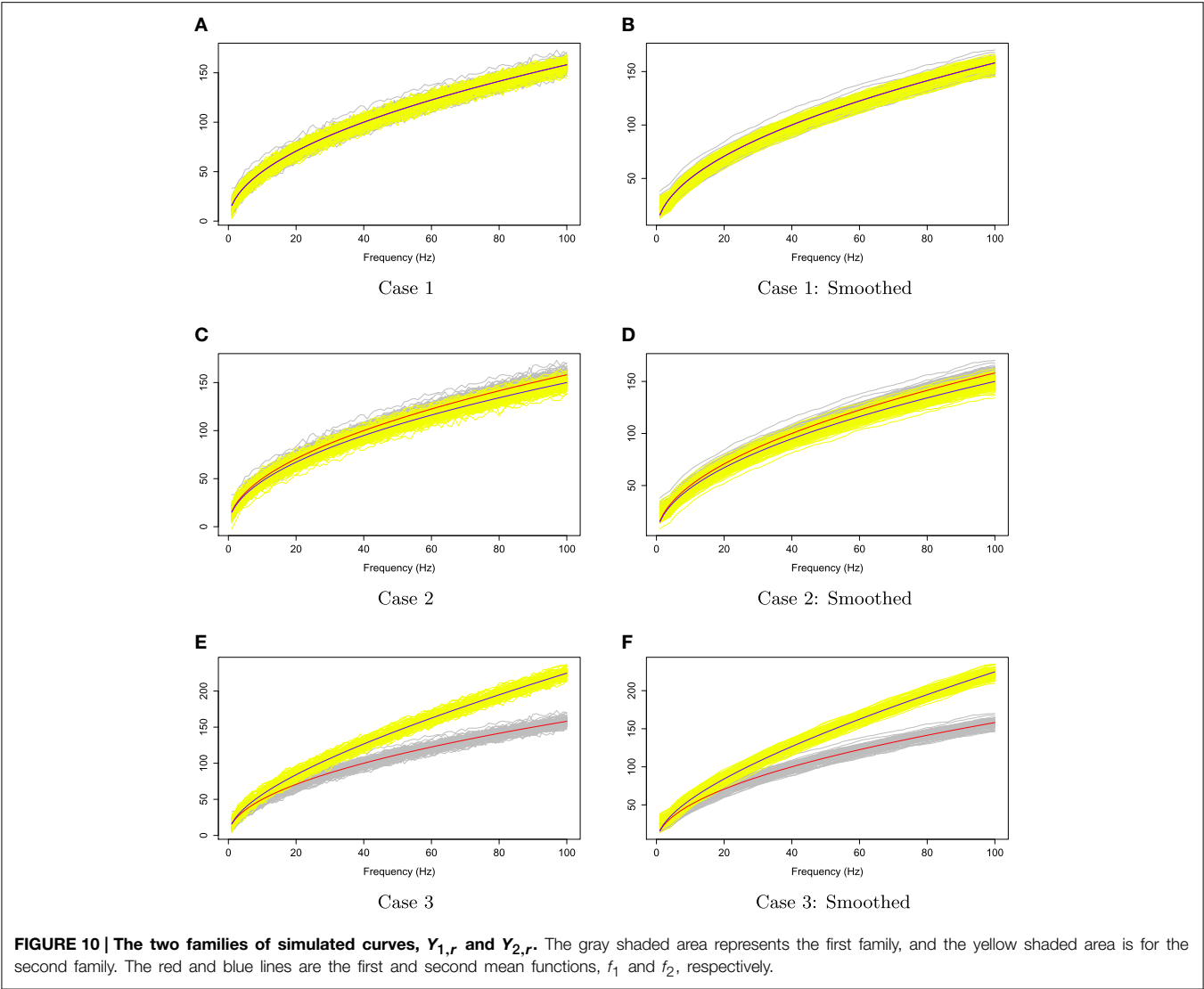


TABLE 1 | Rank sum test study result.

	First case	Second case	Third case
Percentage of time rejecting H_0	44	605	1000

epochs) of the recording with the log spectrum during the late phase (last 60 epochs) of the recording, and identify frequency bands that exhibit any differences between the early vs. late phases. In Section 4.4, we studied the spatial variation of power, at each of the five frequency bands: delta, theta, alpha, beta, and gamma, across all 194 channels, with the goal of identifying regions that exhibit relatively greater proportion of spectral power in each of the five frequency bands of interest. Finally, we compared the spatial variation for each of the five bands during the early vs. late phases of the resting-state EEG recording.

4.2. Functional Medians of the Pre-motor Log Spectral Curves

The log of the bias-corrected periodograms at the representative channel (channel 197) that approximately overlies cortex of the pre-motor region recorded for several traces and the FBPs are displayed in **Figure 11A**. The functional median curve is represented by the black curve, which is located inside the 50% central region, shaded area. The two blue curves outside of the shaded area are the non-outlying maximum and minimum curves. Similar to a FBP, we show in **Figure 11B** the pointwise boxplot (per frequency point), where the black curve is the median obtained by connecting the medians at each frequency point; the blue curves form the central region (50-th percentile region); the green curves are two non-outlying extreme curves. We compared these two median curves in **Figure 11C** and noted a slight discrepancy between these median curves derived using a FBP and the pointwise boxplot, with an emphasis on the low frequency range. The main difference between the functional

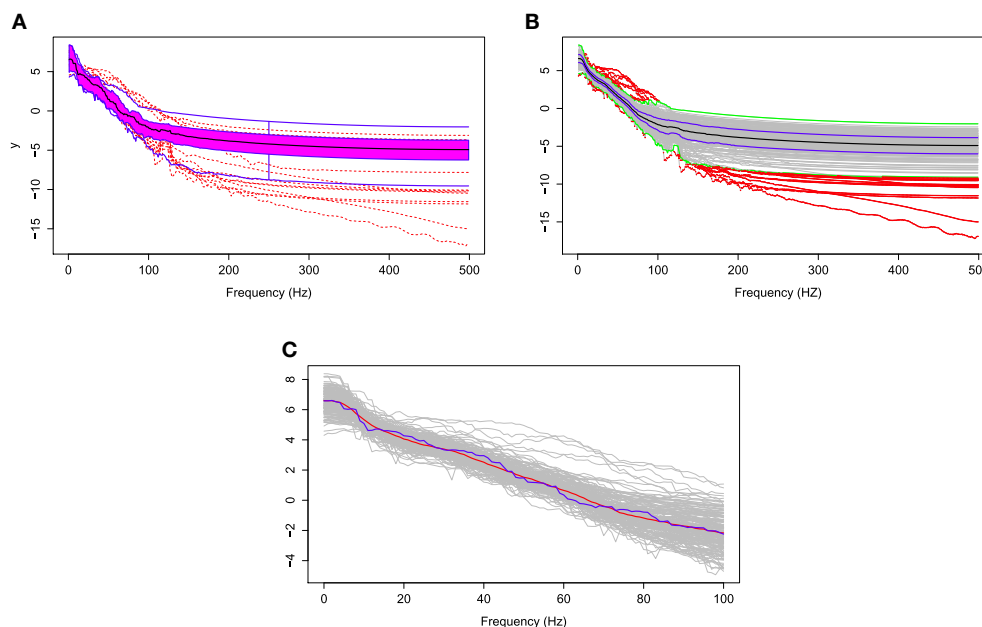


FIGURE 11 | (A) The functional boxplots, **(B)** pointwise boxplots of log bias-corrected periodograms, and **(C)** two median curves obtained by functional boxplot method (blue) and pointwise boxplot method (red) are shown in the same plot.

median and the point-wise median curve is in the interpretation. The former is one of the curves from a recorded epoch, whereas the latter may not be an actual curve. Hence the latter cannot really be interpreted as a “typical” curve from a family of curves formed from several epochs. Moreover, the FBP approach allows us to identify specific epochs that produce “unusual” or outlying log bias-corrected periodogram curves. Note that in the plots, the gray curves are the log bias-corrected periodograms of 160 epochs and the red curves are outliers. **Figure 11B** also shows that these outlying curves are discontinuous around the frequency bin centered at 100 Hz.

4.3. Testing for Stationarity of EEG Epochs Across the Entire Resting-state

In the previous section, the FBP provided descriptive statistics for the log bias-corrected periodograms of 160 epochs from the pre-motor region. Note that there were originally 180 epochs but 20 had to be removed from further analysis due to extra-brain artifact contamination. Our interest now is to test whether resting-state brain activity evolved across the 3 min EEG recording. While there are many ways to characterize such an “evolution” of the underlying brain processes, here we will specifically look into changes on the log spectral curves for early vs. late phases of the resting-state EEG recording. In this case, a change in the log spectral power in early vs. late phases would indicate non-stationarity of the EEG signal across the resting-state recording.

The null hypothesis of stationarity here is that the true median curves of the early and last phrases are identical. We test this hypothesis using the rank sum test with the significance level set

to 0.05. We defined the early phase to include the first 60 epochs (60 s) of the 3 min recording and the late phase to include the last 60 epochs. In **Figure 12**, we display the FBPs and the other descriptive statistics for each phase. A visual inspection suggests that the median curves are only slightly different from each other for electrodes that approximately overlie the pre-motor region. More significant differences are noted for electrodes that approximately overlie the prefrontal region (see **Figure 12C**). Moreover, the rank sum test failed to reject the null hypothesis, as the p -value is 0.56. Therefore, the two median curves are not significantly different and the hypothesis of stationarity in the pre-motor regions is not rejected. This is not entirely unexpected since the whole 3-min recording was purely resting-state. There was no experimental stimulus and the time frame was short.

Next, we use the same testing procedure at this particular channel in the pre-motor region (channel 197) to test the same null hypothesis of non-evolution of the brain process at each of the other channels across the 3 min EEG recording. Among the 194 total channels, 18 channels were identified that demonstrated a significant difference in median curves during the early vs. late phase at a significance level of 0.05. These channels are represented by colored circles in **Figure 13**. Of these 18, channel 29 (approximately overlying the supplementary motor area) has the lowest p -value at 10^{-4} . Since we repeat the same test for 194 channels, we used the Bonferroni correction so that the significance level for each test was set to $0.05/194 = 2 \times 10^{-4}$. Indeed, only channel 29 (anterior supplementary motor area) survived the stringent threshold after the Bonferroni correction.

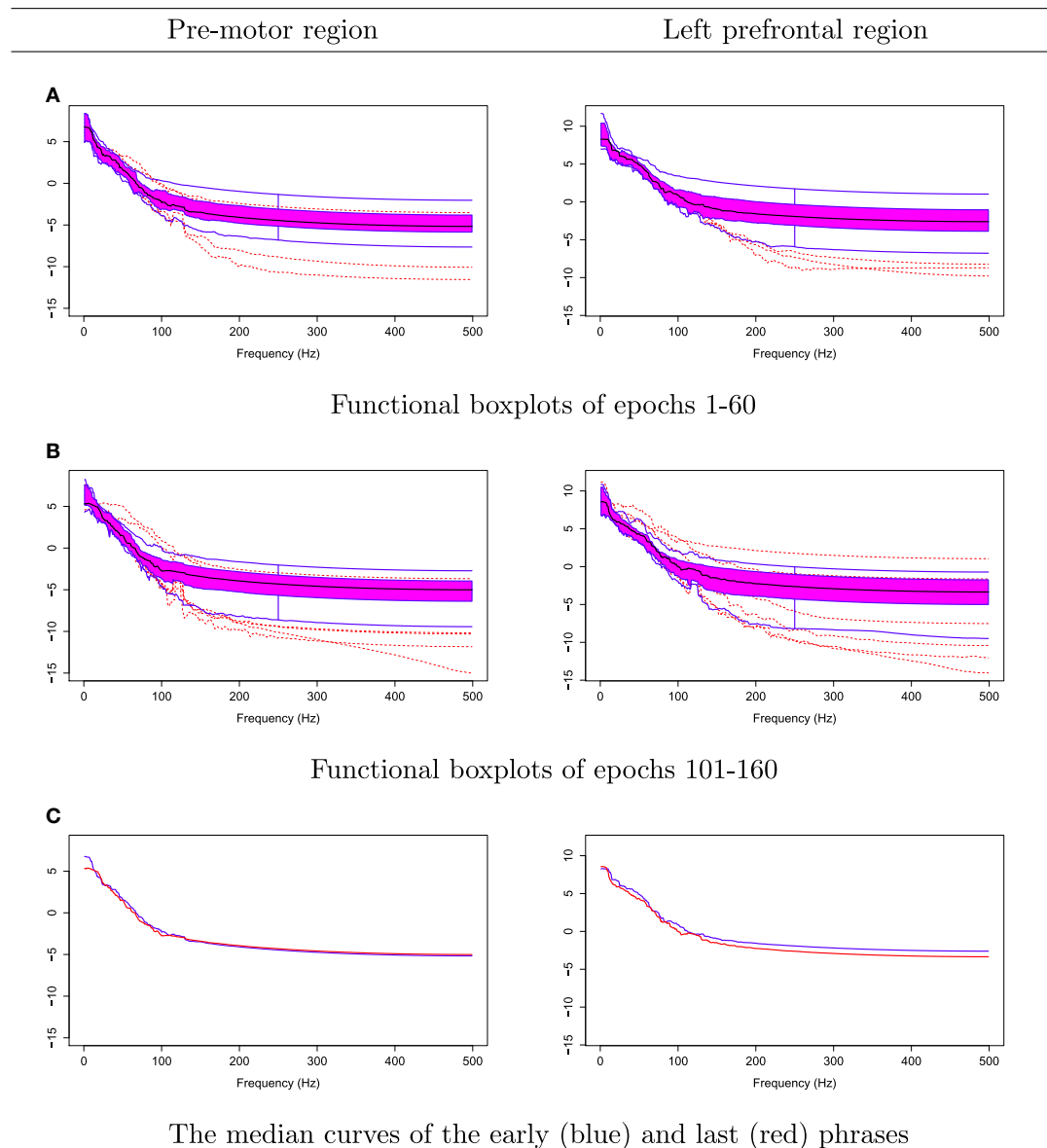


FIGURE 12 | Comparing median curves of the early and last phrases from pre-motor region and left pre-frontal region.

The tests for temporal stationarity at each channel (local spatial tests) revealed several channels having a significant difference between the median curves of the early vs. last phases of the EEG recording. As a next step, we studied stationarity in each of 19 predefined regions of the cortex. In this analysis, the representative EEG signal for each region was obtained by averaging the EEG signal-epochs over all channels within each region.

The plots in **Figure 14** suggest that the median curves for the early vs. late phases of the EEG recording are similar for EEG signals recorded from channels that approximately overlie right pre-motor and anterior supplementary motor regions, but different in the right pre-frontal and left parietal regions. Indeed,

we conclude from the rank sum test that there is significant difference between the early vs. late phases in cluster of channels that approximately overlie the right pre-frontal ($p = 0.01$) and the parietal regions ($p = 0.029$). We found that the right pre-frontal region is significantly non-stationary (i.e., early and late phases differ) at level 0.05 (see **Figure 15**). This result overlaps with the channel-specific tests, in which several of the channels identified to be non-stationary in the single channel tests are included in the predefined right pre-frontal region. In contrast, while the cluster of electrodes that overlie the left parietal region was found to be non-stationary in the region-by-region tests, none of the 18 channels that were identified to be non-stationary in the single channel tests are part of the left parietal cluster. Therefore, the

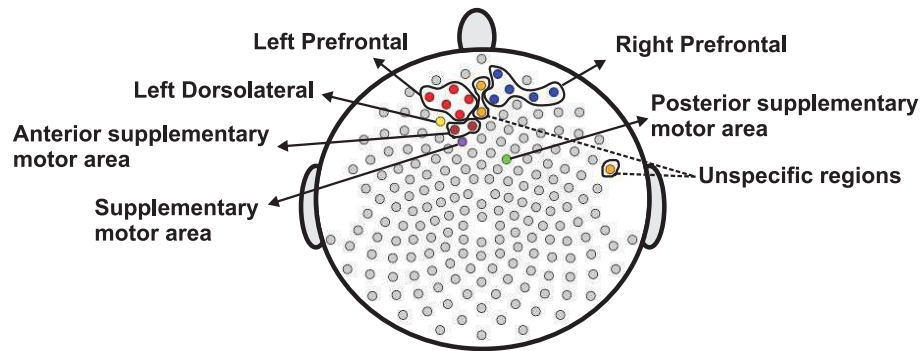


FIGURE 13 | Color circles represent channels, which have significant difference between the median curve of first 60 epochs and the median curve of last 60 epochs at $\alpha = 0.05$. Gray circles

represent channels which do not have significant difference between the median curve of first 60 epochs and the median curve of last 60 epochs at $\alpha = 0.05$.

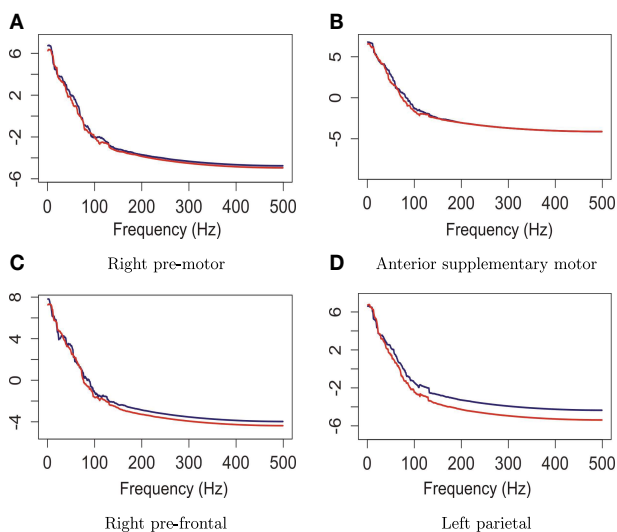


FIGURE 14 | Median curves of the early phase (first 60 epochs, in blue) and the late phase (last 60 epochs, in red) in the right pre-motor, anterior supplementary motor, right pre-frontal and left parietal regions.

additional averaging step across group of channels may improve signal-to-noise in this type of analysis. A similar phenomenon was also noted for predefined clusters of electrodes overlying at the left pre-frontal region.

