

Antonio Canale  
Daniele Durante  
Lucia Paci  
Bruno Scarpa *Editors*

# Studies in Neural Data Science

StartUp Research 2017, Siena, Italy,  
June 25–27

# **Springer Proceedings in Mathematics & Statistics**

Volume 257

## **Springer Proceedings in Mathematics & Statistics**

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including operation research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

More information about this series at <http://www.springer.com/series/10533>

Antonio Canale · Daniele Durante  
Lucia Paci · Bruno Scarpa  
Editors

# Studies in Neural Data Science

StartUp Research 2017, Siena, Italy,  
June 25–27



Springer

*Editors*

Antonio Canale

Department of Statistical Sciences  
University of Padova  
Padua, Italy

Lucia Paci

Department of Statistical Sciences  
Università Cattolica del Sacro Cuore  
Milan, Italy

Daniele Durante

Department of Decision Sciences  
Bocconi University  
Milan, Italy

Bruno Scarpa

Department of Statistical Sciences  
University of Padova  
Padua, Italy

ISSN 2194-1009

Springer Proceedings in Mathematics & Statistics

ISBN 978-3-030-00038-7

<https://doi.org/10.1007/978-3-030-00039-4>

ISSN 2194-1017 (electronic)

ISBN 978-3-030-00039-4 (eBook)

Library of Congress Control Number: 2018961221

Mathematics Subject Classification (2010): S11001, S17030, B18006, L15020

© Springer Nature Switzerland AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

This volume contains a collection of peer-reviewed articles arising from *StartUp Research*. This meeting took place on June 25–27, 2017, at the ancient Certosa di Pontignano (Pontignano Charterhouse), a few kilometers from Siena (Italy). *StartUp Research* was a satellite event of the Statistical Conference of the Italian Statistical Society, held in Florence (Italy) in June 2017. The event was additionally endorsed by the young group of the Italian Statistical Society (<https://youngsis.github.io/>), whose aim is to promote activities and provide a social networking platform for early-career researchers in statistics.

*StartUp Research* was a stimulating experience. It brought together 28 early-career researchers in statistics and seven international professors with the common task of developing novel statistical methods for complex and multimodal brain imaging data. It is, in fact, increasingly common in neuroscience to monitor the brain activity of each subject under different imaging technologies. This motivates the development of novel statistical methods for joint modeling of complex and multimodality data on brain function and structure. The junior researchers, divided into seven groups, focused on brain imaging data from a study of the Enhanced Nathan Kline Institute-Rockland (NKI1) project ([http://fcon\\_1000.projects.nitrc.org/indi/enhanced/](http://fcon_1000.projects.nitrc.org/indi/enhanced/)). This pilot study comprises multimodal imaging data and subject-specific covariates for 24 individuals. In particular, for each subject, the following data are available:

- Structural networks measuring, from diffusion tensor imaging, white matter fiber interconnections among brain regions of interest.
- Functional activity data measuring the dynamic activity of each brain region through changes in the blood oxygen level-dependent signal during resting state functional magnetic resonance imaging.
- Functional networks denoting regions' synchronization in brain activity.

Spatial information on the brain regions of interest and subject-specific data on age, handedness, and psychological traits are also provided. The imaging data were pre-processed and generously provided by Greg Kiar and Eric Bridgeford from NeuroData at Johns Hopkins University, who are gratefully acknowledged.

Motivated by the above dataset, the groups proposed stimulating methods during *StartUp Research* and continued their studies in the following year. More specifically, the contribution “[Understanding Dependency Patterns in Structural and Functional Brain Connectivity Through fMRI and DTI Data](#)” leverages latent variable models and dynamic Bayesian networks to learn, possibly similar, patterns in brain structural and functional connectivity. The contribution “[Hierarchical Graphical Model for Learning Functional Network Determinants](#)” instead adopts a modular approach which combines smoothing procedures, graphical models, and regression methods to relate functional connectivity with regions and subject-specific features. Different directions in the analysis of brain interconnections are proposed in “[Three Testing Perspectives on Connectome Data](#)”. The first focuses on learning structural restrictions in brain functional activity. The second aims at estimating the effective number of white matter fibers via parsimonious models, while the third studies group differences in brain connectivity with subjects’ traits under an object-oriented perspective. Also the work “[An Object Oriented Approach to Multimodal Imaging Data in Neuroscience](#)” analyzes the human brain data as object-valued and provides a wide set of procedures, including clustering, low-dimensional embeddings, and hypothesis testing, to obtain coherent findings in neuroscience. In a similar research direction, the contribution “[Curve Clustering for Brain Functional Activity and Synchronization](#)” focuses on appropriate methods to infer grouping structures and functional outliers in fMRI trajectories. These data are further explored in “[Robust Methods for Detecting Spontaneous Activations in fMRI Data](#)” via novel filtering methods which incorporate heavier tails than classical Gaussian assumptions. Parsimonious, yet flexible, Bayesian dynamic latent factor models are instead considered in the contribution “[Hierarchical Spatio-Temporal Modeling of Resting State fMRI Data](#)” to infer spatial and temporal effects of brain functional activity across multiple regions. A final article by Michele Guindani and Marina Vannucci summarizes the different proposals and opens toward new stimulating research directions in this field.