4.4. The Variation of Spectral Power at Each Frequency Band Across the Entire Cortex

Our goal here is to test whether the spectral power at each frequency band differed across the cortical surface. We first computed the estimate of the spectral power for each channel at each epoch. Starting with the delta band, for each epoch

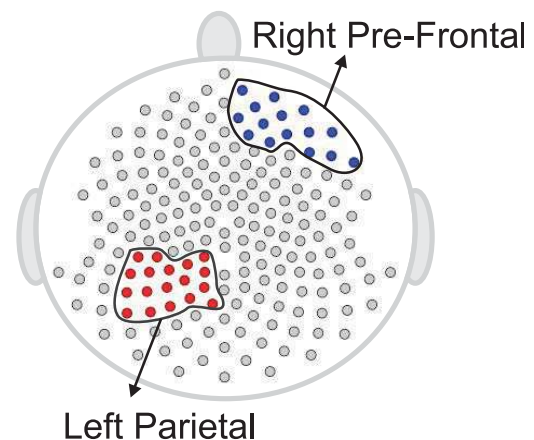
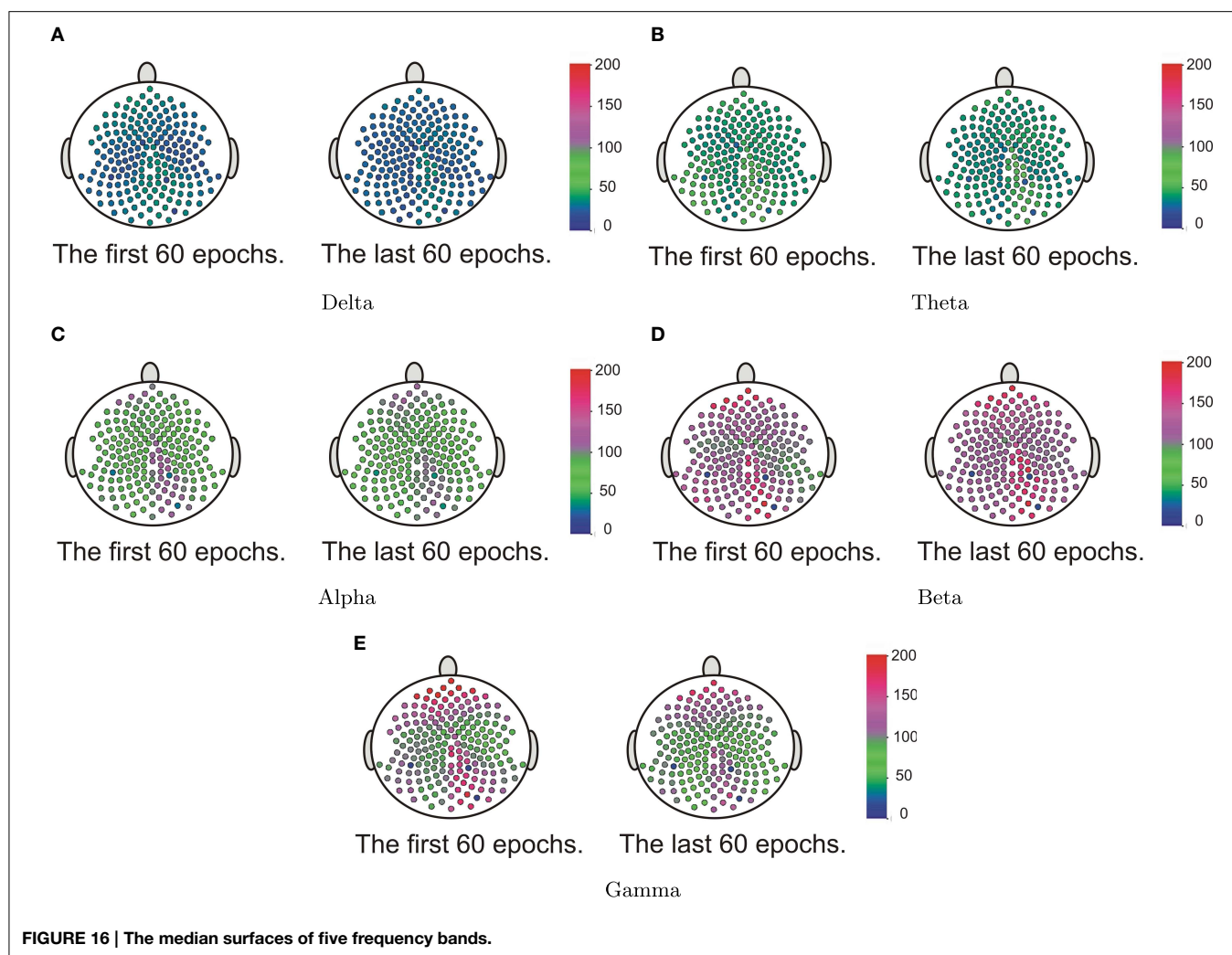


FIGURE 15 | Testing for difference between the early and late phases of the resting-state for each region. The right pre-frontal regions (blue circles) and the left parietal regions (red circles) exhibit significant non-stationarity at level 0.05.

we construct a 2 – D surface plot of the delta power across the entire cortical surface of 194 channels. These surfaces were then grouped according to the early and late phases of the resting-state. We then applied the surface boxplot method for each frequency band to obtain the median surfaces. In **Figure 16**, we present the median surface for five frequency bands in the early and late phases. The color blue represents the low spectral power while red is for high power. In **Figure 16**, it is interesting that even during resting-state there is relatively high spectral power at the beta and gamma bands—which are both associated with higher cognitive processing (Engel and Fries, 2010).



The next step is to test for differences between the early and late phases of the EEG recording for each of the five frequency bands of interest. Using the rank sum test, the delta and alpha bands do not have significant difference between the early and late phases. However, theta, beta and gamma bands show significant differences. In **Figure 17**, the colored regions indicate significant differences between the first and last phases while the gray color regions indicate no significant differences between these two phases. For the theta band, the rank sum test rejected the null hypothesis at only one region which is the cluster of electrodes overlying anterior supplementary motor. For the beta band, the rank sum test identified differences at the left medial parietal region. For the gamma band, there were 13 regions (out of 19) with significant difference between the early and late phases. Since the gamma band is wider than other bands, an estimated spectrum powers' variation across channel in gamma band is expected to be smaller than the estimated spectrum powers' variation in other bands. In Section 4.3, we tested the stationarity for each region. **Figure 15** shows two regions, namely, the right pre-frontal and left parietal, which

are significantly non-stationary across all frequencies between the early and late phases. **Figure 17** shows that the cluster of electrodes overlying the left parietal region exhibits significant non-stationarity in the beta and gamma bands while the cluster of electrodes overlying the right pre-frontal region is significantly non-stationary only in the gamma band.

5. Conclusion

This study has extended the use of the classical boxplot to FBP, which is a new visualization tool to analyze functional neuroimaging data, including EEG. The primary findings from the current study demonstrate the FBP is useful for both characterizing the spectral distribution of both simulated and real EEG data and identifying potential outliers in a continuous EEG signal.

In the current implementation of the FBP, ranked sample curves are used to characterize the EEG spectrum by defining a 50% central region, a median curve, and maximum and minimum non-outlying curves. Thus, the shape, size, and length

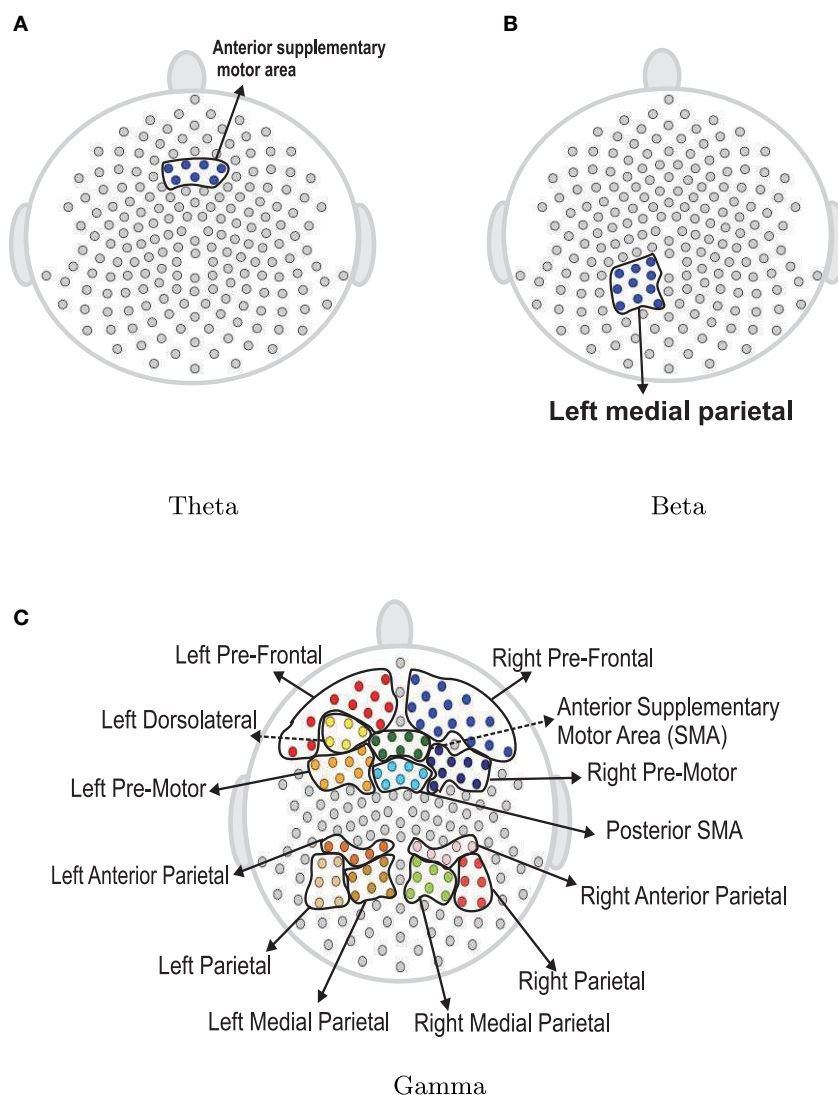


FIGURE 17 | Rank sum test results for regions. Color circles in a bounded curve represents a region, which has a significant difference between the first 60 epochs and the last 60 epochs.

of the FBP can be used to characterize the distribution of the dataset, including the skewedness and degree of variability of the EEG recording. Therefore, potential application of the FBP in this context includes comparing FBP derived from EEG recordings before and after an experimental intervention (e.g., across a period of motor skill training), comparing mean FBP derived from EEG recordings in healthy and diseased experimental groups, and comparing mean FBP derived from EEG recordings during resting-state vs. task.

An additional use of the FBP, as demonstrated by the current results, is to identify potential outliers of the EEG recording. Extra-brain artifacts, including eye blinks, eye movements, heart rhythms measured at pulse points downstream, and muscle movements can cause large deviations in the EEG signal, and represent a significant hurdle in EEG signal processing (Delorme et al., 2007). As a method for identifying

outliers in the EEG signal, the FBP could be used to rapidly identify periods of an EEG recording that show high likelihood for contamination by artifacts. In clinical applications, the continuous EEG recording has demonstrated promise as a method for monitoring neural function in patients who have compromised level of consciousness (Fyntanidou et al., 2012) or changes in neural function in patients undergoing neurosurgical interventions (de Vos et al., 2008). The use of FBP to identify outliers in the EEG recording represents a novel method for determining periods of the EEG recording that represent changes in consciousness in patients with a compromised level of consciousness, or for determining changes in neural function across neurosurgical intervention.

The current study also presents an application of the FBP to examine resting-state EEG data acquired from a single

individual by comparing EEG signals acquired during early vs. late phases of the 3 min EEG recording. This result has important implications for resting-state studies of neural activity, as many neuroimaging studies that examine resting-state brain function assume resting-state neural activity to be static. However, recent studies that examine dynamic changes in resting-state neural activity suggest momentary change in cognitive processes can cause non-stationarity in resting-state function (Chang and Glover, 2010; Hansen et al., 2015). In contrast, the current results show that the majority of channels demonstrate stationarity across the recording period, and provide support for the assumption that the average EEG signal is static across a 3 min EEG recording. Combined with previous findings, the current results suggest that while momentary changes in cognitive processes result in non-stationary fluctuations of the time series, when averaged across a 60 s subset of the complete 3 min EEG recording, the EEG signal is relatively static. This is supported by the current results that show channels which demonstrate non-stationarity of the EEG signal when comparing early and late phases of the recording include electrodes that overlie the right prefrontal region, which is associated with higher-order cognitive processes (Logue and Gould, 2014). Thus, the assumption of stationarity in resting-state functional neuroimaging studies may

be more appropriate for non-cognitive networks, including the motor network. Regardless, further work is needed to determine the minimal time-frame in which EEG signal demonstrate stationarity.

Additional future work is focused on developing a new method for computing confidence bands for the median curve. This method needs to consider the data as a whole. One possible approach is a re-sampling method, in which the notion of band depth is used to construct a 95% confidence band. A potential limitation of the re-sampling method is that there is the potential for multiple curves demonstrating ties with respect to band depth, thus affecting the resultant confidence band. One of the assumptions of the current smoothed periodogram method is that the log bias-corrected periodogram is an unbiased estimator of spectrum. Future work will provide further investigation of this assumption as the current method includes several levels of periodogram manipulation, including smoothing with the gamma generalized crossvalidation, log transformation, and correction by adding Euler Mascheroni constant. In conclusion, the current study presents a novel implementation of the FBP and demonstrates promise as a method for exploratory analysis of complex, high-dimensional neuroimaging datasets, including EEG data.

References

- Bohlin, T. (1973). Comparison of two methods of modeling stationary EEG signals. *IBM J. Res. Dev.* 17, 194–205.
- Bressler, S. L., and Freeman, W. J. (1980). Frequency analysis of olfactory system EEG in cat, rabbit, and rat. *Electroencephalogr. Clin. Neurophysiol.* 50, 19–24.
- Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*, Vol. 36. Philadelphia, PA: SIAM.
- Brockwell, P. J., and Davis, R. A. (2009). *Time Series: Theory and Methods*. New York, NY: Springer Science & Business Media.
- Chang, C., and Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50, 81–98. doi: 10.1016/j.neuroimage.2009.12.011
- de Vos, C. C., van Maarseveen, S. M., Brouwers, P. J., and van Putten, M. J. (2008). Continuous eeg monitoring during thrombolysis in acute hemispheric stroke patients using the brain symmetry index. *J. Clin. Neurophysiol.* 25, 77–82. doi: 10.1097/WNP.0b013e31816ef725
- Delorme, A., Sejnowski, T., and Makeig, S. (2007). Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* 34, 1443–1449. doi: 10.1016/j.neuroimage.2006.11.004
- Engel, A. K., and Fries, P. (2010). Beta-band oscillations-signalling the status quo? *Curr. Opin. Neurobiol.* 20, 156–165. doi: 10.1016/j.conb.2010.02.015
- Fan, J., and Kreutzberger, E. (1998). Automatic local smoothing for spectral density estimation. *Scand. J. Stat.* 25, 359–369.
- Fyntanidou, B., Grosomanidis, V., Aidoni, Z., Thoma, G., Giakoumis, M., Kiurzieva, E., et al. (2012). “Bispectral index scale variations in patients diagnosed with brain death,” in *Transplantation Proceedings*, Vol. 44 (New York, NY: Elsevier), 2702–2705.
- Gao, H.-Y. (1997). Choice of thresholds for wavelet shrinkage estimate of the spectrum. *J. Time Ser. Anal.* 18, 231–251.
- Genton, M. G., Johnson, C., Potter, K., Stenchikov, G., and Sun, Y. (2014). Surface boxplots. *Stat. (Int. Stat. Inst.)* 3, 1–11. doi: 10.1002/sta4.39
- Hansen, E. C., Battaglia, D., Spiegler, A., Deco, G., and Jirsa, V. K. (2015). Functional connectivity dynamics: modeling the switching behavior of the resting state. *Neuroimage* 105, 525–535. doi: 10.1016/j.neuroimage.2014.11.001
- Isaksson, A., Wennberg, A., and Zetterberg, L. H. (1981). Computer analysis of EEG signals with parametric models. *Proc. IEEE* 69, 451–461.
- Jain, S., and Deshpande, G. (2004). “Parametric modeling of brain signals,” in *Proceedings of Technology for Life: North Carolina Symposium on Biotechnology and Bioinformatics, 2004* (Washington, DC: IEEE), 85–91.
- Joundi, R. A., Jenkinson, N., Brittain, J.-S., Aziz, T. Z., and Brown, P. (2012). Driving oscillatory activity in the human cortex enhances motor performance. *Curr. Biol.* 22, 403–407. doi: 10.1016/j.cub.2012.01.024
- Krystal, A. D., Prado, R., and West, M. (1999). New methods of time series analysis of non-stationary EEG data: eigenstructure decompositions of time varying autoregressions. *Clin. Neurophysiol.* 110, 2197–2206.
- Lee, T. C. (1997). A simple span selector for periodogram smoothing. *Biometrika* 84, 965–969.
- Logue, S. F., and Gould, T. J. (2014). The neural and genetic basis of executive function: attention, cognitive flexibility, and response inhibition. *Pharmacol. Biochem. Behav.* 123, 45–54. doi: 10.1016/j.pbb.2013.08.007
- López-Pintado, S., and Romo, J. (2009). On the concept of depth for functional data. *J. Am. Stat. Assoc.* 104, 718–734. doi: 10.1198/jasa.2009.0108
- Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr. Clin. Neurophysiol.* 86, 283–293.
- Makeig, S., Bell, A. J., Jung, T.-P., and Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Adv. Neural Inf. Process. Syst.* 8, 145–151.
- McCullagh, P., and Nelder, J. A. (1989). *Generalized Linear Models*. Vol. 37. (Boca Raton, FL: RC Press), 33–34.
- Nunez, P. L., and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. New York, NY: Oxford University Press.
- Ombao, H. C., Raz, J. A., Strawderman, R. L., and Von Sachs, R. (2001). A simple generalised crossvalidation method of span selection for periodogram smoothing. *Biometrika* 88, 1186–1192. doi: 10.1093/biomet/88.4.1186
- Pawitan, Y., and O'sullivan, F. (1994). Nonparametric spectral density estimation using penalized whittle likelihood. *J. Am. Stat. Assoc.* 89, 600–610.

- Percival, D. B., and Walden, A. T. (1993). *Spectral Analysis for Physical Applications*. Cambridge, UK: Cambridge University Press.
- Pfurtscheller, G., and Aranibar, A. (1979). Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroencephalogr. Clin. Neurophysiol.* 46, 138–146.
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.
- Roopun, A. K., Middleton, S. J., Cunningham, M. O., LeBeau, F. E., Bibbig, A., Whittington, M. A., et al. (2006). A beta2-frequency (20–30 Hz) oscillation in nonsynaptic networks of somatosensory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15646–15650. doi: 10.1073/pnas.0607443103
- Shumway, D. S., and Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*, Vol. 3. New York, NY: Springer.
- Srinivasan, R., and Deng, S. (2012). “Multivariate spectral analysis of EEG: power, coherence and second-order blind identification,” in *Biosignal Processing: Principles and Practices*, eds H. Liang, J. D. Bronzino, and D. R. Peterson (Boca Raton, FL: CRC Press), 3-1–3-24. doi: 10.1201/b12941-4
- Sun, Y., and Genton, M. G. (2011). Functional boxplots. *J. Comput. Graph. Stat.* 20, 316–334. doi: 10.1198/jcgs.2011.09224
- Sun, Y., and Genton, M. G. (2012a). Adjusted functional boxplots for spatio-temporal data visualization and outlier detection. *Environmetrics* 23, 54–64. doi: 10.1002/env.1136
- Sun, Y., and Genton, M. G. (2012b). Functional median polish. *J. Agric. Biol. Environ. Stat.* 17, 354–376. doi: 10.1007/s13253-012-0096-8
- Sun, Y., Genton, M. G., and Nychka, D. W. (2012). Exact fast computation of band depth for large functional datasets: how quickly can one million curves be ranked? *Stat. (Int. Stat. Inst.)* 1, 68–74. doi: 10.1002/sta4.8
- Wahba, G. (1980). Automatic smoothing of the log periodogram. *J. Am. Stat. Assoc.* 75, 122–132.
- Wu, J., Srinivasan, R., Kaur, A., and Cramer, S. C. (2014). Resting-state cortical connectivity predicts motor skill acquisition. *Neuroimage* 91, 84–90. doi: 10.1016/j.neuroimage.2014.01.026

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Ngo, Sun, Genton, Wu, Srinivasan, Cramer and Ombao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

OPEN ACCESS

Edited by:

Han Liu,
Princeton University, USA

Reviewed by:

Xi-Nian Zuo,
Chinese Academy of Sciences, China
Boris Bernhardt,
Max Planck Institute for Human
Cognitive and Brain Sciences,
Germany

***Correspondence:**

Paul M. Thompson,
USC Imaging Genetics Center, Keck
School of Medicine, University of
Southern California, 4676 Admiralty
Way, Suite 400, Marina del Rey,
CA 90292, USA
pthomp@usc.edu

[†]These authors have contributed
equally to this work.

[‡]Data used in preparing this article
were obtained from the Alzheimer's
Disease Neuroimaging Initiative (ADNI)
database (adni.loni.usc.edu). As such,
the investigators within the ADNI
contributed to the design and
implementation of ADNI and/or
provided data but most of them did
not participate in this analysis or
writing this report. A complete list of
ADNI investigators can be found at:
[http://adni.loni.usc.edu/wp-content/
uploads/how_to_apply/ADNI_
Acknowledgement_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 15 April 2015

Accepted: 10 July 2015

Published: 24 July 2015

Citation:

Zhan L, Liu Y, Wang Y, Zhou J,
Jahanshad N, Ye J and Thompson
PM (2015) Boosting brain
connectome classification accuracy in
Alzheimer's disease using higher-order
singular value decomposition.
Front. Neurosci. 9:257.
doi: 10.3389/fnins.2015.00257

Boosting brain connectome classification accuracy in Alzheimer's disease using higher-order singular value decomposition

Liang Zhan^{1†}, Yashu Liu^{2†}, Yalin Wang², Jiayu Zhou², Neda Jahanshad¹, Jieping Ye^{3,4}
and Paul M. Thompson^{1*} for the Alzheimer's Disease Neuroimaging Initiative (ADNI)[‡]

¹ Imaging Genetics Center, Keck School of Medicine, University of Southern California, Marina del Rey, CA, USA, ² School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA, ³ Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA, ⁴ Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA

Alzheimer's disease (AD) is a progressive brain disease. Accurate detection of AD and its prodromal stage, mild cognitive impairment (MCI), are crucial. There is also a growing interest in identifying brain imaging biomarkers that help to automatically differentiate stages of Alzheimer's disease. Here, we focused on brain structural networks computed from diffusion MRI and proposed a new feature extraction and classification framework based on higher order singular value decomposition and sparse logistic regression. In tests on publicly available data from the Alzheimer's Disease Neuroimaging Initiative, our proposed framework showed promise in detecting brain network differences that help in classifying different stages of Alzheimer's disease.

Keywords: Alzheimer's disease, mild cognitive impairment, diffusion MRI, connectome, high-order SVD, classification

Introduction

Alzheimer's disease (AD) is a chronic neurodegenerative disease that involves the accumulation of amyloid plaques and neurofibrillary tangles in the brain. The most common early symptom is difficulty remembering recent events (short-term memory loss). As the disease advances, symptoms often include problems with language, altered affect, disorientation, lack of motivation, problems with self-care, and behavioral abnormalities (Burns, 2009; Burns and Iliffe, 2009). As a patient's condition declines, they may withdraw from family and society. Gradually, more and more bodily functions are lost, ultimately leading to death. Although the speed of progression varies, the average life expectancy following diagnosis is 3–9 years (Querfurth and LaFerla, 2010; Todd et al., 2013). AD has a typical pattern of progression, with anatomical changes that correspond to the types and severity of symptoms. The symptoms, the order in which they appear, and the duration of each clinical stage vary from person to person. Disease progression can be divided into three main stages: normal controls (NC), mild cognitive impairment (MCI) and AD. All of these classifications are defined clinically based on behavioral and cognitive assessments, and although a person with MCI has elevated risk of developing AD, many people with MCI remain stable for some time or may develop other degenerative conditions pathologically distinct from AD, such as vascular dementia or fronto-temporal dementia.

NC represents the subset of the population who are aging normally, and do not have sufficiently severe symptoms to be considered cognitively impaired. MCI involves cognitive impairments, but at a level that is not significant enough to interfere with a person's daily activities (Petersen et al., 1999). MCI is often a transitional stage between normal aging and dementia: every year, around 10–15% of people with MCI progress to probable AD (Grundman et al., 2004). However, not all people with MCI deteriorate cognitively and some even improve. Effective and accurate diagnosis of Alzheimer's disease and its prodromal stage, MCI, are crucial for drug trials, given the urgent need for treatments to resist or slow disease progression.

Many neuroimaging studies have used anatomical measures derived from T1-weighted brain MRI, such as cortical thickness, and volumetric or shape measures of subregions of the brain, to differentiate AD or MCI from NC (Fan et al., 2008; Hua et al., 2008a,b; Gerardin et al., 2009; Magnin et al., 2009; Hua et al., 2010; Cuingnet et al., 2011; Westman et al., 2011; Hua et al., 2013; Gutman et al., 2015).

Moreover, measures derived from functional imaging or cerebrospinal fluid (CSF) assays have also been used to help classify individuals with cognitive impairment vs. healthy controls (De Santi et al., 2001; Morris et al., 2001; Bouwman et al., 2007; Mattsson et al., 2009; Shaw et al., 2009; Fjell et al., 2010). Diffusion weighted MRI is a non-invasive imaging technique that can provide clinical information on white matter integrity in a variety of diseases, such as schizophrenia (Zalesky et al., 2011), autism (Lewis et al., 2014), traumatic brain injury (Dennis et al., 2015b), and even in genetics (Jin et al., 2011, 2013) and sex difference (Jahanshad et al., 2011). The white matter integrity can be analyzed with both the tract-based analysis such as tract-based spatial statistics (Smith et al., 2006), fiber clustering (Jin et al., 2012, 2014), and the parcellation-based connectome analysis (Toga et al., 2012).

In particular, many studies have used diffusion-weighted MRI (DWI) to study AD and MCI. Demirhan et al. (2015) studied the added value of diffusion tensor derived measures, over and above structural MRI, and showed they provided added diagnostic accuracy for classification of disease stages (Demirhan et al., 2015). Nir et al. (2013) found that standard diffusion tensor derived measures were strongly correlated with several clinical ratings that are widely-used in AD research (MMSE, CDR-sob, and ADAS-cog) (Nir et al., 2013). When effect sizes were ranked, *mean diffusivity* (MD) measures tended to outperform fractional anisotropy (FA) measures for detecting group differences in tracts that pass through the temporal lobes and the left hippocampal component of the Cingulum. Diffusivity measures tended to detect the more subtle differences in MCI, even when comparisons of FA measures did not. Jin et al. (2015) also used various diffusion-derived measures to relate fornix degeneration with cognitive decline. MD was also shown to be more sensitive to group differences among AD, MCI, and normal controls than FA (Jin et al., 2015).

Several studies used the ADNI DWI scans to compute structural connectivity measures, including measures of the brain's network properties. Li et al. (2013) proposed a spectral

diffusional connectivity framework to explore the connectivity deficit in AD. Li et al. (2013) The framework was based on studying the eigenvalues of the Laplacian matrix of the diffusion tensor field at the voxel level. The peaks of the diffusional connectivity spectra were shifted in the AD group versus the normal controls. Prasad et al. (2015) ranked several connectivity measures, to see which ones best distinguished AD from normal aging (Prasad et al., 2015). Graph-based network measures—such as small-world properties, clustering, and modularity—helped in differentiating diagnostic subgroups relative to just using the raw connectivity matrices; there was also additional predictive value in computing a very dense connectivity matrix to represent the structural connectivity between all adjacent voxels in the image. This approach, known as “flow-based connectivity analysis” complemented the more standard analysis of large-scale tracts interconnecting cortical and subcortical regions of interest. Even so, brain networks and their features depend to some extent on the choice of field strength (Zhan et al., 2013c; Dennis et al., 2014), scanners (Zhan et al., 2014a), feature space (Zhan et al., 2014b), imaging acquisition parameters (Zhan et al., 2012), fiber tracking parameters (Dennis et al., 2015a), fiber tracking algorithms used to infer the trajectories of pathways in the brain (Zhan et al., 2013b, 2015a,b). Dozens of tractography algorithms are now available (Conturo et al., 1999; Mori et al., 1999; Basser et al., 2000; Lazar et al., 2003; Parker et al., 2003; Behrens et al., 2007; Aganj et al., 2011) yielding visually very different brain networks.

For this study, we adopted the tensor-based *fiber assignment by continuous tracking* (FACT) algorithm (Mori et al., 1999) to compute structural brain networks in a cohort of elderly patients with various levels of cognitive impairment (none, mild, severe). Tensor-based FACT can yield false positive fibers that may add noise to the computed network properties, but it is still one of the most widely used tractography algorithms due to it being simple and flexible. Here we propose a novel framework for network classification, with the goal of improving diagnostic classification by combining diffusion and structural MRI. We also set out to show how this new framework could be applied to networks that might contain false positive fibers (such as those derived from FACT) and used for differentiating different stages of cognition in the stages of Alzheimer's disease.

Methods

Figure 1 summarizes our proposed framework for brain network classification using higher order singular value decomposition (HO-SVD) and sparse logistic regression (Sparse LG). Its two component techniques are explained below.

HO-SVD

Singular value decomposition (SVD) is a powerful tool for dimension reduction that is widely used in machine learning and data mining. The SVD of a matrix $X \in \mathbb{R}^{n \times m}$ is given by $X = U\Sigma V^T$, where $U \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{m \times m}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{n \times m}$ is a rectangular diagonal matrix. The diagonal entries of Σ , known as singular values, are non-negative and assumed to be in descending order.

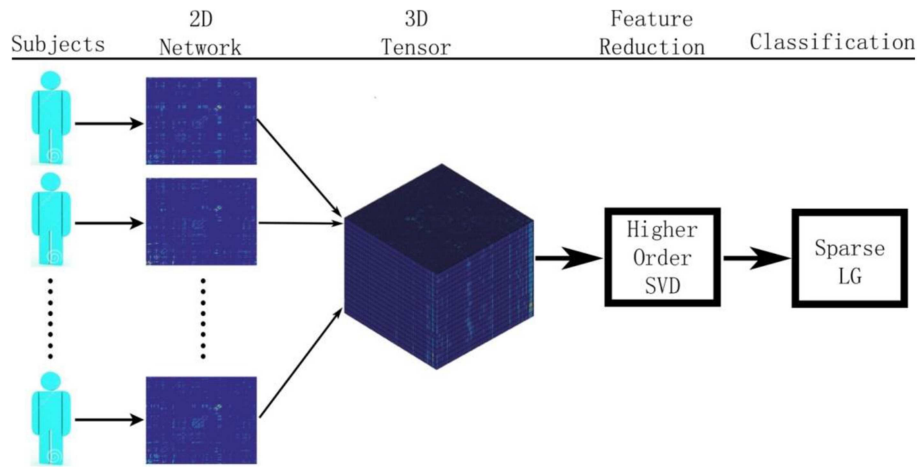


FIGURE 1 | Here we show the workflow used in this paper to classify patients based on their brain structural networks. We model brain networks as connectivity matrices, and then stack

them up, across subjects, as a 3D tensor. We then perform feature reduction and use sparse methods for diagnostic classification.

The higher order SVD (HO-SVD) is one common generalization of SVD from matrices to tensors (De Lathauwer et al., 2000). In HO-SVD, a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is decomposed as

$$\mathcal{X} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \dots \times_N U^{(N)}$$

in which

- (1) $U^{(k)} \in \mathbb{R}^{I_k \times I_k}$, $k = 1, \dots, N$ are orthogonal matrices where the i th column of $U^{(k)}$ is the i th k -mode singular vector.
- (2) $\mathcal{S} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is the core tensor which is of the same size as \mathcal{X} , and has the following properties:
 - For any $1 \leq k \leq N$, let \mathcal{S}_{i_k} and \mathcal{S}_{j_k} be the subtensors obtained by fixing the k th index to i_k and j_k , $1 \leq i_k, j_k \leq I_k$, then $\langle \mathcal{S}_{i_k}, \mathcal{S}_{j_k} \rangle = 0$ for $i_k \neq j_k$;
 - For $1 \leq k \leq N$,

$$\|\mathcal{S}_{i_k=1}\| \geq \|\mathcal{S}_{i_k=2}\| \geq \dots \geq \|\mathcal{S}_{i_k=I_k}\| \geq 0$$

The Frobenius-norms $\|\mathcal{S}_{i_k=i}\|$, $1 \leq i \leq I_k$ are the k -mode singular values.

The k th mode singular matrix $U^{(k)}$ can be obtained as the left singular matrix of the k th mode unfolding matrix of tensor \mathcal{X} . After obtaining all N singular matrices $U^{(1)} \dots U^{(N)}$, the core tensor \mathcal{S} is given by

$$\mathcal{S} = \mathcal{X} \times_1 U^{(1)T} \times_2 U^{(2)T} \dots \times_N U^{(N)T}$$

Inspired by the dimension reduction via SVD in the 2D case, we propose to reduce the dimensions of diffusion MRI derived brain networks, using higher order SVD (HO-SVD).

Similar to the matrix case, the ordering assumption for tensor singular values suggests that most of the information contained in a tensor may be expressed by the first few “components.” Let

the first mode of data tensor \mathcal{X} correspond to the sample size n (i.e., $I_1 = n$) and the remaining modes correspond to feature dimensions. Then, by keeping the largest R_1, \dots, R_N singular values for each mode, a reduced tensor with size $n \times R_2 \times R_3 \times \dots \times R_N$ can be obtained by

$$\tilde{\mathcal{X}} = \tilde{\mathcal{S}} \times_1 \tilde{U}^{(1)}$$

where $\tilde{\mathcal{S}} = \mathcal{X} \times_1 \tilde{U}^{(1)T} \times_2 \tilde{U}^{(2)T} \dots \times_N \tilde{U}^{(N)T}$ is the core tensor with the first R_1, R_2, \dots, R_N singular values kept for each mode, and $\tilde{U}^{(k)} \in \mathbb{R}^{I_k \times R_k}$, $1 \leq k \leq N$. The proposed dimension reduction of the tensor is also analogous to principal components analysis (Mocks and Verleger, 1986) for a matrix input. Instead of the original tensor, we propose to use the reduced tensor $\tilde{\mathcal{X}}$ as the new input data for classification. Figure 2 illustrates the basic idea of HO-SVD and feature reduction.

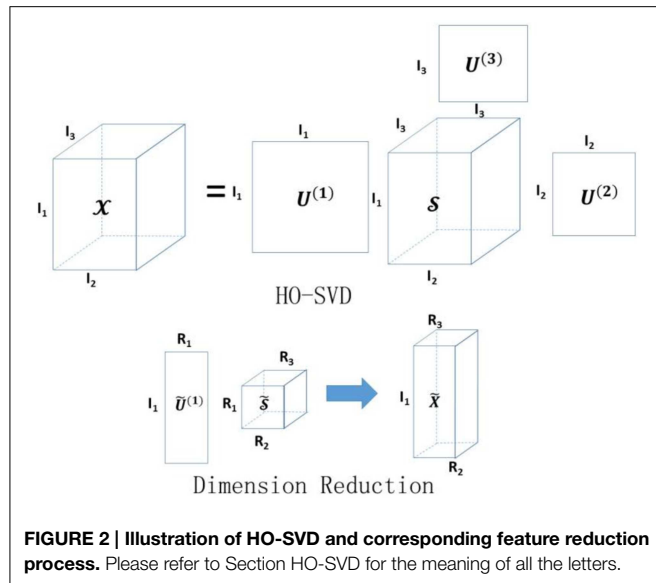
Sparse Logistic Regression

Let $x \in \mathbb{R}^m$ be a sample vector and $y \in \{-1, +1\}$ be a binary outcome. The logistic regression model is given by:

$$\text{Prob}(y|x) = \frac{1}{1 + \exp(-y(x^T w + c))}$$

where $w \in \mathbb{R}^m$ and $c \in \mathbb{R}$ are coefficients, and $\text{Prob}(y|x)$ is the posterior probability. Given n samples $\{x_1, x_2, \dots, x_n\}$, the empirical logistic loss is measured by the negative log-likelihood and the average logistic loss is given by

$$\begin{aligned} \mathcal{L}(w, c) &= -\frac{1}{n} \log \prod_{i=1}^n \text{Prob}(y_i | x_i) \\ &= \frac{1}{n} \sum_{i=1}^n \log(1 + \exp(-y_i(x_i^T w + c))) \end{aligned}$$



The unknown coefficients w and c can be computed by minimizing the logistic loss, which involves a smooth convex optimization problem. However, when dimension m is far larger than the sample size n , solving the logistic regression problem is ill-posed, and the learned model may suffer from the over-fitting problem.

Sparse logistic regression embeds the feature selection into classification using the Lasso penalty (Tibshirani, 1996, 2011) which results in a sparse solution for w . The sparse logistic regression problem is formulated as:

$$\min_{w, c} \mathcal{L}(w, c) + \lambda \|w\|_1$$

where the l_1 norm of w , i.e., $\|w\|_1$, is the Lasso penalty and $\lambda > 0$ is the regularization parameter that controls the sparsity of the solution.

Experiments

Subject Demographics and Image Acquisition

We analyzed brain imaging data from 202 participants in ADNI2, the second stage of the North American Alzheimer's Disease Neuroimaging Initiative (ADNI) (<http://adni.loni.usc.edu>). Participant information including performance on the mini-mental state exam (MMSE) and the clinical dementia rating (CDR) are summarized in **Table 1**. Subjects are divided into three broad diagnostic categories based on the standard criteria outlined on the ADNI website (<http://www.adni-info.org/scientists/ADNIGrant/ProtocolSummary.aspx>).

- Normal Control (NC) subjects: MMSE scores between 24–30 (inclusive), CDR of 0, and non-depressed.
- MCI participants: MMSE scores between 24 and 30 (inclusive), a memory complaint, have objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical

TABLE 1 | Summary of ADNI data used in this study.

	AD	MCI	NC	Total
Number	39	112	51	202
Age (y)	75.56 ± 9.11	71.68 ± 9.89	69.69 ± 15.43	71.92 ± 11.54
Sex	25M	71M	22M	118M

Memory II, a CDR of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia.

- AD subjects: MMSE scores between 20 and 26 (inclusive), CDR of 0.5 or 1.0, and meeting NINCDS/ADRDA criteria for probable AD.

T1-weighted and diffusion MRI were acquired from each participant using 3-tesla GE Medical Systems scanners. 3D T1-weighted images were collected using spoiled gradient echo (SPGR) sequences with the following parameters: 256×256 acquisition matrix; voxel size = $1.2 \times 1.0 \times 1.0 \text{ mm}^3$; TI = 400 ms; TR = 6.98 ms; TE = 2.85 ms; flip angle = 11° . 5 T2-weighted volumes with no diffusion sensitization (b_0 images) and 41 diffusion-weighted volumes ($b = 1000 \text{ s/mm}^2$), were collected with the following parameters: 128×128 matrix; TR = 9050 ms, isotropic voxels, of size 2.7 mm; number of slices = 59; scan time = 9 min. Additional details of the protocols are available at http://adni.loni.usc.edu/wp-content/uploads/2010/05/ADNI2_GE_3T_T2_22.0_T2.pdf. The diffusion MRI protocol for ADNI was chosen after a detailed evaluation of different protocols that could be performed in a reasonable amount of time; we reported these comparisons previously (Jahanshad et al., 2010; Zhan et al., 2013a). All T1-weighted MR and DWI images were visually checked for quality assurance to exclude scans with excessive motion and/or artifacts, and all scans were included.