We would like to thank the early-career participants, Emanuele Aliverti, Gaia Bertarelli, Alessandra Cabassi, Alessia Caponera, Andrea Cappozzo, Alessandro Casa, Alice Corbella, Federico Crescenzi, Marta Crispino, Silvia D’Angelo, Francesco Denti, Jacopo Di Iorio, Roberta Falcone, Federico Ferraccioli, Matteo Fontana, Laura Forastiere, Francesca Gasperoni, Anastasiia Gorshechnikova, Tullia Padellini, Sally Paganin, Michele Peruzzi, Alexios Polymeropoulos, Saverio Rancati, Tommaso Rigon, Dutta Ritabrata, Massimiliano Russo, Andrea Sottosanti, and Marco Stefanucci for their enthusiasm and dedication to this stimulating experience. Also, the group leaders Alessio Farcomeni, Alan Gelfand, Alessandra Luati, Antonietta Mira, Piercesare Secchi, Marian Scott, and Ernst Wit are warmly acknowledged for their fundamental and inspiring contribution in leading the groups, both personally and scientifically.

Finally, we would like to thank the Italian Statistical Society, the Department of Economics and Statistics of the University of Siena, and the Department of Statistical Sciences of the University of Bologna for supporting *StartUp Research*. We are also grateful to the referees for their thoughtful revisions.

Padua, Italy  
Milan, Italy  
Milan, Italy  
Padua, Italy  
July 2018

Antonio Canale  
Daniele Durante  
Lucia Paci  
Bruno Scarpa

# Contents

<b>Understanding Dependency Patterns in Structural and Functional Brain Connectivity Through fMRI and DTI Data .....</b>	1
Marta Crispino, Silvia D'Angelo, Saverio Ranciatì and Antonietta Mira	
<b>Hierarchical Graphical Model for Learning Functional Network Determinants .....</b>	23
Emanuele Aliverti, Laura Forastiere, Tullia Padellini, Sally Paganin and Ernst Wit	
<b>Three Testing Perspectives on Connectome Data .....</b>	37
Alessandra Cabassi, Alessandro Casa, Matteo Fontana, Massimiliano Russo and Alessio Farcomeni	
<b>An Object Oriented Approach to Multimodal Imaging Data in Neuroscience .....</b>	57
Andrea Cappozzo, Federico Ferraccioli, Marco Stefanucci and Piercesare Secchi	
<b>Curve Clustering for Brain Functional Activity and Synchronization .....</b>	75
Gaia Bertarelli, Alice Corbella, Jacopo Di Iorio, Anastasia Gorshechnikova and Marian Scott	
<b>Robust Methods for Detecting Spontaneous Activations in fMRI Data .....</b>	91
Francesca Gasperoni and Alessandra Luati	
<b>Hierarchical Spatio-Temporal Modeling of Resting State fMRI Data .....</b>	111
Alessia Caponera, Francesco Denti, Tommaso Rigon, Andrea Sottosanti and Alan Gelfand	
<b>Challenges in the Analysis of Neuroscience Data .....</b>	131
Michele Guindani and Marina Vannucci	

# About the Editors

**Antonio Canale** is an Assistant Professor of Statistics in the Department of Statistical Sciences, University of Padova (Italy). His research areas cover Bayesian non-parametric methods, functional data analysis, statistical learning, and data mining. He is the author of a number of papers on methodological and applied statistics and has served on the scientific committees of national and international conferences. He was the coordinator of the young group of the Italian Statistical Society (y-SIS) in 2015.

**Daniele Durante** is an Assistant Professor of Statistics in the Department of Decision Sciences, Bocconi University (Italy), and a Research Affiliate at the Bocconi Institute for Data Science. His research is characterized by an interdisciplinary approach at the intersection of Bayesian methods, modern applications, and statistical learning to develop flexible and computationally tractable models for complex data. He is the coordinator of the young group of the Italian Statistical Society (y-SIS).

**Lucia Paci** is an Assistant Professor of Statistics in the Department of Statistical Sciences, Università Cattolica del Sacro Cuore, Milan (Italy). Her research focuses mainly on spatial and spatiotemporal modeling under the Bayesian framework, with applications in the environmental and economic sciences. She was the coordinator of the young group of the Italian Statistical Society (y-SIS) in 2016.

**Bruno Scarpa** is an Associate Professor of Statistics in the Department of Statistical Sciences, University of Padova (Italy). He teaches data mining at the master level and statistical methods for big data at the undergraduate level. His research interests include methodological developments motivated by real data applications. He is the author and coauthor of numerous papers and books in the fields of methodological and applied statistics and data mining.

# Understanding Dependency Patterns in Structural and Functional Brain Connectivity Through fMRI and DTI Data



Marta Crispino, Silvia D'Angelo, Saverio Ranciati and Antonietta Mira

**Abstract** Neuroscience and neuroimaging have been providing new challenges for statisticians and quantitative researchers in general. As datasets of increasing complexity and dimension become available, the need for statistical techniques to analyze brain related phenomena becomes prominent. In this paper, we delve into data coming from functional Magnetic Resonance Imaging (fMRI) and Diffusion Tensor Imaging (DTI). The aim is to combine information from both sources in order to learn possible patterns of dependencies among regions of interest (ROIs) of the brain. First, we infer positions of these regions in a latent space, using the observed structural connectivity provided by the DTI data, to understand if physical spatial coordinates suitably reflect how ROIs are effectively interconnected. Secondly, we inspect Granger causality in the fMRI data in order to capture patterns of activations between ROIs. Then, we compare results from the analysis on these datasets, to find a link between functional and structural connectivity. Preliminary findings show that latent space positions well reflect hemisphere separation of the brain but are not perfectly connected to all the other structural partitions (that is, lobe, cortex, etc.); furthermore, activations of ROIs inferred from fMRI data are tied to observed structural connections derived from DTI scans.

---

M. Crispino  
Univ. Grenoble Alpes, Inria, CNRS, LJK, 38000 Grenoble, France  
e-mail: [marta.crispino@inria.fr](mailto:marta.crispino@inria.fr)

S. D'Angelo  
Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy  
e-mail: [silvia.dangelo@uniroma1.it](mailto:silvia.dangelo@uniroma1.it)

S. Ranciati (✉)  
Department of Statistical Sciences, University of Bologna, Bologna, Italy  
e-mail: [saverio.ranciati2@unibo.it](mailto:saverio.ranciati2@unibo.it)

A. Mira  
Institute of Computational Science, Università della Svizzera italiana, Lugano, Switzerland  
e-mail: [antonietta.mira@usi.ch](mailto:antonietta.mira@usi.ch)

A. Mira  
Department of Science and High Technology, Università dell’Insubria, Como, Italy

**Keywords** Network analysis · Resting state fMRI · DTI · Latent space models  
Penalized weighted regression

## 1 Motivating Real World Dataset

Advances in neuroimaging have led to an increase in the availability of data to study complex systems (for instance, neurological processes in human brain). It is now possible to collect data considering different aims and assimilating different sources, a strategy that better captures the underlying dynamics of the phenomenon at study. The main interest lies in unraveling the mechanisms originating structural and functional brain activity, and simultaneously in understanding how these aspects are intertwined with patients covariates: for instance, how significant and relevant are the differences in brain connectivity and activity among subjects with heterogeneous characteristics. From different available multimodal brain imaging frameworks, here we focus on functional Magnetic Resonance Imaging (fMRI), and Diffusion Tensor Imaging (DTI).

The datasets we consider for the analysis were collected during a pilot study of the Enhanced Nathan Kline Institute-Rockland Sample project (information about the project itself is available at [http://fcon\\_1000.projects.nitrc.org/indi/enhanced/](http://fcon_1000.projects.nitrc.org/indi/enhanced/)). Data consist of 24 subjects, whose brain activity and structural connectivity were captured through DTI and resting state fMRI scan. The raw data were preprocessed and the scanned areas of the brain were parceled to determine a set of regions of interest (ROIs). An overview of the preprocessing steps is given in [4, 29]. Interest in analyzing rich and significant data from neuroimaging received a huge boost in the last decades, through a significant spillover of network analysis into neuroscience [2]. In an attempt to both distill information from complex systems and to infer the main mechanisms underlying brain activity, methods and concepts from network analysis were used into the framework of brain data. Concepts and terms such as hubs, centrality, hierarchy, node connectedness, and so forth, became both vocabulary and methodological tools shared by network analysis and neuroscience communities. A comprehensive review of this bridging between network analysis and neuroimaging is provided in [2]; more recently, in [30] an overview on connection between network properties and brain imaging data is also discussed, with emphasis on detection of neurological disorders via changes in the network structure itself. In [25, 29], Statistics and network science are directly tied to the study of functional and structural connectivity, and how both could help in deepening our understanding of interactions and, possibly, causality, among regions of the brain. In particular, for causal inference with an emphasis on fMRI data, refer to [22, 26].

From a statistical point of view, we find interest in exploiting—in a synergized approach—all the information at disposal. This can be done in many ways: for example, one could inject the number of white matter fibers for each pair of regions (DTI data) as a covariate information in a model capturing the correlation in the fMRI data. Alternatively, one might describe the statistical properties of an assumed underlying

network model originating the DTI dataset, then use some network's properties as aid in discovering (in fMRI data) patterns of synchronization between regions of the brain. Given the variety of questions and ways to address them, we adopt a statistical framework unifying these two ways of looking at brain connectivity, with the intent to assimilate different kinds of information captured by these different technologies, DTI and fMRI. The main goal of this work is thus to combine results from structural and functional observed data, in order to enhance the interpretation of each separate findings. In particular, our aim is to assess if the two datasets give 'coherent' answers with respect to the behavior we expect from the phenomenon: the patterns of activation among ROIs from fMRI data should be tied to structural connectedness highlighted by DTI data. The two datasets obviously share information on the overall activity of the brain but, from a modeling perspective, they bring different contributions to the whole research framework. In particular, data for the structural connectivity should depict the 'hardware' reference for us to understand which regions of the brain are physically connected. On the other hand, data from resting fMRI should provide insights on the dynamic counterpart of signals commuting between ROIs, and thus a different aspect on the concept of connectivity. For these reasons, we consider separate statistical models for the two available datasets.

First, we analyze the structural connectivity information provided by the DTI dataset, with the aid of models and tools coming from network analysis' framework. In particular, we investigate the idea that a statistical interpretation of the topology of the network differs from the physical observed topology, represented by the spatial coordinates of the ROIs. For this reason, we resort on latent space models which allow us to infer positions of the ROIs, in terms of how close they are, directly from the data on white matter fibers and their structural connectivity. Latent space models for social network analysis have been introduced in [15]. In their work, the authors assume that the observed network data depend on a set of latent variables. Indeed, the nodes are assumed to be in a  $p$ -dimensional latent space. Then, the probability that two nodes are joined by an edge in the network depends on some function of the unknown latent coordinates of the nodes. In the case of *distance latent space models*, this function is generally assumed to be the Euclidean distance: the smaller the distance, the greater the probability of an edge. While for *projection latent space model*, the function considers the angle formed between two nodes in the bilinear latent space: the smaller the angle the higher the probability that the dyad is connected. This class of models can take into account some of the typical features of network analysis, such as the presence of degree heterogeneity and of group structure. Indeed [14] introduced sender and receiver effects for networks and [13] proposed a clustering model for the nodes in the latent space. A different approach that makes use of latent variable to model the dependency structure observed in network data is *stochastic block modeling* [21, 27]. Stochastic block models are particularly suited to cluster the nodes into blocks. From this rich literature, we mainly draw from the contribution of [15] in order to gain insights about spatial organizations of the ROIs.