Network Computation

Each subject's brain network was computed with the method described in Zhan et al. (2013c). In brief, each subject's DWI was preprocessed (corrected for eddy current distortion and motion as well as removal of non-brain tissue) using the FSL toolbox (<http://fsl.fmrib.ox.ac.uk/>). Then, whole brain tractography was computed using tensor-based fiber assignment by continuous tracking (FACT) algorithm (Mori et al., 1999) implemented in diffusion toolkit (<http://trackvis.org/dtk/>). 113 cortical and subcortical regions-of-interest (ROIs) were defined using the Harvard Oxford Cortical and Subcortical probabilistic atlas (Desikan et al., 2006). For each pair of ROIs, the number of detected fibers connecting them was determined from the FACT tractography. A fiber was considered to connect two ROIs if it intersected both of them. This process was repeated for all ROI pairs, to compute a whole brain fiber connectivity matrix. This matrix is symmetric, by definition, and has a zero diagonal, i.e., we did not consider self-connections. **Figure 3** illustrates the overall process to compute the brain networks.

To avoid bias in subsequent analyses, we normalized each subject's matrix by dividing each entry by its maximum value,

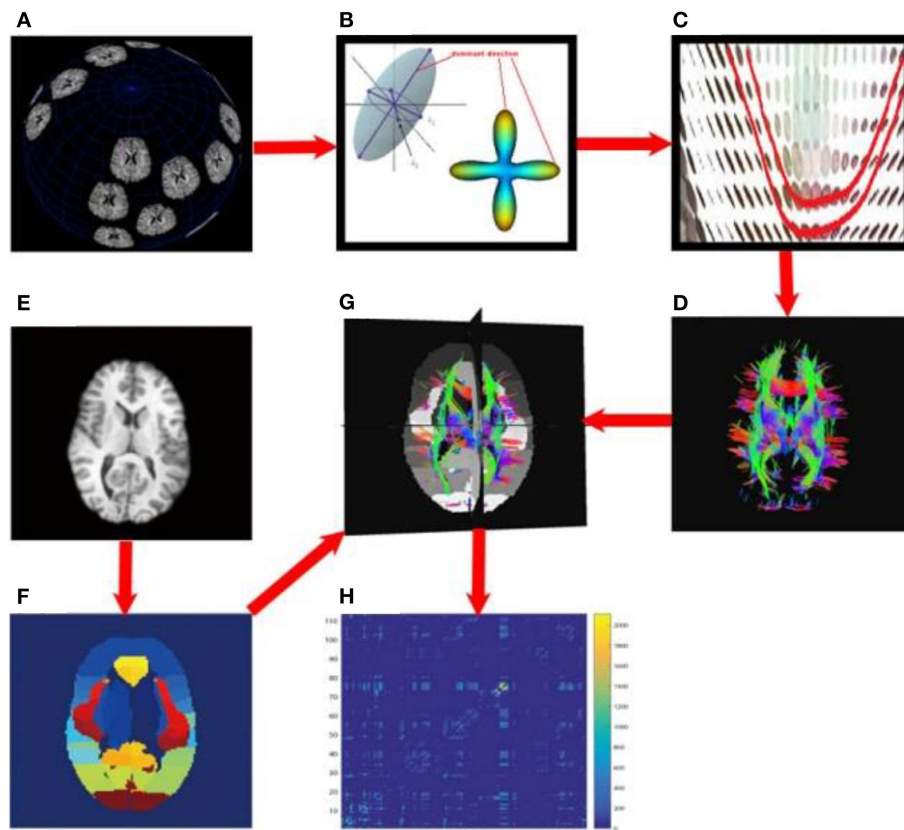


FIGURE 3 | Flowchart to compute structural brain networks. (A)

Diffusion MRI: the MR signal was sampled after applying gradients in a set of directions uniformly distributed on a spherical surface; (B) Modeling: the diffusion process was modeled using a tensor model, or by fitting orientation distribution functions, and then the dominant direction was identified; (C) Fiber tracking: a fiber streamline was generated, connecting as far as possible the dominant directions of neighboring voxels under some constraints (e.g., a threshold on the maximum turning angle); (D)

whole brain tractography, tracking fibers from a set of seeds across the whole brain; here, the color indicates the fiber directions, *red* for left and right, *blue* for superior and inferior, *green* for anterior and posterior; (E) T1-weighted brain MRI; (F) brain parcellation: here we defined 113 ROIs using the Harvard Oxford Cortical and Subcortical probabilistic atlas; (G) aligning the whole brain tractography and 113 ROIs; (H) the resulting un-normalized brain network, counting the number of detected fibers connecting each pair of ROIs.

as matrices derived from different subjects have different scales and ranges. This normalized network served as the input for the following analyses.

Network Analysis and Confounds Removing

For each of the 202 subjects' 113×113 normalized networks, we calculated standard network metrics. Five common global network measures, including *modularity* (MOD), *mean clustering coefficient* (MCC), *characteristic path length* (CPL), *global efficiency* (GLOB), and *small-worldness* (SW), were computed using the Brain Connectivity Toolbox (BCT) (Rubinov and Sporns, 2010). We used the weighted version of these measures. Definitions and mathematical equations for all of these metrics may be found at the BCT website (<https://sites.google.com/site/bctnet/>).

We used the generalized linear model (GLM) to remove confounds related to age and sex, across all subjects. Element-wise residual 113×113 networks were used as well as the residuals from the global network measures. From now on, we will refer to

these networks as the “GLM-adjusted” networks and the residuals from the global network measures as GLM-adjusted network measures.

Feature Extraction

Using the 113×113 GLM-adjusted networks computed in Section Network Analysis and Confounds Removing, we compared three feature extraction methods:

- (1) *Raw features*: for each subject, the feature vector was constructed by stacking all the entries of the upper triangular portion of the matrix (as the brain matrix is symmetric). So each subject has 6328 ($= 113 \times 112/2$) features. Thus, we obtained a matrix with 202 (subjects) by 6328 (features) as the input to Sparse LG.
- (2) *SVD*: We first built the raw feature matrix and then center this matrix by subtracting the means for each column. We use the top k principal components as the input for Sparse LG. The SVD method is essentially equivalent to PCA. Let us assume that the rows of a data matrix represent samples

and the columns of data matrix represent different features. One can always use SVD to decompose an arbitrary matrix to reduce the dimensionality, by using its top k singular values and left singular vectors. When the feature columns of the data matrix are all centered, it can be easily verified that SVD is exactly the same as PCA. In our paper, we center the data matrices first, so we are essentially comparing HO-SVD with PCA.

- (3) *HO-SVD*: We reduced the dimension of data tensor to $202 \times 15 \times 15$ by keeping the largest k singular values for each mode. Then, we constructed the feature vector for each subject by stacking the entries of the reduced data matrix. This constructed feature vector then serves as the input for Sparse LG.

Our empirical tests showed that the performances of SVD and HO-SVD are stable when k is set between 10 and 30. In this paper, we report the performances obtained after setting $k = 15$.

Experiment Design

Three comparisons, including (1) AD vs. MCI, (2) AD vs. NC, and (3) MCI vs. NC, were evaluated on the extracted features using three types of assessments:

- (1) *Assessment of element-wise brain connectivity matrices*: Each GLM-adjusted network cell value is compared across the different groups (AD, MCI, and NC).
- (2) *Assessment of global network measures*: Each of the five GLM adjusted network measures was compared across the AD, MCI, and NC groups.
- (3) *Assessment of feature extraction methods*: We will use the three feature extraction methods described in Section Feature Extraction to extract features from GLM-adjusted networks and conduct sparse logistic regression to evaluate the classification performance when classifying the different groups. The rows in the extracted data matrix correspond to samples, and the columns represent features. We first normalize each feature column of the input data matrix by subtracting the mean and dividing it by the standard deviation. As the outcome has an imbalanced distribution, we used undersampling techniques to mitigate the bias. Undersampling used in this paper is a special case of subsampling, and our method is essentially the bagging procedure where a bunch of models built on subsampled dataset are averaged. Different from a general subsampling strategy, undersampling requires the numbers of positive and negative samples to be the same. Thus, undersampling does not introduce any additional prior knowledge of the data distribution, so it is less likely to create classifiers that favor the majority class. For each evaluation procedure, we first randomly split data into two parts: first, 85% of the samples were used for training, and the remaining 15% of the samples were used for testing. We next ran a 5-fold cross-validation on the training data alone to select a model parameter (the LASSO parameter) and we then re-trained a model on the whole training data with the selected parameter to produce the final prediction model. We then made predictions on the test data using the

final model and evaluate the performance. We repeated the training/test procedure 20 times. We report the mean and standard deviations of the classification performances including measures of accuracy, sensitivity, specificity, and the area under the curve (AUC). The Sparse LG model was implemented using the Sparse Learning with Efficient Projections package (Liu et al., 2009).

Results and Discussions

Assessment of Element-Wise Brain Connectivity Matrices

After removing age and sex effects, the GLM-adjusted brain network are used to estimate differences among the different diagnostic groups. To quantify these differences, we conducted Student's t -tests on each cell of the GLM-adjusted network for the three different tests (AD vs. NC, AD vs. MCI and MCI vs. NC). Since there are 6328 ($= 113 \times 112/2$) cells in each GLM-adjusted network, a Bonferroni correction was adopted to account for multiple comparisons and the threshold for statistical significance was set to $0.05/6328 \approx 7.9 \times 10^{-6}$. **Figure 4** shows the highlighted P map from a Student's t -test. Red elements in the matrices represent the connections with uncorrected $P < 0.001$. White elements in the matrices indicate connections that differ significantly between groups after Bonferroni correction. It is interesting that in the comparison between MCI and NC, there are four connections with significant uncorrected P -values. These connections involve brain stem, left thalamus, left putamen, left superior temporal gyrus posterior division and left hippocampus. There are considerable literatures reporting the involvement of several of these regions in degenerative neurological disorders such as Alzheimer's Disease. For example, in 2009, Simic and his colleagues reported early changes in Alzheimer's disease in the serotonergic nuclei of the brain stem, even though the brain stem would not normally appear in the set of regions with preferential atrophy in AD Simic et al. (2009) Also, the reduced volume of putamen and thalamus have been reported in Alzheimer's Disease. de Jong et al. (2008) Even so, the hippocampus is more typically one of the first brain regions to be affected by Alzheimer's Disease. Our results indicate the connection patterns among these regions may also be affected by the disease. Thus, this result deserves further investigation.

However, no significant differences were detected on an element-wise level between MCI and NC, after correction, still suggesting that it is challenging differentiate these groups based on the GLM-adjusted networks. In contrast, there were 21 significant connections for the classification task of discriminating AD vs. NC and 7 significant connections for the task AD vs. MCI. These results are consistent with our previous studies (Zhan et al., 2015b), where we found that there is an approximate order of difficulty in these differentiating tasks, with the hardest task being: MCI vs. NC > AD vs. MCI > AD vs. NC. Furthermore, comparing the P map between AD vs. NC and AD vs. MCI in **Figure 4**, we did not find any points are repeated in both P maps, which suggests the raw brain network cell values may not be ideal for studying of the progressive process

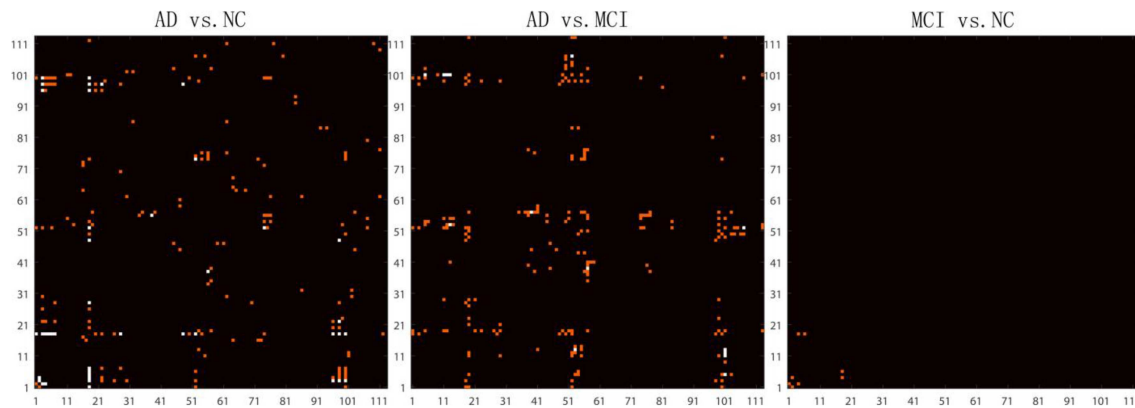


FIGURE 4 | Highlighted matrices showing Student *t*-test *P* maps for three diagnostic comparisons: left: AD vs. NC, middle: AD vs. MCI and right: MCI vs. NC are displayed. Each matrix is 113×113 , corresponding to 113 ROI connectivity pattern. The ROIs are indexed from 1 to 113. Please refer to Zhan et al. (2013c) for corresponding numbers. Each cell of the GLM-adjusted network represents the connectivity, after removing the effects of age and sex at each element. The red points in these matrices

highlight the location of uncorrected $P < 0.001$. Multiple comparisons were adjusted for by Bonferroni correction and the significance threshold was set to 7.9×10^{-6} . The white points in these matrices highlight the location of the significant differences (after Bonferroni correction) in the network cell between the groups. The greatest number of connections were different when comparing controls and AD, but no connections survive Bonferroni correction when testing differences between controls and MCI.

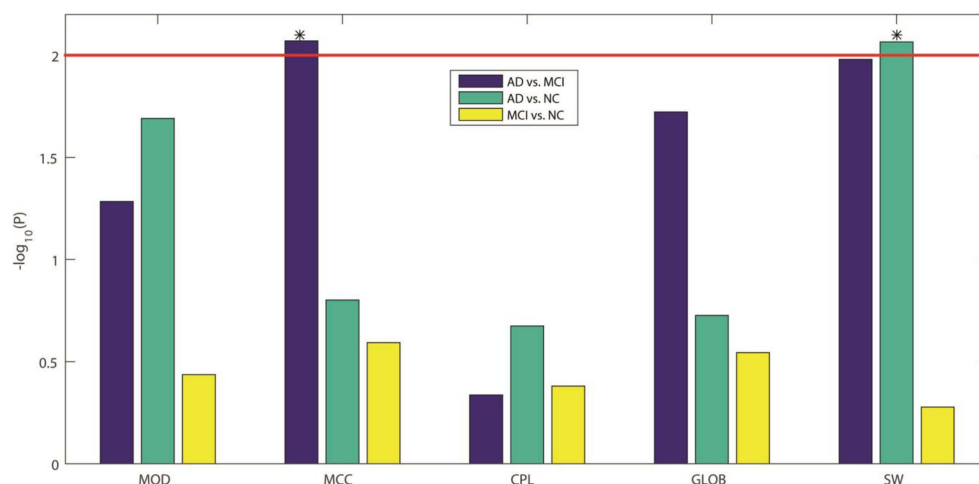


FIGURE 5 | $-\log_{10}(P)$ values from our three inter-group comparisons using Student's *t*-test for each of five standard global network measures. From left to right, the five network measures are: MOD, modularity; MCC, mean clustering coefficient; CPL, characteristic path length; GLOB, global efficiency; and SW, small worldness, respectively. The colors indicate which groups are being compared: blue

for AD vs. MCI, green for AD vs. NC, and yellow for MCI vs. NC. Bonferroni correction was used to account for multiple comparisons, so the adjusted threshold is 2, indicated by the red line in the figure. Only values above the line are statistically significant given this threshold. Our results show that only MCC can differentiate AD from MCI and only SW can differentiate AD from NC.

of Alzheimer's disease. Thus, we went on to investigate network measures, in the next section.

Assessment of Global Network Measures

Here we compared the five GLM-adjusted global network measures (MOD, MCC, CPL, GLOB, and SW) between the diagnostic groups. **Figure 5** shows the $-\log_{10}(P)$ values computed from the *t*-test between groups for these network measures, in each of the three diagnostic tasks. We again adopted a Bonferroni correction to account for the 5 comparisons in

each task, so the adjusted significance threshold at the $\alpha = 0.05$ level is 0.01 ($=0.05/5$). We marked this adjusted threshold with a red horizontal line [$2 = -\log_{10}(0.01)$]. Our results showed that SW can be used to differentiate AD from NC while MCC can differentiate AD from MCI. As in Section Assessment of Element-wise Brain Connectivity Matrices, no measure was able to statistically distinguish between MCI and NC, which again indicates that more sensitive brain imaging features are needed to distinguish MCI from NC, at least in samples of this size.

Assessment of Feature Extraction Methods

Here we conducted more advanced feature extraction methods and classification techniques as described in Sections Feature Extraction and Experiment Design to better distinguish diagnostic classes. We firstly applied McNemar's test (McNemar, 1947) to confirm there are significant differences between different feature extraction methods. (Please refer to Supplementary Table for the results of the McNemar's test). Then we started to rank these feature extraction methods. **Table 2** summarizes the classification performance, and **Table 3** lists the Student's *t*-test *P*-values. The column *SVD* > *Raw* in **Table 3** indicates statistical differences in classification performance of the SVD and Raw feature sets; there was no detectable difference in classification performance between these feature sets for all three diagnostic tasks. Therefore, for this particular set of tasks and this dataset, performing SVD does not improve classification performance. SVD can reduce the dimension of the data, perhaps also reducing the noise, it may still discard useful information that may be vital for classification.

A similar result is seen when using HO-SVD in the classification task AD vs. NC. For task AD vs. NC in **Table 3**, both SVD and HO-SVD feature sets performed similarly to the raw feature set. One possible explanation for this could be that the AD and NC groups are the most biologically different, so they are easier to differentiate than the other two, as is evident in **Table 2**. The classification performance is already quite good for raw features and there is little room for improvement.

On the other hand, our proposed HO-SVD had a significant advantage in accuracy for the other two differentiation tasks. As listed in **Table 3**, HO-SVD performed significantly better than raw features for accuracy and specificity for AD vs. MCI; and in accuracy and sensitivity for the task, MCI vs. NC.

Brain networks derived from FACT-based tractography often include a substantial number of false positive fibers generated. Our experimental results suggest that HO-SVD is quite effective in handling feature reduction for these noisy networks, especially

in the more challenging task of differentiating cognitively healthy controls from MCI.

Alzheimer's disease involves structural atrophy detectable on MRI, as well as pathological amyloid depositions and metabolic alterations in the brain. In this study, we compared the brain network properties in different stages of Alzheimer's disease using different analysis methods. In our first two assessments using element-wise brain connectivity matrices and global network measures, respectively, we were unable to differentiate the diagnostic classes MCI and NC. But while within our HO-SVD framework, the classification performance was significantly improved compared to using raw features. The choice of tractography algorithms can also affect the generated brain network, but in our previous studies (Zhan et al., 2015b), we presented a very detailed paper that was not able to detect any significant difference in classification accuracy, using brain networks generated from different tractography methods. This was extremely surprising to us, as some tractography methods lead to a much sparser representation of brain connectivity than others. But it seemed like they were all somewhat sensitive to disease effects and their accuracy was hard to distinguish even in a DTI sample of a reasonable size. In the meantime, we also conducted similar studies using different network derived from different tractography algorithms, the accuracy was also boosted by HO-SVD in compared to SVD or raw. Because of these, we only presented the result from the most common tractography algorithm, FACT, and focused our analysis on the features from the networks and classification algorithms best suited for distinguishing between the various stages of neurodegeneration. Taken together, it seems like using HO-SVD makes more difference than the tract tracing method, at least among the ones we analyzed, which were all quite well validated and widely used. Of course the possibility

TABLE 2 | Classification Performance for our three feature extraction methods.