Second, we deal with temporal information provided by the dataset on resting state fMRI, which comprises of time series of brain activity. In particular, we exploit a weighted linear regression model that encodes a type of causality between

observations at different time points and for different ROIs. This causality, called *Granger causality*, is tied to the estimation of parameters in the weighted linear regression. Granger Causality [9, 10] was first introduced in the Econometric literature, specifically developed for time series analysis. This notion of causality is grounded on the rather obvious intuition that the origin of a cause should necessarily precede in time its effect. In particular, it states that, given two sets of time series data,  $V_1$  and  $V_2$ , the series  $V_1$  (Granger) causes  $V_2$  if past values of  $V_1$  are helpful in predicting the future values of  $V_2$ . It is important to underline that Granger causality is not intended as causality in a deep sense: it just measures whether one time series is likely to influence the other one, that is, if  $V_1$  provides more information about future values of  $V_2$  than past values of  $V_2$  alone. As such, Granger causality not always overlaps with actual causality, but it is still a useful instrument to infer whether two series are related by some, generally unknown, phenomenon. Recently, the notion of Granger causality entered into the network literature on multivariate time series, with the objective of learning sparse sets of Granger causal relationships between univariate series [11, 12, 17, 24, 31]. In this paper, we exploit this notion in order to infer any existing relationship between different brain regions.

In light of the above discussion, the remainder of the paper is organized as follows: in Sect. 2 we provide some exploratory statistics on the datasets, to summarize the salient features of the data we are modeling; in Sect. 3, we outline a modeling approach to static network data along with the results we get by applying this methodology to the DTI dataset (Sect. 3.1). Then, a time-varying dynamic linear model formulation for the time-series dataset is presented in Sect. 4, together with preliminary results from the fMRI dataset (Sect. 4.1). Finally, in Sect. 5, we discuss the results so far obtained, also providing a glimpse of future developments.

## 2 Descriptive Analysis

As mentioned in Sect. 1, we here analyze two core datasets. The first one, DTI, refers to subjects' brain structural connectivity, measured using DTI. For the scan of each individual (and re-scan, if available), a  $70 \times 70$  matrix reports the observed count of white matter fibers connecting pairs of ROIs; a structural 'NA' (not assigned) value is reported for self-connectivity, which is the diagonal of the aforementioned matrix. The number of white matter fibers is thus an observed measure of connectivity between brain regions. To compare our results with a standard brain representation, we refer to the atlas for brain parcellation in [6] (also reported in Appendix 6), that has a total of  $n = 68$  ROIs. To match the analyzed regions with those of the Desikan atlas, two of the 70 ROIs in the data labeled as "unknown" are discarded, when comparing the results. Moreover, two of the regions in the data refer to the corpus callosum (left and right), while the Desikan atlas does not represent this region but instead considers the insula, which includes the corpus callosum together with the lateral ventricles. Therefore, when comparing our results with the Desikan atlas representation of the brain, we refer to the corpus callosum as insula. The

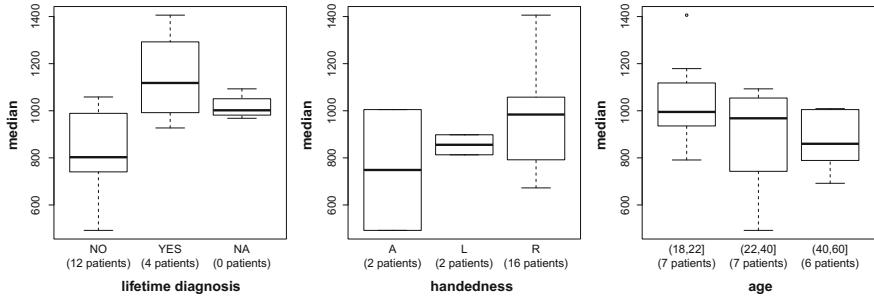
second dataset, fMRI, comprises of resting state dynamic functional activities of the ROIs, measured via blood-oxygen-level dependent (BOLD) technique through fMRI, at  $T = 404$  equally spaced time points with in-between lags of 1400 (ms). A third collateral dataset was produced by computing, from the fMRI dataset, time-wise correlations among the ROIs, resulting in a  $70 \times 70$  matrix for each subject (and each re-scan, when available). Additional information is provided in the form of covariates. For the subjects, characteristics available are: status of current (single episode/recurrent depressive disorder, cannabis abuse, anxiety, social phobia) or lifetime (alcohol abuse, drug dependence, Attention Deficit Hyperactivity Disorder, eating disorder, major depressive disorder) mental disorder; handedness (left, right, ambidextrous); age. For the ROIs, their lobe and hemisphere memberships are recorded, together with the physical spatial coordinates of the centroids used in the atlas. Information is not always available for all individuals and, in some cases, not every subject has a re-scan dataset to be paired with the original scan, causing missing values in the reported observations.

We here provide some descriptive statistics to familiarize the reader with the data at hand. In Table 1 we focus on the DTI dataset. The data for 4 subjects (labeled 6, 17, 20 and 22) are not available, and the disease diagnosis is missing for four patients (labeled 3, 4, 5, 6). We also notice that the seven patients who are diagnosed with lifetime disease, are also diagnosed with current disease. More importantly, three out of four patients with a diagnosis of current disease have missing data. The descriptive analysis in this section therefore focuses only on differences among patients with diagnosis of lifetime disease (four available out of seven), and patients with no such diagnosis (twelve available out of thirteen). The most salient feature of Table 1 is that the variability within subject of the number of white matter fibers is extremely high, with a range of values between zero and several thousands; distribution of the number of white matter fibers for each patient is highly skewed, and the median values are always much smaller than the mean values. We check if adjusting marginally (that is, one at a time) for covariates can help explain the distribution of the median number of white matter fibers across the patients. For example, the boxplot in Fig. 1 (left panel) represents the median number of white matter fibers, stratified by the lifetime disease diagnosis of the subjects (YES/NO).

From this plot, we notice that there is a difference, in terms of median number of white matter fibers, among patients with and without lifetime disease diagnosis. This result is confirmed by a Wilcoxon rank sum test which rejects the null hypothesis ( $p$ -value = 0.0077) of equal medians in the two groups. We also performed a Wilcoxon rank sum test to assess whether there is a difference between subjects with/without diagnosis of lifetime disease in terms of the percentage of zeros in the adjacency matrix (column 8 of Table 1). The null hypothesis in this case is not rejected. These findings suggest that the number of white matter fibers may be particularly informative about disease status, while the presence/absence of white matter fibers is not. There is a huge literature (see e.g [18, 20, 23]) that studies associations between mental disorders and DTI of white matter fibers. Analogous considerations can be

**Table 1** Summary table of the DTI dataset. Grey rows refer to patients who received a lifetime diagnosis. Empty cells correspond to non-available data. Columns (Statistics) from left to right: minimum value, maximum value, mean, median, standard deviation, interquartile range, coefficient of variation, percentage of zeros in the adjacency matrix

	Statistics			Covariates								
	Min	Max	Mean	Median	sd	iqr	cv	% of 0s	Age	Handed	Current diagnosis	Lifetime diagnosis
(computed on the non-0 entries)												
Patient 1	1	29078	2340	789	3814	2709	163	0.53	57	R	NO	NO
Patient 2	1	38041	2802	1009	4562	3315	163	0.63	52	R		
Patient 3	1	33081	2978	968	4822	3389	162	0.62	32	R		
Patient 4	1	39783	3068	1093	4771	4218	155	0.57	36	R		
Patient 5	1	34807	2840	995	4426	3428	156	0.67	22	R		
Patient 6									27	R	NO	NO
Patient 7	1	34481	2844	927	4738	3010	167	0.65	60	R	NO	YES
Patient 8	1	36809	3375	1057	5351	4101	159	0.61	21	R	YES	YES
Patient 9	1	37286	2366	898	4207	3019	164	0.60	21	L	NO	NO
Patient 10	1	33858	3123	1058	4949	3690	158	0.62	30	R	NO	NO
Patient 11	1	31682	2198	492	3848	2634	175	0.53	27	A	NO	NO
Patient 12	1	29288	2832	1005	4383	3260	155	0.63	48	A	NO	NO
Patient 13	1	41524	3530	1406	5340	4508	151	0.60	22	R	NO	YES
Patient 14	1	32377	2805	973	4446	3461	158	0.63	19	R	NO	NO
Patient 15	1	26388	2499	792	3826	3238	153	0.64	57	R	NO	NO
Patient 16	1	28797	2251	672	3484	3043	155	0.65	25	R	NO	NO
Patient 17									38	R	YES	YES
Patient 18	1	27501	1988	692	3070	2556	154	0.66	46	R	NO	NO
Patient 19	1	44968	3645	1179	6037	4500	166	0.62	22	R	NO	YES
Patient 20									32	L	YES	YES
Patient 21	1	36270	2888	791	4839	3627	168	0.60	22	R	NO	NO
Patient 22									42	R	YES	YES
Patient 23	1	36857	2585	813	4191	3095	162	0.60	31	L	NO	NO
Patient 24	1	41699	2990	1050	4862	3441	163	0.63	36	R	NO	NO



**Fig. 1** Boxplots of the median number of white matter fibers stratified by the lifetime disease diagnosis (left), handedness (middle), age (right)

drawn for other covariates, such as age or handedness<sup>1</sup> (see Fig. 1, middle and right panels). However, in this paper we do not model directly the number of white matter fibers (see Sect. 3), or the potential impact of covariates on these counts. Rather, we look at absence/presence of fibers since our focus is mainly in joining information coming from the two datasets. Nevertheless, we are willing to investigate this aspect in a future development of the current analysis (see also Sect. 5).