		Raw	SVD	HO-SVD
AD vs. NC	Accuracy	0.7104 ± 0.0816	0.7000 ± 0.1008	0.7125 ± 0.1020
	Sensitivity	0.6750 ± 0.1910	0.6750 ± 0.1985	0.7167 ± 0.1881
	Specificity	0.7222 ± 0.0845	0.7083 ± 0.1080	0.7111 ± 0.1134
	AUC	0.7611 ± 0.1162	0.7694 ± 0.1140	0.7806 ± 0.1117
AD vs. MCI	Accuracy	0.6288 ± 0.0454	0.6259 ± 0.0456	0.6894 ± 0.0612
	Sensitivity	0.5750 ± 0.1750	0.5417 ± 0.1418	0.6083 ± 0.1816
	Specificity	0.6329 ± 0.0484	0.6323 ± 0.0476	0.6956 ± 0.0691
	AUC	0.6195 ± 0.0997	0.6359 ± 0.0807	0.6520 ± 0.1086
MCI vs. NC	Accuracy	0.5311 ± 0.0529	0.5383 ± 0.0610	0.5734 ± 0.0533
	Sensitivity	0.5304 ± 0.0649	0.5384 ± 0.0707	0.5754 ± 0.0568
	Specificity	0.5375 ± 0.1677	0.5375 ± 0.1411	0.5563 ± 0.1489
	AUC	0.5585 ± 0.0892	0.5653 ± 0.0884	0.6104 ± 0.0777

TABLE 3 | Student's *t* test *P*-values are shown for comparing the SVD and the HO-SVD feature sets to the Raw feature set for each of the diagnostic classification tasks.

		SVD > Raw	HO-SVD > Raw	HO-SVD > SVD
AD vs. NC	Accuracy	0.6393	0.4718	0.3494
	Sensitivity	0.5000	0.2456	0.2499
	Specificity	0.6734	0.6363	0.4686
	AUC	0.4101	0.2963	0.3786
AD vs. MCI	Accuracy	0.5805	0.0005	0.0003
	Sensitivity	0.7440	0.2790	0.1017
	Specificity	0.5165	0.0010	0.0009
	AUC	0.2860	0.0267	0.0554
MCI vs. NC	Accuracy	0.3473	0.0082	0.0102
	Sensitivity	0.3562	0.0125	0.0382
	Specificity	0.5000	0.3553	0.3425
	AUC	0.4051	0.0286	0.0474

A Bonferroni correction was adopted here, to account for multiple testing. As there are four measures including Accuracy, Sensitivity, Specificity, and AUC, the corrected *P* threshold in each column is $0.05/4 = 0.0125$. *P* < 0.0125 are marked in red.

remains that someone will develop a better algorithm in the future.

Conclusion

In this study, we proposed a novel framework to differentiate different stages of cognitive impairment—from no impairment in healthy controls to mild cognitive impairment and ultimately Alzheimer's disease, using diffusion MRI derived structural networks in conjunction with a sparse machine learning method. Experimental results indicate that our proposed framework performed better than more traditional methods (direct comparisons of matrix elements or singular value decomposition; SVD) in our network classification tests. Future studies will extend this framework to multi-task classification to better detect earlier stages of Alzheimer's disease, as well as including data from other modalities (anatomical MRI, PIB-PET) that may further improve classification.

Acknowledgments

Algorithm development and image analysis for this study was funded, in part, by grants to PT from the NIBIB (R01 EB008281, R01 EB008432), by the NIA, NIBIB, NIMH, the National Library of Medicine, and the National Center for Research Resources (AG016570, AG040060, EB01651, MH097268, LM05639, RR019771 to PT. PT is also supported, in part, by U54 EB020403 (the “Big Data to Knowledge” Initiative), supported by a cross-NIH Consortium. This work is also supported by R21 AG043760 to YW, R01 LM010730 to JY, and by the National Science Foundation (IIS-0812551, IIS-0953662 to JY). Data collection and sharing for this project was funded

by ADNI (NIH Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amorfis Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California. This research was also supported by NIH grants P30 AG010129 and K01 AG030514 from the National Institute of General Medical Sciences.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnins.2015.00257>

References

- Aganj, I., Lenglet, C., Jahanshad, N., Yacoub, E., Harel, N., Thompson, P. M., et al. (2011). A Hough transform global probabilistic approach to multiple-subject diffusion MRI tractography. *Med. Image Anal.* 15, 414–425. doi: 10.1016/j.media.2011.01.003
- Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., and Aldroubi, A. (2000). *In vivo* fiber tractography using DT-MRI data. *Magn. Reson. Med.* 44, 625–632. doi: 10.1002/1522-2594(200010)44:4<625::AID-MRM17>3.0.CO;2-O
- Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F., and Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *Neuroimage* 34, 144–155. doi: 10.1016/j.neuroimage.2006.09.018
- Bouwman, F. H., van der Flier, W. M., Schoonenboom, N. S., van Elk, E. J., Kok, A., Rijmen, F., et al. (2007). Longitudinal changes of CSF biomarkers in memory clinic patients. *Neurology* 69, 1006–1011. doi: 10.1212/01.wnl.0000271375.37131.04
- Burns, A. (2009). Alzheimer's disease: on the verges of treatment and prevention. *Lancet Neurol.* 8, 4–5. doi: 10.1016/S1474-4422(08)70271-0
- Burns, A., and Iliffe, S. (2009). Alzheimer's disease. *BMJ* 338:b158. doi: 10.1136/bmj.b158
- Conturo, T. E., Lori, N. F., Cull, T. S., Akbudak, E., Snyder, A. Z., Shimony, J. S., et al. (1999). Tracking neuronal fiber pathways in the living human brain. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10422–10427. doi: 10.1073/pnas.96.18.10422
- Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehericy, S., Habert, M. O., et al. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. *Neuroimage* 56, 766–781. doi: 10.1016/j.neuroimage.2010.06.013
- de Jong, L. W., van der Hiele, K., Veer, I. M., Houwing, J. J., Westendorp, R. G., Bollen, E. L., et al. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimer's disease: an MRI study. *Brain* 131, 3277–3285. doi: 10.1093/brain/awn278
- De Lathauwer, L., De Moor, B., and Vandewalle, J. (2000). A Multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.* 21, 1253–1278. doi: 10.1137/S0895479896305696
- Demirhan, A., Nir, T., Zavaliangos-Petropulu, A., Jack, C., Weiner, M., Bernstein, M., et al. (2015). “Feature selection improves the accuracy of classifying alzheimer disease using diffusion tensor images,” in *ISBI2015* (Brooklyn, NY).
- Dennis, E., Jin, Y., Kernan, C., Babikian, T., Mink, R., Babikian, T., et al. (2015a). “White matter integrity in traumatic brain injury: effects of permissible fiber turning angle,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: from Nano to Macro IEEE International Symposium on Biomedical Imaging* (New York, NY), 930–933.
- Dennis, E. L., Jin, Y., Villalon-Reina, J. E., Zhan, L., Kernan, C. L., Babikian, T., et al. (2015b). White matter disruption in moderate/severe pediatric traumatic brain injury: advanced tract-based analyses. *Neuroimage Clin.* 7, 493–505. doi: 10.1016/j.nicl.2015.02.002
- Dennis, E., Zhan, L., Jahanshad, N., Mueller, B., Jin, Y., Lenglet, C., et al. (2014). “Rich club analysis of structural brain connectivity at 7 tesla versus 3 tesla,” in *Computational Diffusion MRI and Brain Connectivity*, eds T.

- Schultz, G. Nedjati-Gilani, A. Venkataraman, L. O'Donnell, and E. Panagiotaki (Nagoya: Springer International Publishing), 209–218.
- De Santi, S., de Leon, M. J., Rusinek, H., Convit, A., Tarshish, C. Y., Roche, A., et al. (2001). Hippocampal formation glucose metabolism and volume losses in MCI and AD. *Neurobiol. Aging* 22, 529–539. doi: 10.1016/S0197-4580(01)00230-5
- Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021
- Fan, Y., Batmanghelich, N., Clark, C. M., Davatzikos, C., and Alzheimer's Disease Neuroimaging, I. (2008). Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *Neuroimage* 39, 1731–1743. doi: 10.1016/j.neuroimage.2007.10.031
- Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., et al. (2010). CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. *J. Neurosci.* 30, 2088–2101. doi: 10.1523/JNEUROSCI.3785-09.2010
- Gerardin, E., Chetelat, G., Chupin, M., Cuingnet, R., Desgranges, B., Kim, H. S., et al. (2009). Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *Neuroimage* 47, 1476–1486. doi: 10.1016/j.neuroimage.2009.05.036
- Grundman, M., Petersen, R. C., Ferris, S. H., Thomas, R. G., Aisen, P. S., Bennett, D. A., et al. (2004). Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Arch. Neurol.* 61, 59–66. doi: 10.1001/archneur.61.1.59
- Gutman, B. A., Wang, Y., Yanovsky, I., Hua, X., Toga, A. W., Jack, C. R. Jr., et al. (2015). Empowering imaging biomarkers of Alzheimer's disease. *Neurobiol. Aging* 36(Suppl. 1), S69–S80. doi: 10.1016/j.neurobiolaging.2014.05.038
- Hua, X., Hibar, D. P., Ching, C. R., Boyle, C. P., Rajagopalan, P., Gutman, B. A., et al. (2013). Unbiased tensor-based morphometry: improved robustness and sample size estimates for Alzheimer's disease clinical trials. *Neuroimage* 66, 648–661. doi: 10.1016/j.neuroimage.2012.10.086
- Hua, X., Lee, S., Hibar, D. P., Yanovsky, I., Leow, A. D., Toga, A. W., et al. (2010). Mapping Alzheimer's disease progression in 1309 MRI scans: power estimates for different inter-scan intervals. *Neuroimage* 51, 63–75. doi: 10.1016/j.neuroimage.2010.01.104
- Hua, X., Leow, A. D., Lee, S., Klunder, A. D., Toga, A. W., Lepore, N., et al. (2008a). 3D characterization of brain atrophy in Alzheimer's disease and mild cognitive impairment using tensor-based morphometry. *Neuroimage* 41, 19–34. doi: 10.1016/j.neuroimage.2008.02.010
- Hua, X., Leow, A. D., Parikshak, N., Lee, S., Chiang, M. C., Toga, A. W., et al. (2008b). Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects. *Neuroimage* 43, 458–469. doi: 10.1016/j.neuroimage.2008.07.013
- Jahanshad, N., Aganj, I., Lenglet, C., Joshi, A., Jin, Y., Barysheva, M., et al. (2011). “Sex differences in the human connectome: 4-Tesla high angular resolution diffusion imaging (HARDI) tractography in 234 young adult twins,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: From Nano to Macro IEEE International Symposium on Biomedical Imaging* (Chicago, IL), 939–943.
- Jahanshad, N., Zhan, L., Bernstein, M. A., Borowski, B. J., Jack, C. R., Toga, A. W., et al. (2010). “Diffusion tensor imaging in seven minutes: determining trade-offs between spatial and directional resolution,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: From Nano to Macro IEEE International Symposium on Biomedical Imaging* (Rotterdam), 1161–1164.
- Jin, Y., Shi, Y., Joshi, S., Jahanshad, N., Zhan, L., de Zubicaray, G., et al. (2011). Heritability of white matter fiber tract shapes: a HARDI study of 198 twins. *Multimodal Brain Image Analysis* 2011, 35–43. doi: 10.1007/978-3-642-24446-9_5
- Jin, Y., Shi, Y., Zhan, L., de Zubicaray, G. I., McMahon, K. L., Martin, N. G., et al. (2013). “Labeling white matter tracts in hardi by fusing multiple tract atlases with applications to genetics,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: From Nano to Macro IEEE International Symposium on Biomedical Imaging* 2013 (San Francisco, CA), 512–515.
- Jin, Y., Shi, Y., Zhan, L., Gutman, B. A., de Zubicaray, G. I., McMahon, K. L., et al. (2014). Automatic clustering of white matter fibers in brain diffusion MRI with an application to genetics. *Neuroimage* 100, 75–90. doi: 10.1016/j.neuroimage.2014.04.048
- Jin, Y., Shi, Y., Zhan, L., Li, J., de Zubicaray, G., McMahon, K., et al. (2012). Automatic population HARDI white matter tract clustering by label fusion of multiple tract atlases. *Multimodal Brain Image Analysis* 7509, 147–156. doi: 10.1007/978-3-642-33530-3_12
- Jin, Y., Shi, Y., Zhan, L., and Thompson, P. (2015). “Automated multi-atlas labeling of the fornix and its integrity in Alzheimer's disease,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: from nano to macro IEEE International Symposium on Biomedical Imaging* (New York, NY), 140–143.
- Lazar, M., Weinstein, D. M., Tsuruda, J. S., Hasan, K. M., Arfanakis, K., Meyerand, M. E., et al. (2003). White matter tractography using diffusion tensor deflection. *Hum. Brain Mapp.* 18, 306–321. doi: 10.1002/hbm.10102
- Lewis, J. D., Evans, A. C., Pruett, J. R., Botteron, K., Zwaigenbaum, L., Estes, A., et al. (2014). Network inefficiencies in autism spectrum disorder at 24 months. *Transl. Psychiatry* 4, e388. doi: 10.1038/tp.2014.24
- Li, J., Jin, Y., Shi, Y., Dinov, I. D., Wang, D. J., Toga, A. W., et al. (2013). Voxelwise spectral diffusional connectivity and its applications to Alzheimer's disease and intelligence prediction. *Med. Image Comput. Comput. Assist. Interv.* 16, 655–662. doi: 10.1007/978-3-642-40811-3_82
- Liu, J., Ji, S., Ye, J. (2009). *SLEP: Sparse Learning with Efficient Projections*. Arizona State University. Available online at: <http://www.public.asu.edu/~jye02/Software/SLEP>
- Magnin, B., Mesrob, L., Kinkingnehun, S., Pelegrini-Issac, M., Colliot, O., Sarazin, M., et al. (2009). Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. *Neuroradiology* 51, 73–83. doi: 10.1007/s00234-008-0463-x
- Mattsson, N., Zetterberg, H., Hansson, O., Andreasen, N., Parnetti, L., Jonsson, M., et al. (2009). CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. *JAMA* 302, 385–393. doi: 10.1001/jama.2009.1064
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12, 153–157. doi: 10.1007/BF02295996
- Mocks, J., and Verleger, R. (1986). Principal component analysis of event-related potentials: a note on misallocation of variance. *Electroencephalogr. Clin. Neurophysiol.* 65, 393–398. doi: 10.1016/0168-5597(86)90018-3
- Mori, S., Crain, B. J., Chacko, V. P., and van Zijl, P. C. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann. Neurol.* 45, 265–269.
- Morris, J. C., Storandt, M., Miller, J. P., McKeel, D. W., Price, J. L., Rubin, E. H., et al. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. *Arch. Neurol.* 58, 397–405. doi: 10.1001/archneur.58.3.397
- Nir, T. M., Jahanshad, N., Villalon-Reina, J. E., Toga, A. W., Jack, C. R., Weiner, M. W., et al. (2013). Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging. *Neuroimage Clin.* 3, 180–195. doi: 10.1016/j.nicl.2013.07.006
- Parker, G. J., Haroon, H. A., and Wheeler-Kingshott, C. A. (2003). A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements. *J. Magn. Reson. Imaging* 18, 242–254. doi: 10.1002/jmri.10350
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* 56, 303–308. doi: 10.1001/archneur.56.3.303
- Prasad, G., Mackey, L., Joshi, S., Mezher, A., Galvis, J., Ragothaman, A., et al. (2015). “Evolving the human connectome for efficient classification of disease,” in *Submit to MICCAI2015* (Munich).
- Querfurth, H. W., and LaFerla, F. M. (2010). Alzheimer's disease. *N. Engl. J. Med.* 362, 329–344. doi: 10.1056/NEJMra0909142
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Shaw, L. M., Vanderstichele, H., Knapić-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., et al. (2009). Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann. Neurol.* 65, 403–413. doi: 10.1002/ana.21610
- Simic, G., Stanic, G., Mladinov, M., Jovanov-Milosevic, N., Kostovic, I., and Hof, P. R. (2009). Does Alzheimer's disease begin in the brainstem? *Neuropathol. Appl. Neurobiol.* 35, 532–554. doi: 10.1111/j.1365-2990.2009.01038.x

- Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31, 1487–1505. doi: 10.1016/j.neuroimage.2006.02.024
- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. B* 58, 267–288.
- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc.* 73, 273–282. doi: 10.1111/j.1467-9868.2011.00771.x
- Todd, S., Barr, S., Roberts, M., and Passmore, A. P. (2013). Survival in dementia and predictors of mortality: a review. *Int. J. Geriatr. Psychiatry* 28, 1109–1124. doi: 10.1002/gps.3946
- Toga, A. W., Clark, K. A., Thompson, P. M., Shattuck, D. W., and Van Horn, J. D. (2012). Mapping the human connectome. *Neurosurgery* 71, 1–5. doi: 10.1227/NEU.0b013e318258e9ff
- Westman, E., Simmons, A., Zhang, Y., Muehlboeck, J. S., Tunnard, C., Liu, Y., et al. (2011). Multivariate analysis of MRI data for Alzheimer's disease, mild cognitive impairment and healthy controls. *Neuroimage* 54, 1178–1187. doi: 10.1016/j.neuroimage.2010.08.044
- Zalesky, A., Fornito, A., Seal, M. L., Cocchi, L., Westin, C. F., Bullmore, E. T., et al. (2011). Disrupted axonal fiber connectivity in schizophrenia. *Biol. Psychiatry* 69, 80–89. doi: 10.1016/j.biopsych.2010.08.022
- Zhan, L., Franc, D., Patel, V., Jahanshad, N., Jin, Y., Mueller, B. A., et al. (2012). How do spatial and angular resolution affect brain connectivity maps from diffusion mri? *Proc. IEEE Int. Symp. Biomed. Imaging* 1–6. doi: 10.1109/isbi.2012.6235469
- Zhan, L., Jahanshad, N., Ennis, D. B., Jin, Y., Bernstein, M. A., Borowski, B. J., et al. (2013a). Angular versus spatial resolution trade-offs for diffusion imaging under time constraints. *Hum. Brain Mapp.* 34, 2688–2706. doi: 10.1002/hbm.22094
- Zhan, L., Jahanshad, N., Faskowitz, J., Zhu, D., Prasad, G., Marin, N. G., et al. (2015a). “Heritability of Brain network in 853 twins and sibilings,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: From Nano to Macro IEEE International Symposium on Biomedical Imaging*, (Brooklyn, NY), 449–453.
- Zhan, L., Jahanshad, N., Jin, Y., Nir, T. M., Leonardo, C., Bernstein, M. A., et al. (2014a). “Understanding scanner upgrade effects on brain integrity & connectivity measures,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: From Nano to Macro IEEE International Symposium on Biomedical Imaging* (Beijing), 234–237.
- Zhan, L., Jahanshad, N., Jin, Y., Toga, A. W., McMahon, K. L., de Zubicaray, G. I., et al. (2013b). “Brain network efficiency and topology depend on the fiber tracking method: 11 tractography algorithms compared in 536 subjects,” in *Proceedings/IEEE International Symposium on Biomedical Imaging: From Nano to Macro IEEE International Symposium on Biomedical Imaging* (San Francisco, CA), 1134–1137.
- Zhan, L., Mueller, B. A., Jahanshad, N., Jin, Y., Lenglet, C., Yacoub, E., et al. (2013c). Magnetic resonance field strength effects on diffusion measures and brain connectivity networks. *Brain Connect.* 3, 72–86. doi: 10.1089/brain.2012.0114
- Zhan, L., Nie, Z., Ye, J., Wang, Y., Jin, Y., Jahanshad, N., et al. (2014b). “Multiple stages classification of Alzheimer's disease based on structural brain networks using generalized low rank approximations (GLRAM),” in *Computational Diffusion MRI*, ed L. O'Donnell, G. Nedjati-Gilani, Y. Rathi, M. Reiser, and T. Schneider (Boston, MA: Springer International Publishing), 35–44.
- Zhan, L., Zhou, J., Wang, Y., Jin, Y., Jahanshad, N., Prasad, G., et al. (2015b). Comparison of 9 tractography algorithms for detecting abnormal structural brain networks in Alzheimer's disease. *Front. Aging Neurosci.* 7:48. doi: 10.3389/fnagi.2015.00048