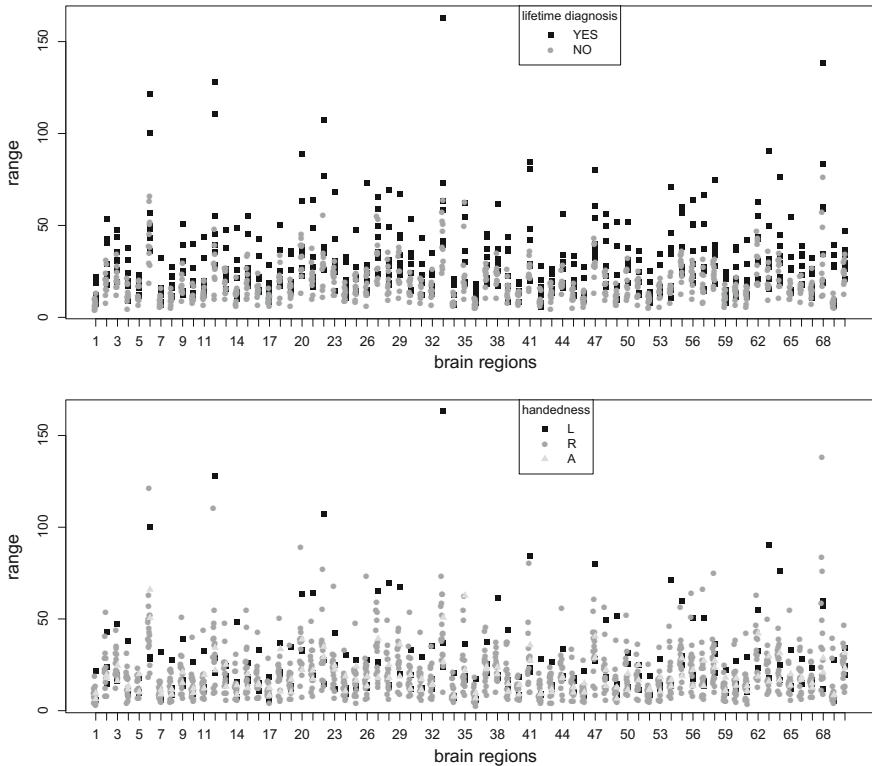
The fMRI dataset reports the dynamic activity time series of each brain region, for each subject in the study. Data for 2 subjects (1 and 21) are not available, whereas for other 11 subjects (labeled 4, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20) only the first scan is observed. In Fig. 2 we report some plots produced as follows:

- for each subject and each region of the brain, we compute the range of the activation levels along the time series;
- we then plot the obtained values versus the regions on the  $x$ -axis, with different symbols related to the user-specific covariates (as in the legend).

The range for patients who are diagnosed with a lifetime disease (black squares in the upper panel of Fig. 2) seems to be much higher than the one of the patients who did not receive a positive diagnosis (grey circles). Differences are noticeable also inspecting the handedness plot, in the lower panel of Fig. 2. We see that the range of activation is smaller for ambidextrous (A) patients (light-grey triangles).

The third dataset reports synchronization in brain activity for each pair of brain regions, obtained from the correlation in the dynamic functional activity of the fMRI dataset. In the heatmap of Fig. 3 is depicted the adjacency matrix corresponding to the functional network measuring correlation in brain activity between pairs of regions. The upper triangular panel is built averaging values of the patients who were not diagnosed a lifetime disease, while the lower triangular one refers to patients with lifetime disease. The structure of the two triangular panels is very similar, as indicated by the dark cells common in the two panels, and by the evident subdivision of both plots into three areas. However, it is evident an overall higher correlation for patients with diagnosis, in particular regarding pairs of brain regions that do not activate

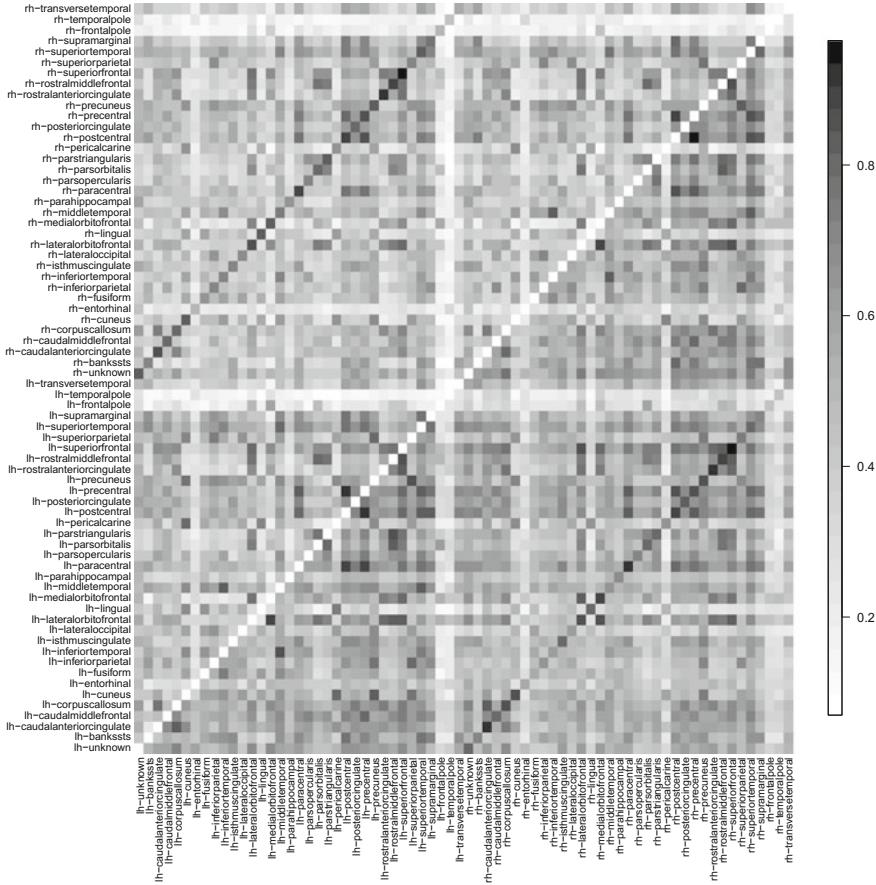
<sup>1</sup> Handedness is the dominance of one hand over the other, or the unequal distribution of fine motor skills between the left and right hands.