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Zhan, Liu, Wang, Zhou, Jahanshad, Ye and Thompson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Multimodal Imaging Signatures of Parkinson's Disease

F. DuBois Bowman^{1*}, Daniel F. Drake¹ and Daniel E. Huddleston²

¹ Department of Biostatistics, The Mailman School of Public Health, Columbia University, New York, NY, USA, ² Department of Neurology, Emory University, Atlanta, GA, USA

OPEN ACCESS

Edited by:

Han Liu,
Princeton University, USA

Reviewed by:

Andrew L. Alexander,
University of Wisconsin School of
Medicine and Public Health, USA
Baxter P. Rogers,
Vanderbilt University, USA

*Correspondence:

F. DuBois Bowman
dubois.bowman@columbia.edu

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 30 October 2015

Accepted: 15 March 2016

Published: 18 April 2016

Citation:

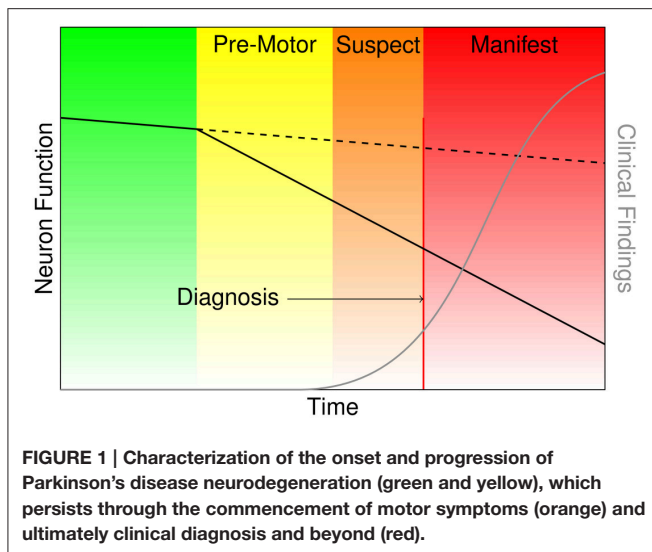
Bowman FD, Drake DF and
Huddleston DE (2016) Multimodal
Imaging Signatures of Parkinson's
Disease. *Front. Neurosci.* 10:131.
doi: 10.3389/fnins.2016.00131

Parkinson's disease (PD) is a complex neurodegenerative disorder that manifests through hallmark motor symptoms, often accompanied by a range of non-motor symptoms. There is a putative delay between the onset of the neurodegenerative process, marked by the death of dopamine-producing cells, and the onset of motor symptoms, creating an urgent need to develop biomarkers that may yield early PD detection. Neuroimaging offers a non-invasive approach to examining the potential utility of a vast number of functional and structural brain characteristics as biomarkers. We present a statistical framework for analyzing neuroimaging data from multiple modalities to determine features that reliably distinguish PD patients from healthy control (HC) subjects. Our approach builds on elastic net, performing regularization and variable selection, while introducing additional criteria centering on parsimony and reproducibility. We apply our method to data from 42 subjects (28 PD patients and 14 HC). Our approach demonstrates extremely high accuracy, assessed via cross-validation, and isolates brain regions that are implicated in the neurodegenerative PD process.

Keywords: multimodal imaging, MRI, prediction, classification, penalized regression, Parkinson's disease, biomarker

INTRODUCTION

Parkinson's disease (PD) is a devastating, progressive movement disorder affecting 7–10 million individuals worldwide (Parkinson's Disease Foundation, 2015). PD usually affects people over 50 years of age, but a subset of patients experience early onset. The hallmark pathology of PD is the loss of dopaminergic neurons in the substantia nigra pars compacta (SNpc), but the disease manifests with a diversity of symptoms referable to multi-system neuropathology. The clinical features of PD include the classic motor symptoms of tremor, rigidity, bradykinesia, and gait impairment, as well as a host of non-motor symptoms (Kalia and Lang, 2015). At the time of PD diagnosis it has been estimated based on histopathology that over 50% of dopamine neurons in the SNpc have died (Fearnley and Lees, 1991). Braak et al. (2003) posit a process of phased pathology of PD, which suggests that early neurodegeneration occurs in lower brainstem structures and progresses in ascending fashion, in particular affecting the locus coeruleus in Stage II and SNpc in Stage III. Further progression extends to higher-level sensory association areas and prefrontal cortical regions, eventually impacting first order sensory association areas, premotor regions, and primary sensory and motor fields (Del Tredici and Braak, 2013). The putative delay in the onset of motor symptoms leading to PD diagnosis is portrayed in **Figure 1**, and the corresponding neurodegeneration occurring throughout this pre-motor period represents a missed opportunity for early therapeutic intervention that may significantly slow or halt the progression of PD related decline.



There is so far no reliable method to accurately diagnose PD in its pre-motor stages, and addressing this unmet need is a key challenge in the field of PD biomarker development. Many pre-motor symptoms of PD are non-specific, including depression, anxiety, constipation, and excessive daytime sleepiness (Tolosa and Pont-Sunyer, 2011). REM sleep behavior disorder (RBD) in the absence of dementia, hallucinations, autonomic dysfunction or parkinsonian motor symptoms, referred to as idiopathic RBD (iRBD), portends a high likelihood of eventual conversion to a synucleinopathy: PD, multiple system atrophy or Lewy body dementia (Iranzo et al., 2013). However, simply identifying iRBD does not allow prediction of the specific clinical phenotype a patient will develop, and the duration to phenoconversion is variable from the time of iRBD diagnosis (Iranzo et al., 2013; Postuma et al., 2015), which makes pre-motor PD study design in this group more challenging. Furthermore, clinical presentation with iRBD prior to evidence of a broader neurodegenerative syndrome is also relatively uncommon and most PD patients have not sought treatment for iRBD prior to phenoconversion with motor symptoms. Other strategies for pre-motor diagnosis of PD have included combining clinical features, such as olfactory loss and family history, with dopamine transporter radionuclide imaging (The Parkinson At-Risk Study or PARS), or an algorithmic approach to develop a cohort enriched with an at-risk genotype, such as LRRK2 G2019S mutation (Tolosa and Pont-Sunyer, 2011; Foroud et al., 2015). While it appears likely that a multi-tiered screening process to identify pre-motor or asymptomatic at risk subjects has promise, inclusion of neuroimaging in a cost-effective manner for in vivo confirmation of PD associated brain pathology may speed up and improve the efficiency of these studies. MRI is a fraction of the cost of radionuclide imaging (Fiandaca et al., 2014), and allows efficient collection of multiple types of disease-relevant brain measurements, including assessment of structural and functional connectivity, which are expected to be impacted by the degeneration of the widely projecting catecholamine neurons affected in PD. Here we leverage advanced statistical methods

to identify robust candidate biomarkers and profiles from a large number of MRI features to differentiate patients with early to moderate PD from controls. Because the neurodegeneration process is already advanced at the time of PD diagnosis, a highly robust biomarker in early to moderate (motor) PD patients is likely to be detectable in the pre-motor state as well. Therefore, the outputs of this study may serve as candidate neuroimaging biomarkers in future studies of pre-motor or asymptomatic PD.

Our research is driven by a broad initiative called the Parkinson's Disease Biomarker Program (PDBP) at the National Institutes of Health's (NIH's) National Institute of Neurological Disorders and Stroke to identify early stage biomarkers for PD. In the context of our study, we regard biomarkers in a general sense, defined by an NIH Biomarkers Definitions Working Group as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention," although more specific molecular definitions have been proposed (Strimbu and Tavel, 2010). A first step in identifying early stage PD neuroimaging biomarkers is to determine neural characteristics that reliably distinguish patients with mild to moderate PD from healthy control subjects.

Neuroimaging has shown early promise for identifying alterations associated with PD. There is emerging evidence of cortical thinning in PD patients determined from T1 MRI (Lee et al., 2013; Zarei et al., 2013; Zhang et al., 2015). Vaillancourt et al. (2009) established neuroimaging correlates of PD through decreases in fractional anisotropy generated from diffusion tensor imaging (DTI) data within caudal regions of the substantia nigra. Du et al. (2011) found that augmenting fractional anisotropy measures of the substantia nigra with its transverse relaxation rate, $R2^*$, improved the discrimination of PD patients from controls over that of using fractional anisotropy alone. Kahan et al. (2014) target effective connectivity in PD using resting state fMRI (rs-fMRI) in patients with deep brain stimulation, suggesting that subthalamic nucleus modulates major components of the motor cortico-striato-thalamo-cortical loop.

Single modality neuroimaging renders only a partial view toward understanding the neural basis for PD. When targeting classification or prediction, simultaneously examining data from multiple imaging modalities stands to increase accuracy, to provide a more complete picture of the multiple neuropathophysiologic manifestations of PD, and to determine the relative predictive strengths of the PD-related functional and structural changes.

We conduct a novel multimodal imaging investigation that seeks to identify functional and structural changes in mild to moderate PD, which collectively yield high prediction accuracy in dissociating patients from healthy control subjects. Of note, our goals extend beyond simply achieving high prediction accuracy. We aim to contribute to PD biomarker discovery efforts by determining the potential involvement of specific brain regions in the disease process, whether novel or previously studied, which will help to direct future research. Therefore, we balance our objective of high accuracy with criteria of parsimony and reproducibility.

We utilize elastic net, an advanced statistical learning technique, building in novel refinements to enhance performance and to achieve desired levels of parsimony and reproducibility. Elastic net blends both L_1 and L_2 penalties, applied here in context of logistic regression, to perform both regularization and variable selection (Zou and Hastie, 2005). We apply the analysis techniques to a set of measures based on magnetic resonance imaging (MRI), including structural T1 images, rs-fMRI, and DTI from PD patients and healthy control subjects. We perform cross-validation to assess accuracy. Overall, the approach achieves extremely high accuracy and reveals key neuroimaging contributors that help to reliably distinguish PD patients from healthy controls.

EXPERIMENTAL DATA AND METHODS

Experimental Data

All subject records and data, collected under the auspices of a previous study, were supplied de-identified, stripped of any protected health information (PHI) and personally identifiable information (PII). Accordingly, this research qualifies as Research of Existing Data, Records, Specimens [Basic Exempt Criteria 45 CFR 46.101(b)(4)], and has been deemed “Not Human Subjects Research” (HS Code 10 in IPMAC II as referenced in the manual chapter 7410) by NIH and Columbia University Medical Center Institutional Review Board (Protocol: IRB-AAO0062).

We consider data from 42 subjects, including 28 PD patients and 14 healthy control (HC) subjects. The data include a collection of magnetic resonance (MR) derived scans characterizing different structural and functional properties of the brain as well as demographic measures. Specifically, we use T1-weighted anatomical MRI scans, rs-fMRI, and DTI. The mean age of the subjects is 65.0 years (9.0 years standard deviation), and the subjects include 21 males and 21 females. The mean age is 61.9 years (8.7 years standard deviation) for PD patients and 71.4 years (5.8 years standard deviation) for controls (a significant difference, with $p < 0.001$). The PD group has 13 females (46.4% of PD patients) and the control group has 8 females (57.1% of controls), reflecting a small sex difference between groups, although not statistically significant ($p = 0.74$). The mean Unified Parkinson's Disease Rating Scale (UPDRS) Part III (motor) score for these patients was 19.4 (standard deviation 10.2). The mean duration of disease was 7.7 years (standard deviation 3.3 years), although the duration was not calculable for 5 patients due to missing data.

All scans were captured with a Siemens Trio Tim 3T MRI scanner; the first 36 subjects were scanned with a 12 channel head coil and the remaining 6 subjects (5 PD and 1 control) were scanned with a 32 channel head coil (we control for this difference in the statistical analyses). The structural T1 scans were acquired using MPRAGE (TR = 2600 ms, TE = 3 ms, 192 sagittal slices at 1 mm; 256×232 1 mm isotropic pixels). Echo planar imaging (EPI) was used to acquire 140 frames of rs-fMRI scans (TR = 3000 ms, TE = 30 ms, 48 axial slices at 3 mm, 128×128 2 mm isotropic pixels) for each subject. DTI data were captured

using a biphasic approach with consecutive left-to-right and right-to-left phase scans. The first thirty six subjects underwent DTI scans (TR = 8700ms, TE = 94ms, 64 axial slices at 2 mm, 128×128 2 mm isotropic pixels) comprised of 64 directions ($B = 1000\text{s/mm}^2$), with three leading and three trailing B0 scans. The remaining 6 subjects followed a DTI protocol (TR = 3292 ms, TE = 97.6 ms, 66 axial slices at 2 mm, 92×106 2 mm pixels) comprised of 128 directions ($B = 1000\text{s/mm}^2$), with six leading and five trailing B0 scans.

We implemented standard neuroimaging preprocessing steps including voxel-based morphometry (VBM) on the anatomical T1 scan, using the VBM toolbox (Gaser, 2010) under SPM8, produced voxel-wise estimates of gray matter density in MNI space, along with subject-specific native-to-MNI DARTEL transformations (and their inverses) and gray matter, white matter, and cerebral spinal fluid segmentations. The inverse transformations were used to map MNI-defined parcellations back to each subject's native space. Resting state preprocessing, performed with AFNI (Cox, 1996), consisted of a despiking stage, slice time correction, motion correction, spatial normalization to MNI and smoothing by 6mm FWHM. The resulting rs-fMRI time courses were orthogonalized relative to Legendre polynomials orders 0 through 3; motion parameters and their derivatives; and global white matter and ventricular cerebral spinal fluid (CSF) signals. Finally, the time courses were filtered to the band 0.01–0.1 Hz.

A t -test applied to the resting state scans shows no difference in mean temporal SNR between PD (54.4 ± 2.9) and control (53.9 ± 4.9) subjects ($p = 0.92$), with standard error of the mean used to express variability. Similarly, a Wilcoxon rank sum test shows no significant difference in the maximum absolute displacement over the duration of the scan between PD ($1.67 \text{ mm} \pm 0.21 \text{ mm}$) and control ($1.31 \text{ mm} \pm 0.17 \text{ mm}$) subjects ($p = 0.53$). Finally, another motion-related quantity, the average motion per TR, also does not differ significantly between PD ($0.10 \text{ mm} \pm 0.01 \text{ mm}$) and control ($0.08 \text{ mm} \pm 0.01 \text{ mm}$) subjects ($p = 0.38$).

For DTI scans, each subject's two opposing phase DTI scans were combined to estimate the susceptibility-induced off-resonance field using a method similar to that described in Andersson et al. (2003) as implemented in FSL (Smith et al., 2004) and the two images were combined into a single corrected one. The resulting composite scan was corrected for eddy currents. After preprocessing, we have $121 \times 145 \times 121$ voxels (1.5 mm isotropic) for DTI and VBM and $91 \times 109 \times 91$ (2 mm isotropic) for rs-fMRI.

Methods

Modality-Specific Data Representations

The first step is to determine the spatial scale for data representations. The imaging data from MRI, rs-fMRI, and DTI are acquired at a voxel level. We utilize a popular neuroanatomic parcellation of the brain, the Automated Anatomical Labeling (AAL) (Tzourio-Mazoyer et al., 2002) system, to define 90 brain regions. For MRI and rs-fMRI, we further refine the standard AAL parcellation by defining subregions to yield more homogeneous collections of voxels within subregions. This refinement of the AAL-90 parcellation uses a hierarchical

clustering algorithm to subdivide each region based on a metric that combines distance, structural and functional connectivity, and tissue type to identify homogeneous subregions of the encompassing region. The resulting extended parcellation produces 290 subregions (AAL-290), with a given subregion falling entirely within a single AAL region. The regional parcellations appear in **Figure 2**.

We generate data representations (or features) for each imaging modality and specify the spatial scale. **Figure 3** provides a conceptual overview describing the multiple modalities generating data, estimates obtained from each reflecting particular structural or functional properties of the brain, the spatial scale for each summary, and ultimately the features constituting the global set of potential neuroimaging markers of PD. We use 290 regions from the extended AAL map (AAL-290) to compute regional averages of local volumetric MRI measures, specifically from voxel-based morphometry (VBM) (see **Table 1**). We use rs-fMRI data to generate both localized and connectivity features. To quantify the power concentrated at low frequencies for fMRI data, we use fractional amplitude of low frequency fluctuation (fALFF), which calculates the ratio of the power spectrum at low-frequencies (0.01–0.10 Hz) to that of the entire frequency range (Zou et al., 2008). We compute fALFF at a voxel level, for all voxels, and average within each of the 290 subregions. We quantify functional connectivity (FC) by calculating pairwise correlations between the average time courses within each pair of the 290 subregions. We compute fractional anisotropy (FA) for each voxel and obtain regional summaries by averaging over each of the AAL-90 regions. Thus our summary measure will increase both as a function of the restricted diffusion in the regional white matter and the proportion of white matter within a region. We calculate structural connectivity (SC) derived from DTI, using anisotropy to constrain tracking. We use FSL to perform estimation of the diffusion tensor (BEDPOSTX) and tractography (PROBTRACKX) (Behrens et al., 2007).

We perform marginal screening to reduce the 46,580 features prior to analysis by eliminating features that are unlikely to carry strong predictive power. Screening typically improves the performance and facilitates implementation of subsequent modeling by eliminating sources of noise and reducing data dimensionality. Toward our goal of attaining reproducible PD biomarkers, we perform a bootstrap screening procedure for each

feature independently using logistic regression, with modality-specific screening thresholds. The bootstrap procedure isolates a set of viable markers, after accounting for sampling variability, which increases the likelihood that the identified features will emerge in other samples. Specifically, our screening rule selects features satisfying the following:

$$p^* = \frac{1}{B} \sum_{b=1}^B I[p_b < p_0] \geq r.$$

In our analysis, we perform independent screening within each of $B = 100$ bootstrap samples indexed by $b = 1, \dots, B$ to obtain a corresponding p -value p_b , apply designated modality-specific thresholds p_0 , and determine features that are selected in at least $r = 0.75$ proportion of the bootstrap samples. After our screening process, we retained 24 regional VBM features ($p_0 = 0.2$), 6 fALFF ($p_0 = 0.2$), 225 FC estimates across the brain ($p_0 = 0.05$), 6 regional FA measures ($p_0 = 0.2$), and 10 SC estimates ($p_0 = 0.2$), giving 271 features in total (**Table 1**).