**Fig. 2** Range of activation for each brain region (x-axis). Top-panel: symbols refer to the lifetime diagnosis covariate; Bottom-panel: symbols refer to the handedness covariate: Left (L), Right (R), and Ambidextrous (A)

simultaneously in patients without diagnosis: in fact, many light cells of the upper triangular panel are darker in the lower triangular panel.

As a final insight, we show in Fig. 4 the adjacency matrix corresponding to the functional network measuring correlation in brain activity between pairs of regions, for a subject picked at random (patient labeled as 2). The black stars reported on some cells indicate whether there is physical connection in terms of white fibers: the presence of the star corresponds to a non-zero entry in the adjacency matrix of dataset DTI. From a rapid inspection, it seems that there is some relationship between the physical connectivity and the correlation in brain activity, as almost all dark cells are labeled with a star. We therefore check if this result holds for all the patients. We compute, separately for each patient, the fraction of dark starred cells, that is, cells corresponding to a value of correlation  $> 0.8$ , among all starred cells. After averaging across all the patients, we obtain an average fraction greater than 0.75, with no different pattern tied to covariate information such as mental disorder diagnosis or handedness. This means that, on average, three in four starred cells are also highly correlated in terms of brain activity as measured in the fMRI dataset.



**Fig. 3** Heatplot of the adjacency matrix of the network measuring correlation in brain activity between pairs of regions. Upper triangular: averaged across subjects who did not receive a diagnosis of lifetime disease (scan 1); Lower triangular: averaged across subjects who did receive a diagnosis of lifetime disease (scan 1)

### 3 Latent Space Model for DTI Dataset

We model each of the 20 patients, those whose data are available, separately: to lighten the notation, no subscript or superscript for the subject is used. Data consist of an  $n \times n$  adjacency matrix  $\mathbf{N}$ , collecting information on  $n = 70$  units which are the ROIs of the brain. Notice that, for the latent space analysis we use  $n = 70$  but to make the comparison with the Desikan atlas  $n$  is, ex-post, reduced to 68 by dropping the two ROI that in the original data are labeled as “unknown”.

Matrix  $\mathbf{N}$  is symmetric, with  $n(n - 1)/2$  distinct elements, where a generic element  $N_{i,j} \in \mathbb{N}$  represents the number of white matter fibers connecting the pair of



**Fig. 4** Heatplot of the adjacency matrix of the network measuring correlation in brain activity between pairs of regions, in patient 2. The black stars indicate whether there was physical connection in terms of white matter fiber (the presence of a star indicates a non-zero entry in the adjacency matrix of dataset DTI)

regions  $i$  and  $j$ . There is a correspondence between the matrix  $\mathbf{N}$  and the adjacency matrix of a weighted undirected graph, whose vertices are the ROIs and the edges represent connections between them. Edges' weights are the number of white matter fibers. Given the limited knowledge-domain we possess, especially about variability of the data, we follow the approach used in [7] and focus the analysis on the dichotomized version of  $\mathbf{N}$ . This new data matrix  $\mathbf{Z}$  has the same dimension as  $\mathbf{N}$  but each element is defined as:

$$Z_{i,j} = \begin{cases} 1 & \text{if } N_{i,j} > 0 \text{ i.e. if at least one white matter fiber connects } i \text{ and } j; \\ 0 & \text{if } N_{i,j} = 0 \text{ i.e. if no white matter fiber connects } i \text{ and } j. \end{cases}$$

Thus, we are here disregarding the weight of the edges: in the model we will treat equally pairs of brain regions connected by only one white matter fiber and pairs connected by many (possibly thousands of) white matter fibers. This is a strong assumption and we plan to relax it in future developments. Following the idea of [15], for each  $i = 1, \dots, n$  and  $j \neq i$  we model the probability of observing an edge between  $(i, j)$  as

$$P(Z_{i,j} | \mathbf{x}_i, \mathbf{x}_j, \beta) = \text{Bern}\left(Z_{i,j}; \pi_{i,j} = \frac{\exp\{\beta - ||\mathbf{x}_i - \mathbf{x}_j||\}}{1 + \exp\{\beta - ||\mathbf{x}_i - \mathbf{x}_j||\}}\right) \quad (1)$$

where  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are  $p$ -dimensional vectors representing the positions of ROI  $i$  and ROI  $j$ , respectively, in a  $p$ -dimensional latent space. The term  $||\mathbf{x}_i - \mathbf{x}_j||$  represents the  $\ell^2$  distance between pair  $(i, j)$  of ROIs, and encodes the idea that the probability  $\pi_{i,j}$  of observing a connection between regions  $i$  and  $j$  depends on how close/far they are in the  $p$ -dimensional latent space. The choice of  $p$  is a research direction by itself. A few proposals have been made in the literature to tackle this issue, for instance [8]. For simplicity and ease of visualization, we decide to work with  $p = 2$  throughout the rest of the manuscript. The coefficient  $\beta$  controls the highest possible edge probability value in the network. Indeed, if a couple of nodes  $(i, j)$  are at zero distance in the latent space, the probability of them being connected would be  $\exp\{\beta\}/(1 + \exp\{\beta\})$ .