Statistical Learning and Prediction Methods

We propose an analytic approach that uses imaging data from multiple modalities and demographic information to classify subjects as either PD patients or HCs. We present an approach that builds on elastic net with refinements to encourage parsimony and reproducibility. Let $D_i = 1$, if the i th subject has PD and $D_i = 0$, if subject i is a healthy control, $i = 1, \dots, n$. The predictors and an intercept term are arrayed in a vector $X_i = (1, X_{i1}, \dots, X_{ip})'$, with p denoting the number of predictors following screening. We standardize each of the predictors so that $\sum_{i=1}^n x_{ij} = 0$ and $(1/n) \sum_{i=1}^n x_{ij}^2 = 1$. We let $\pi_i = \Pr(D_i = 1 | X_i)$ represent the probability that subject i has PD, given a set of predictors, and use logistic regression to model $\log[\pi_i/(1 - \pi_i)] = X_i' \beta$. The elastic net procedure applied to logistic regression maximizes the likelihood function

$$\max_{\beta} \left\{ \frac{1}{n} \sum_{i=1}^n [D_i \log(\pi_i) + (1 - D_i) \log(1 - \pi_i)] - \lambda \sum_{j=0}^p \left[\frac{1}{2} (1 - \alpha) \beta_j^2 + \alpha |\beta_j| \right] \right\}.$$

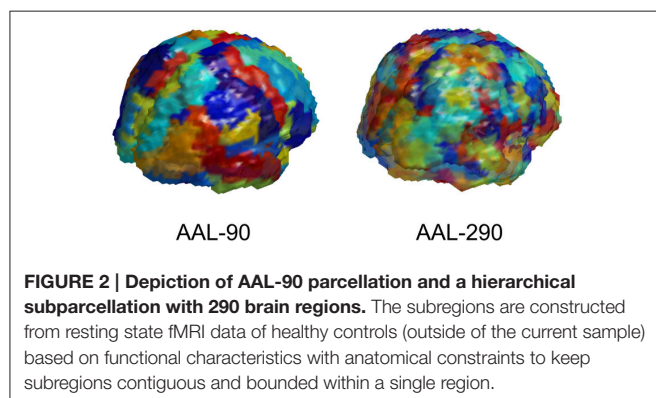
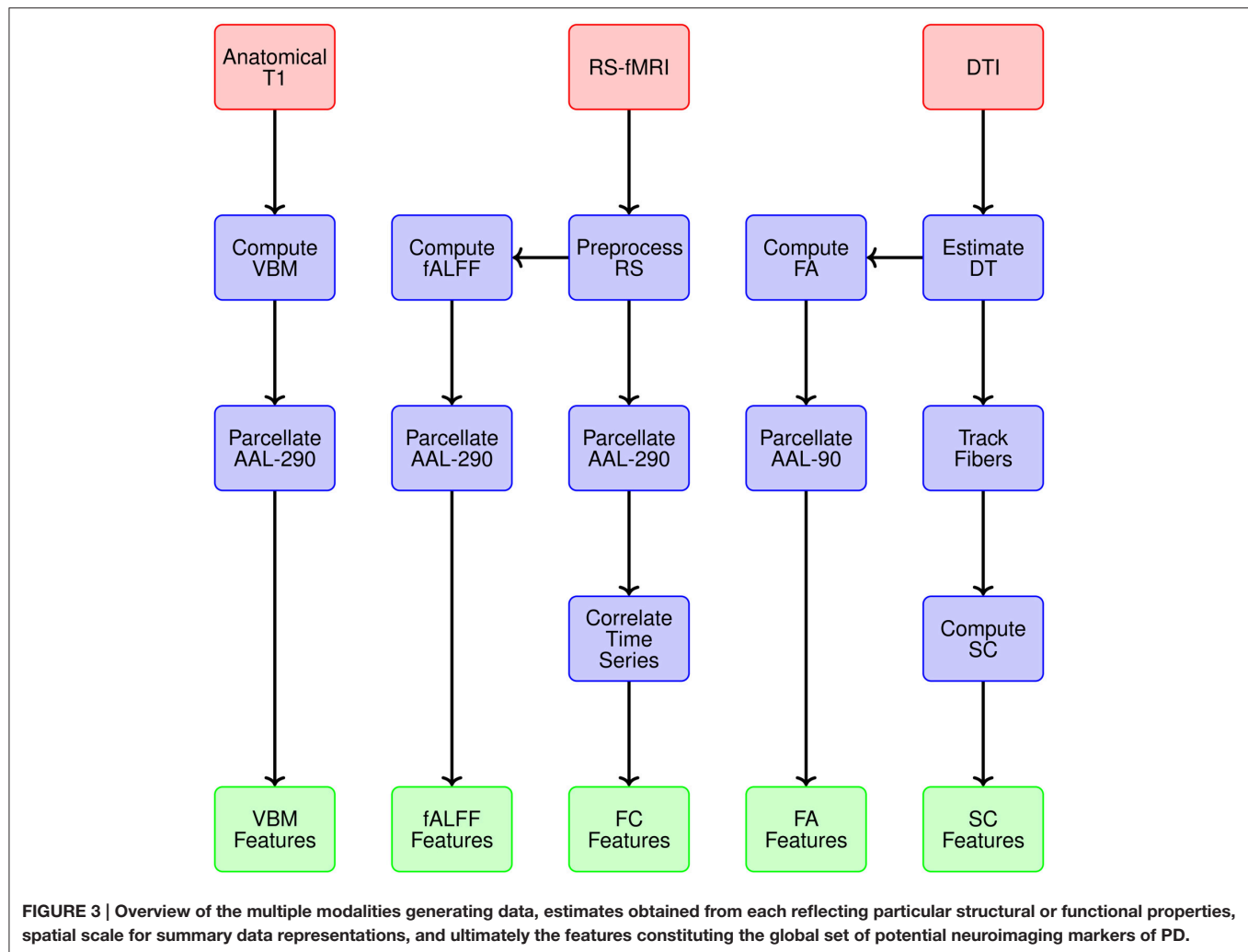


TABLE 1 | Description of modalities, corresponding features, spatial scale, and screening-based dimension reduction.

Modality	Feature	Spatial Scale	Number of Features	Post-screening
MRI	VBM	AAL-290	290 regions	24
rs-fMRI	fALFF	AAL-290	290 subregions	6
rs-fMRI	FC	AAL-290	41,905 subregion pairs	225
DTI	FA	AAL-90	90 regions	6
DTI	SC	AAL-90	4005 region pairs	10
Total			46,580	271



From the large set of variables, the method performs shrinkage and variable selection by blending ridge-regression ($\alpha = 0$) using an L_2 penalty and the lasso ($\alpha = 1$) using an L_1 penalty (Zou and Hastie, 2005). The parameters α and λ are determined by optimizing an objective function via cross-validation, e.g., minimizing the cross-validation error.

The penalized framework, implemented here in context of a logistic model, points to the predictive ability of a specific subset of imaging and demographic variables that constitute a signature for PD in our sample. Ridge regression shrinks the coefficients and tends to draw the coefficients of correlated predictors towards each other. The lasso tends to pull many coefficients near zero, with a small subset of coefficients with larger magnitudes, therefore serving as a useful tool for variable selection. We perform covariate adjustment for demographic variables (age and sex) and scan differences (head coil) in our models. The elastic net penalty is particularly useful when $p \gg n$, and when the set of predictors includes some highly correlated variables, which poses a challenge for L_1 penalization alone.

We modify the usual optimization procedure for the tuning parameters, when necessary, to promote parsimony, accuracy,

and reproducibility (see Results section for details). Our procedure defines a restricted or bounded tuning parameter space, B , in which to optimize (α, λ) . Specifically, we consider

$$B = \{(\alpha, \lambda) \mid p \leq p_1, AUC \geq q_1\},$$

where AUC represents the area under the receiver operating characteristic (ROC) curve. Inducing parsimony may sacrifice accuracy, so the subspace B incorporates a lower bound on AUC as a measure of accuracy.

We evaluate accuracy using an iterated k -fold cross validation scheme for model training and testing to promote reproducibility. A typical implementation of k -fold cross validation splits the data into k groups, trains the model by fitting the data from $k - 1$ groups (training set), and uses the estimates obtained to predict the disease status of each subject in the remaining group (validation set). The process then rotates the training sets and validation sets until testing has been performed on each group, hence each subject. Variability is inherent in k -fold cross validation, which is not typically accounted for in practice. For example,

by constructing the folds differently, one may obtain a different estimate of accuracy and detect the involvement of different predictors. To encourage the identification of neuroimaging markers of PD that are reproducible and have high predictive strength and to account for variability in the cross-validation process, we implement an iterated framework. Specifically, we implement two-fold cross-validation and repeat the process 100 times, randomly assigning subjects to folds in each iteration. This process results in 200 training samples.

The cross-validation approach presents an important advantage in the context of our quest to identify likely PD biomarkers, allowing us to gauge the overall importance of each feature by virtue of its average predictive effect. Since the features were standardized, coefficient strengths are comparable: a larger average coefficient strength indicates a greater predictive effect. At each (α, λ) in B , we aim to select the top 10% of features based on these coefficient strengths. Let M be the random variable representing the magnitude of a predictive effect at (α, λ) . Conceptually, imaging features β_j satisfying $\Pr(M \geq |\beta_j|) \leq \tau_S$, and which contribute to high predictive accuracy, are regarded as strong candidates for potential biomarkers. In practice, we use the cross-validation process to estimate the empirical distribution function of M and determine predictors that have the most sizable effects (on average) across 200 training samples. So for our data, we specifically seek to determine the features satisfying $|\beta_j| \geq \xi_{0.10}$, where $|\beta_j|$ is the average magnitude of the j th effect and $\xi_{0.10}$ is defined by $\Pr(M \geq \xi_{0.10}) = 0.10$.

Moreover, we track the consistency with which these predictors with sizable effects are selected for specific values (α, λ) and ultimately choose features that are consistently strong across various combinations (α, λ) . Let $S(\alpha, \lambda)$ represent the set of features satisfying the above condition for coefficient strength, i.e., $S(\alpha, \lambda) = \{\beta_j | |\beta_j| \geq \xi_{0.10} \text{ and } \Pr(M \geq \xi_{0.10}) = 0.10; (\alpha, \lambda)\}$. Our procedure selects features

$$C = \left\{ \beta_j \mid \left[\frac{1}{\#B} \sum_{(\alpha, \lambda) \in B} I[\beta_j \in S(\alpha, \lambda)] \right] \geq \tau_C \right\},$$

where the notation $\#B$ denotes the cardinality or number of elements in set B , and I is the indicator function. We set $\tau_C = 0.90$, effectively taking the features that were selected to be in set $S(\alpha, \lambda)$ in 90% or more of the points in B . The set of features in C are deemed to have high predictive strength, to be extremely parsimonious, to have high likelihood of emerging in other samples, and to be robust over a range of values in the tuning parameter space. These properties aid the delivery of potential PD biomarkers that can be investigated further in future research. In the application of our methods to the multimodal imaging data of PD patients and healthy controls discussed below, we explore further reductions of the set C .

RESULTS

We applied the methods above to our multimodal imaging data. We consider a 51×151 grid of elastic net tuning parameters,

with $\alpha \in [0, 1]$ and $\lambda \in [10^{-5}, 10^1]$, with 25 points per decade. For every (α, λ) pair, we fit elastic net to half the subjects, then apply the resulting model to the other half of the subjects to predict their disease status. We control for head coil, sex, and age in the model fit. Then we swap sets of subjects and perform the operation again; i.e., two-fold cross validation. We compare the result of the predictions with true disease status to compute the ROC curve and associated AUC value. Finally, we perform this operation 100 times at every (α, λ) in the grid and record the average AUC and various statistics on the model coefficients for each of the 271 features.

The resulting average AUC values in the (α, λ) grid are shown in **Figure 4A**. Point A indicates the (α, λ) combination with the maximum average area under the curve, $AUC = 0.989$. The corresponding average ROC curve (black) is shown in **Figure 4B**, along with the individual ROC curves from each cross-validation fit, indicating the degree of variability across samples. Point A, at $\alpha = 0.02$, is very close to ridge regression and, correspondingly, there is only a slight degree of feature selection. The average number of nonzero coefficients over the 200 training samples is 245.3 (out of 271). Moreover, no feature is consistently excluded over the 200 samples. So, while on average the models achieve remarkable accuracy in distinguishing PD patients from healthy controls, the large number of contributing variables involved does not advance our goal of identifying potential biomarkers that can be considered in future research to explore possible biological mechanisms. Therefore, despite attaining high prediction accuracy, our pursuit of potential markers prompts us to seek additional parsimony.

We proceed by constructing a bounded search region, $B = \{(\alpha, \lambda) \mid p \leq 75, AUC \geq 0.9\}$, for the tuning parameters to induce parsimony (see **Figure 4A**). The white boundary partitions the search region so that the area to the right has, on average, $p \leq 75$ variables. To the left, the black trace defines the area with average $AUC \geq 0.9$ to ensure that we retain a sufficient level of accuracy. The operating points between the two lines make up set B .

From the previously described elastic net with repeated two-fold cross-validation, **Figure 5A** shows scatter plots of the mean absolute coefficient of each standardized feature vs. the proportion of instances the feature is retained (i.e., has a nonzero coefficient) over the 200 training samples. Each plot corresponds to operating points A, B, C, and D in **Figure 4A**. At point A, we see that the mean absolute coefficient values are relatively large, and that every feature is selected 75% or more of the 200 trials. Points B, C, and D explore different extremes of our bounded search region. As alpha increases, the rate at which features are selected decreases. At large lambda (point B), the mean coefficient values are small. In each panel, the horizontal line indicates the threshold $\xi_{0.10}$ signifying the top 10% with the strongest predictive features (based on mean absolute coefficient value). **Figure 5B** shows an enlarged plot at point E, a representative point near the middle of the search region. Using color, the plot illustrates the distribution associated with the different modalities. At point E, modalities FC, SC, and VBM yield the most predictive features.

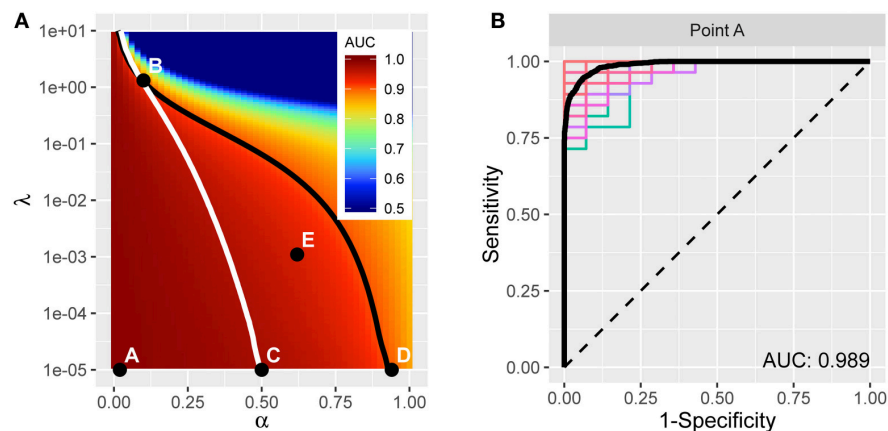


FIGURE 4 | (A) AUC for different tuning parameters, with each point averaged over 100 applications of two-fold cross-validation. The point A reflects the tuning parameter value yielding the maximum AUC, and is depicted in the curves in **(B)**. The traces define a restricted space of tuning parameters. Above and to the right of the white trace yields no more than an average of 75 predictors, and below and to the left of the black trace reflects at least 0.90 AUC on average. **(B)** ROC curve (in black) reflecting high prediction accuracy based on 271 imaging predictors; AUC is 0.989. The colored curves highlight the variability associated with each separate CV sample.

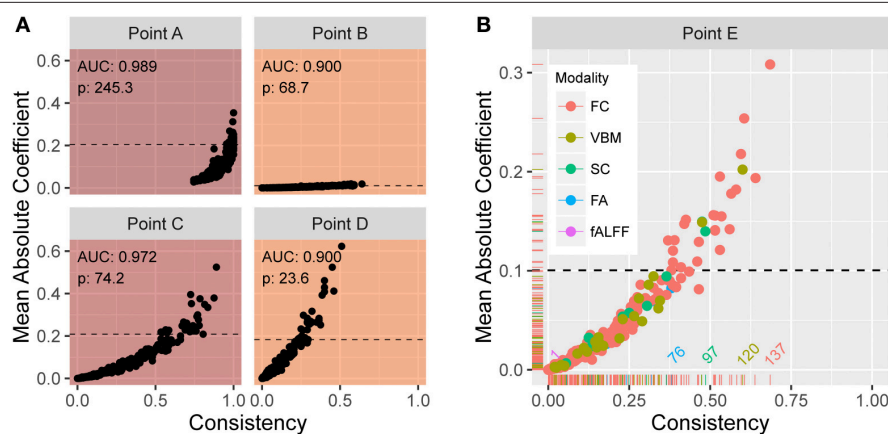


FIGURE 5 | (A) Plots of the mean absolute coefficient (standardized) vs. the proportion of times the feature is retained over 200 training samples at (α, λ) corresponding to points (B–E) in **Figure 4A**. The enlarged plot shown in **(B)** is point E from **Figure 4A**, with colors depicting the modality. The reference lines in all plots reveal the 10% of values with strongest predictive power over the training samples. At point E, modalities FC, SC, and VBM yield the most predictive features.

Considering all operating points in the search area, the resulting set C includes 24 features, each of which is consistently among the most predictive for at least 90% ($\tau_C = 0.9$) of the operating points in B. The features are listed in **Table 2** and include 21 FC, 1 SC, and 2 regional VBM measures.

The 24 features contribute extremely strong predictive power. Using logistic regression, still controlling for head coil, sex, and age, one can achieve perfect separation between PD patients and HC using subsets of as few as three of these multimodal imaging features. In fact, out of all possible three-feature models, three of them achieve perfect separation between the groups, and comprise an aggregate of eight separate features. The three models and the associated map of features are presented in **Figure 6**. No model of less than

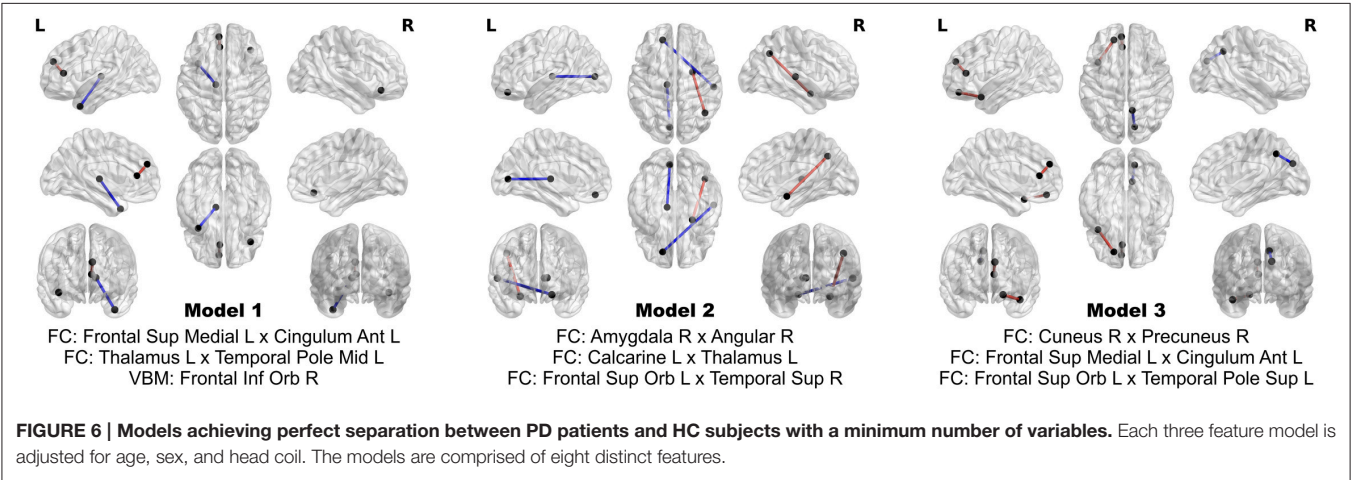
three features achieves perfect separation; however many such models exist when more than three out of the 24 features are considered.

Performing univariate screening as a separate step prior to cross validation of the elastic net could potentially impact variable selection and classification performance. To examine this possibility for our analysis, we performed univariate screening using the same 200 (2x100) cross validation training samples constructed in the elastic net stage. Note that for this sensitivity analysis, we did not additionally implement our bootstrap procedure within each of the cross validation samples, given the computational cost. We track the number of times each feature's corresponding p -value falls below the designated modality-specific threshold across the 200 samples. Features that pass the threshold in 75% or more of the 200 cross

TABLE 2 | List of 24 features that are consistently the most predictive across a restricted tuning parameter space for (α, λ) in the elastic net procedure.

Feature	Upper 10% predictive strength over (α, λ) (%)	Upper 10% selection rate over (α, λ) (%)	Direction of effect
FC: Frontal Sup Orb L x Temporal Pole Sup L	94.4	99.4	1
FC: Amygdala R x Angular R*	97.5	46.0	1
FC: Amygdala R x Lingual L	100.0	100.0	1
FC: Calcarine L x Thalamus L	100.0	100.0	-1
FC: Cingulum Ant R x Cingulum Post L	100.0	100.0	1
FC: Cuneus R x Precuneus R	100.0	100.0	-1
FC: Frontal Inf Orb R x Temporal Mid R	99.7	90.7	1
FC: Frontal Inf Orb R x Temporal Mid L	100.0	100	1
FC: Frontal Inf Tri R x Temporal Pole Mid R	100.0	100.0	1
FC: Frontal Mid Orb L x Hippocampus L	100.0	100.0	1
FC: Frontal Sup Medial L x Cingulum Ant L	95.2	0.85	1
FC: Frontal Sup Orb L x Insula L	100.0	100.0	1
FC: Frontal Sup Orb L x Parietal Inf L	100.0	100.0	1
FC: Frontal Sup Orb L x Temporal Sup R	100.0	100.0	-1
FC: Occipital Mid L x Occipital Inf R	99.1	98.0	1
FC: Occipital Sup L x Temporal Mid R	100.0	100.0	-1
FC: Occipital Sup R x Precuneus R	100.0	61.2	-1
FC: Temporal Mid R x Temporal Pole Mid R	100.0	100.0	1
FC: Temporal Sup R x Temporal Pole Mid L	100.0	99.7	1
FC: Thalamus L x Temporal Pole Mid L	100.0	100.0	-1
SC: Calcarine L x Precuneus_R	100.0	87.5	-1
VBM: Frontal Inf Orb R	100.0	100.0	-1
VBM: Frontal Mid R	100.0	98.9	-1

Predictive strength, for a given (α, λ) , was computed as the mean absolute coefficient (normalized) across 200 training samples. 18 features were retained at 100% of the tuning parameter values considered. The list includes 21 FC features, 2 regional VBM measures, and 1 SC measure. *Two distinct FC links between these regions.



validation samples are regarded to pass univariate screening in our sensitivity analysis.

Nearly all (23 of 24) features that emerged from the elastic net stage in our original analysis were also selected by this revised cross-validation univariate screening. The excluded feature, the functional connectivity between the right frontal inferior orb and the middle temporal lobe, fell just below our threshold by being selected in only 73% of the 2×100 cross-validation trials. Thus our findings suggest that the final panel of 24 features is not

substantially impacted by the decoupling of screening and cross validation, perhaps buffered by the addition of our bootstrap screening procedure which accounts for sampling variability.

High dimensional prediction and classification methods are subject to inflated measures of accuracy for the particular data set under consideration, resulting in findings that are not reproducible on independent data sets. We take measures to minimize the risk of overfitting and to assess the potential influence of overfitting in our analysis. In particular, we perform

iterated two-fold cross-validation in the elastic net procedure with 271 variables as outlined above. To assess the presence and potential influence of overfitting, we conduct a null randomization experiment in which we combine all subjects and then randomly assign the subjects to one of two groups, with the group sizes matched to the actual sizes of the PD and control groups in our experimental data. In our null data, one would not expect to observe systematic between-group differences given the random mixing of PD and control subjects. We repeat our analysis on the null data and, as expected, obtain an ROC curve that roughly tracks a 45° line (chance), giving no indication of inflated accuracy due to our methods.

We note that this null hypothesis verification was performed only within the cross-validated elastic net analysis, using the 271 features that had previously passed screening and hence been deemed to be strongly associated to disease status. Given the massive number of features relative to the number of subjects in the analysis, it is entirely possible that other sets of features could be found, which have high explanatory affinity for random (perhaps meaningless) subject groupings. We assume that the PD (vs. control) labels reflect true manifestation of disease and thus that the identified features are strong candidates for PD biomarkers.

DISCUSSION

Our analysis provides a broad multimodal view of prevailing alterations in PD, which serve as accurate and reliable predictors. We identify 24 neural manifestations of PD, which contribute extremely strong predictive power in these subjects (**Table 2**). FC from resting-state fMRI emerges as the most prominent modality. Decreased SC between the left calcarine area and the right precuneus also is an indicator aiding dissociation of PD and HC subjects. VBM calculated from anatomical T1-MRI scans also contributes to accurate prediction, with PD patients revealing reduced volume in the right inferior orbital frontal cortex (OFC) and right middle frontal gyrus. Our findings support a previous report of volumetric changes in gray matter associated with PD, including bilateral OFC and the right inferior frontal gyrus (rIFG) (Xia et al., 2013) and are consistent with reports of cortical thinning in PD (without dementia) in the middle frontal gyrus and other regions including inferior and superior parietal areas, superior frontal, superior temporal, precuneus, pre- and postcentral, and fusiform regions (Zhang et al., 2015).

Bilaterally, the middle temporal pole (MTP) exhibited strong discriminatory power and consistency. For PD patients, the right MTP shows increased FC with the rIFG pars triangularis (rIFG-PT) and with the right middle temporal gyrus (MTG). The left MTP exhibits decreased FC with the left thalamus in PD patients and increased FC the right superior temporal sulcus (STS). The OFC, which is linked to inhibitory control, exhibits functional connections with several regions that are predictive of PD. Right inferior areas of the OFC show increased FC bilaterally with the MTG. The left middle areas of the OFC show increased FC with the left hippocampus. The left superior OFC exhibits increased FC in PD patients with the left insula, the left inferior

parietal region, and the left superior temporal pole; and decreased FC with the right STS. Our analysis also reveals increased FC between the left medial superior frontal gyrus, which includes the supplementary motor area (SMA) and the preSMA, and the left ACC. These FC alterations all yield strong power to dissociate PD patients from controls.

PD symptoms have been linked to FC between the pars triangularis and the orbito-frontal cortex, specifically with FC shown to be positively associated with the Movement Disorder Society (MDS) Unified Parkinson's Disease Rating Scale (UPDRS), part II, entitled Motor Aspects of Experiences of Daily Living (Yoo et al., 2015). Also, some of the OFC functional connections map anatomical tracts known from macaque monkey studies between the orbitofrontal cortex and limbic areas including insular cortex and the hippocampus (Cavada et al., 2000).

Discriminatory power is also drawn from decreased FC in PD patients between the left calcarine and thalamus and between the right cuneus and precuneus as well as increased FC for PD patients between right anterior cingulate cortex (ACC) and the posterior cingulate cortex (PCC) and between the right amygdala and both the left lingual gyrus and the right angular gyrus. In contrast to our work examining specific connections between pairs of brain regions, graph-theoretic approaches seek to characterize whole brain topological properties of brain networks (Simpson et al., 2013). In this complementary view, Göttlich et al. (2013) show that the degree of whole-brain connectivity was decreased in the occipital lobe (cuneus and calcarine), but increased in the superior parietal cortex, PCC, supramarginal gyrus and supplementary motor area.

Several of the above regions have been identified for PD-related alterations or dysfunction. The amygdala plays a key role in memory, decision-making, and emotional response. The amygdala may undergo a loss of gray-matter volume in PD due to neurodegeneration (Harding et al., 2002). Hu et al. (2015) found that, relative to healthy controls, depressed PD patients exhibited decreased right amygdala FC with the left gyrus rectus, left inferior OFC, and right putamen. The FC alterations in the amygdala may be driven by the severe pathological changes that occur in this region and the major projections to the prefrontal cortex and limbic system (hippocampus and entorhinal region), among others (Braak et al., 1994). Van Eimeren et al. (2009) detected different deactivation patterns in the PCC and the precuneus in PD patients relative to healthy controls.

Our analysis considers extremely high-dimensional data and eventually selects a small number of variables, representing just 0.051% of the original features considered, with subsets representing 0.0064% of the features achieving perfect separation of the PD patients and the HC subjects. An important step in the route to developing reliable biomarkers is to validate the identified features in independent data sets. We take many steps to encourage reproducibility here within our sample, but ultimately adoption of biomarkers requires external validation. It is likely that many other useful predictors are present in the data, so our results do not preclude the possibility that important predictive information may be gleaned from a neuroimaging modality/feature that was ultimately excluded from our final

model. Moreover, our data representations may have excluded potentially useful markers. For example, we focused our analysis on AAL ROIs, opting to maintain consistency of the regions across modalities (with possibly nested subregions). AAL regions are predominantly composed of gray matter and contain relatively less white matter (median across regions and subjects is roughly 15%). As such, we may have excluded potentially predictive markers from DTI-related measures (FA or SC) sampled from regions dominated by white matter. Also, the features extracted from the multimodal imaging data reflect particular characteristics at a selected spatial scale. Some data reduction is necessary, e.g., to limit the data from generating billions of features. We cannot determine in advance, which spatial scale will extract maximal information for the purpose of dissociating PD patients from controls.

AUTHOR CONTRIBUTIONS

FB conceptualized methodology and guided evolution of the research. FB also wrote the majority of the paper. DD processed the imaging data and performed statistical analyses on the results.

REFERENCES

- Andersson, J. L. R., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* 20, 870–888. doi: 10.1016/S1053-8119(03)00336-7
- Behrens, T. E. J., Johansen-Berg, H., Jbabdi, S., Rushworth, M. F. S., and Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations. *What can we gain?* *Neuroimage* 23, 144–155. doi: 10.1016/j.neuroimage.2006.09.018
- Braak, H., Braak, E., Yilmazer, D., de Vos, R. A., Jansen, E. N., Bohl, J., et al. (1994). Amygdala pathology in Parkinson's disease. *Acta Neuropathol.* 88, 493–500. doi: 10.1007/BF00296485
- Braak, H., Del Tredici, K., Rüb, U., de Vos, R. A., Jansen Steur, E. N., and Braak, E. (2003). Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* 24, 197–211. doi: 10.1016/S0197-4580(02)00065-9
- Cavada, C., Company, T., Tejedor, J., and Cruz-Rizzolo, R. J., Reinoso-Suarez (2000). The anatomical connections of the macaque monkey orbitofrontal cortex. *Cereb. Cortex* 10, 220–242. doi: 10.1093/cercor/10.3.220
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014
- Del Tredici, K., and Braak, H. (2013). Dysfunction of the locus coeruleus-norepinephrine system and related circuitry in Parkinson's disease-related dementia. *J. Neurol. Neurosurg. Psychiatry* 84, 774–783. doi: 10.1136/jnnp-2011-301817
- Du, G., Lewis, M. M., Styner, M., Shaffer, M. L., Sen, S., Yang, Q. X., et al. (2011). Combined R2* and diffusion tensor imaging changes in the substantia nigra in Parkinson's disease. *Mov. Disord.* 26, 1627–1632. doi: 10.1002/mds.23643
- Fearnley, J. M., and Lees, A. J. (1991). Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain* 114, 2283–2301
- Fiandaca, M. S., Mapstone, M. E., Cheema, A. K., and Federoff, H. J. (2014). The critical need for defining preclinical biomarkers in Alzheimer's disease. *Alzheimers Dement.* 10(Suppl. 3), 196–212. doi: 10.1016/j.jalz.2014.04.015
- Foroud, T., Smith, D., Jackson, J., Verbrugge, J., Halter, C., Wetherill, L., et al. (2015). Novel recruitment strategy to enrich for LRRK2 mutation carriers. *Mol. Genet. Genomics* 3, 404–412. doi: 10.1002/mgg3.151
- Gaser, C. (2010). *Author Archive for Christian Gaser. Structural Brain Mapping Group*. Available online at: <http://www.neuro.uni-jena.de/vbm/>
- Göttlich, M., Münte, T. F., Heldmann, M., Kasten, M., Hagenah, J., and Krämer, U. M. (2013). Altered resting state brain networks in Parkinson's disease. *PLoS ONE* 8:e77336. doi: 10.1371/journal.pone.0077336
- Harding, A. J., Stimson, E., Henderson, J. M., and Halliday, G. M. (2002). Clinical correlates of selective pathology in the amygdala of patients with Parkinson's disease. *Brain* 125, 2431–2445. doi: 10.1093/brain/awf251
- Hu, X., Song, X., Yuan, Y., Li, E., Liu, J., Liu, W., et al. (2015). Abnormal functional connectivity of the amygdala is associated with depression in Parkinson's disease. *Mov. Disord.* 30, 238–244. doi: 10.4137/JEN.S33444
- Iranzo, A., Tolosa, E., Gelpi, E., Molinuevo, J. L., Valldeoriola, F., Serradell, M., et al. (2013). Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study. *Lancet Neurol.* 12, 443–453. doi: 10.1016/S1474-4422(13)70056-5
- Kahan, J., Urner, M., Moran, R., Flandin, G., Marreiros, A., Mancini, L., et al. (2014). Resting state functional MRI in Parkinson's disease: the impact of deep brain stimulation on 'effective' connectivity. *Brain* 137, 1130–1144. doi: 10.1093/brain/awu027
- Kalia, L. V., and Lang, A. E. (2015). Parkinson's disease. *Lancet* 386, 892–912. doi: 10.1016/s0140-6736(14)61393-3
- Lee, E. Y., Sen, S., Eslinger, P. J., Wagner, D., Shaffer, M. L., Kong, L., et al. (2013). Early cortical gray matter loss and cognitive correlates in non-demented Parkinson's patients. *Parkinsonism Relat. Disord.* 19, 1088–1093. doi: 10.1016/j.parkreldis.2013.07.018
- Parkinson's Disease Foundation (2015). *Statistics on Parkinson's*. Available online at: http://www.pdf.org/en/parkinson_statistics
- Postuma, R. B., Iranzo, A., Hogl, B., Arnulf, I., Ferini-Strambi, L., Manni, R., et al. (2015). Risk factors for neurodegeneration in idiopathic rapid eye movement sleep behavior disorder: a multicenter study. *Ann. Neurol.* 77, 830–839. doi: 10.1002/ana.24385
- Simpson, S. L., Bowman, F. D., and Laurienti, P. J. (2013). Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Stat. Surv.* 7, 1–36. doi: 10.1214/13-SS103
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, 208–219. doi: 10.1016/j.neuroimage.2004.07.051
- Strimbu, K., and Tavel, J. A. (2010). What are Biomarkers? *Curr. Opin. HIV AIDS* 5, 463–466. doi: 10.1097/COH.0b013e32833ed177
- Tolosa, E., and Pont-Sunyer, C. (2011). Progress in defining the premotor phase of Parkinson's disease. *J. Neurol. Sci.* 310, 4–8. doi: 10.1016/j.jns.2011.05.027

ACKNOWLEDGMENTS

This research was funded by a grant from the NINDS (U18 NS082143) at NIH as part of the Parkinson's Disease Biomarker Program. Funding support leading to the generation of the dataset came from the William N. and Bernice E. Bumpus Foundation Early Career Investigator Innovation Award (BFIA 2011.3, Huddleston), the Emory University Morris K. Udall Center for Parkinson's Disease Research (P50-NS071669), and the Emory Alzheimer's Disease Research Center (P50-AG025688). Additional significant contributions to the acquisition of the data were made by Stewart A. Factor and Rebecca McMurray of the Emory Department of Neurology, and by Jason Langley and Xiaoping Hu of the Emory Department of Biomedical Engineering.

- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Vaillancourt, D. E., Spraker, M. B., Prodoehl, J., Abraham, I., Corcos, D. M., Zhou, X. J., et al. (2009). High-resolution diffusion tensor imaging in the substantia nigra of de novo Parkinson disease. *Neurology* 72, 1378–1384. doi: 10.1212/01.wnl.0000340982.01727.6e
- Van Eimeren, T., Monchi, O., Ballanger, B., and Strafella, A. P. (2009). Dysfunction of the default mode network in Parkinson disease: a functional magnetic resonance imaging study. *Arch. Neurol.* 66, 877–883. doi: 10.1001/archneurol.2009.97
- Xia, J., Wang, J., Tian, W., Ding, H., Wei, Q., Huang, H., et al. (2013). Magnetic resonance morphometry of the loss of gray matter volume in Parkinson's disease patients. *Neural Regen. Res.* 8, 2557–2565. doi: 10.1371/journal.pone.0131133
- Yoo, K., Chung, S. J., Kim, H. S., Choung, O.-H., Lee, Y. B., Kim, M. J., et al. (2015). Neural substrates of motor and non-motor symptoms in Parkinson's disease: a resting fMRI study. *PLoS ONE* 10:e0125455. doi: 10.1371/journal.pone.0125455
- Zarei, M., Ibarretxe-Bilbao, N., Compta, Y., Hough, M., Junque, C., Bargallo, N., et al. (2013). Cortical thinning is associated with disease stages and dementia in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 84, 875–881. doi: 10.1007/s00429-014-0785-x
- Zhang, L., Wang, M., Sterling, N., Lee, E., Eslinger, P., Wagner, D., et al. (2015). "Cortical thinning and cognitive impairment in Parkinson's disease without dementia," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*.
- Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B* 67, 301–320. doi: 10.1111/j.1467-9868.2005.00503.x
- Zou, Q. H., Zhu, C. Z., Yang, Y., Zuo, X. N., Long, X. Y., Cao, Q. J., et al. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *J. Neurosci. Methods* 172, 137–141. doi: 10.1016/j.jneumeth.2008.04.012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Bowman, Drake and Huddleston. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read,
for greatest visibility



COLLABORATIVE PEER-REVIEW

Designed to be rigorous
– yet also collaborative,
fair and constructive



FAST PUBLICATION

Average 85 days from
submission to publication
(across all journals)



COPYRIGHT TO AUTHORS

No limit to article
distribution and re-use



TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles



SUPPORT

By our Swiss-based
editorial team



IMPACT METRICS

Advanced metrics
track your article's impact



GLOBAL SPREAD

5'100'000+ monthly
article views
and downloads



LOOP RESEARCH NETWORK

Our network
increases readership
for your article

Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on