Although we do possess spatial coordinates of the ROIs, we believe that computing Euclidean distance directly on those covariates can mask (if not hinder) the effective spatial dependency in the data. The argument is that brain surface is not flat, nor regular, and a simple distance in the physical space between the centroids of the ROI is not representative of how structural proximity is reflected into the observed white matter fibers counts. To circumvent this limitation, we decide to let the data organize themselves inside a latent space, and we use this projected spatial configuration as our proxy for spatial dependency patterns. We collect all latent positions  $\{\mathbf{x}_i\}$  as row vectors of an  $n \times p$  matrix  $\mathbf{X}$ , with  $p = 2$  as previously mentioned. Given that no self-loops  $(i, i)$  are allowed, the diagonal elements of  $\mathbf{Z}$  are not considered; moreover, symmetry of  $\mathbf{Z}$  makes it possible to focus on either the lower- or upper- triangular part of the matrix. As a result, the likelihood can be expressed as:

$$\mathcal{L}_{\mathbf{Z}}(\boldsymbol{\theta}, \mathbf{X}) = P(\mathbf{Z} | \mathbf{X}, \boldsymbol{\theta}) = \prod_{j=1}^n \prod_{i < j} P(Z_{i,j} | \mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}), \quad (2)$$

where we assume independence between ROIs conditional on the set of parameters  $(\boldsymbol{\theta}, \mathbf{X})$ , which contains also  $\beta$ . Notice that the presence of isolated nodes in the data, that is, regions that are not linked by white fiber matter to any other region, would correspond to an unbounded likelihood. Indeed, the distance in the latent space between the isolated region and the others could potentially be infinite. However, isolated regions are not present in the data. The assumed prior probability distribution for  $\mathbf{X}$  is a Multivariate Normal Distribution for each  $\mathbf{x}_i$ , as in [15]. Indeed, we assume

for each unit  $i = 1, \dots, n$ , that  $\mathbf{x}_i \sim N(\mathbf{x}_i; \mathbf{0}, \Sigma)$ , where the variance-covariance matrix  $\Sigma$  is equal to  $\sigma^2 I_p$ , with  $\sigma^2$  measuring the dispersion of the latent positions in the 2-dimensional space. The corresponding hierarchical formulation is then

$$\begin{aligned} P(\sigma^2 | \boldsymbol{\theta}_{\sigma^2}) &= \text{IG}(\sigma^2; \boldsymbol{\theta}_{\sigma^2}), \\ P(\mathbf{x}_i | \sigma^2) &= N(\mathbf{x}_i; \mathbf{0}, \sigma^2 I_p) \quad \forall i = 1, \dots, n, \\ P(\beta | \boldsymbol{\theta}_\beta) &= N(\beta; \boldsymbol{\theta}_\beta), \\ P(Z_{i,j} | \mathbf{x}_i, \mathbf{x}_j, \beta) &= \text{Bern}(Z_{i,j}; \pi_{i,j}) \quad \forall i, j, \end{aligned}$$

denoting with: IG the Inverse-Gamma density function;  $\boldsymbol{\theta}_{\sigma^2}$  and  $\boldsymbol{\theta}_\beta$  the vectors of hyper-parameters;  $\pi_{i,j}$  the probability of a link being present between pair  $(i, j)$ , as defined in Eq. 1. The associated likelihood  $\mathcal{L}_Z(\boldsymbol{\theta}, \mathbf{X})$  is the one in Eq. 2 with  $\boldsymbol{\theta} = (\boldsymbol{\theta}_\beta, \boldsymbol{\theta}_{\sigma^2})$ . The latent positions are assumed to be a priori independent, conditionally on  $\sigma^2$ , that is,  $P(\mathbf{X} | \sigma^2) = \prod_{i=1}^n P(\mathbf{x}_i | \sigma^2)$ .

The model, which we name Model I, can be fitted with the `latentnet` R package [16], which provides, the currently adopted, default options to set the hyperparameters.

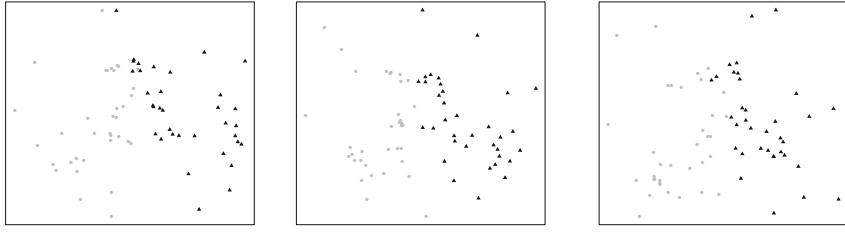
### 3.1 Results on the DTI Dataset

We model the presence of white matter fibers connecting the ROIs according to the approach described in Sect. 3, and estimate their latent positions via posterior means for each of the 20 subjects.

In Fig. 5 we plot the estimated latent coordinates for subjects 8, 9 and 24. Subjects 9 and 24 are non-diagnosed with mental disorder, and of age 21 and 36 respectively. Subject 8 is 21 years old and is diagnosed with major depressive disorder (single episode), cannabis abuse, and eating disorder. Both subjects 8 and 24 are right handed, while 9 is left handed. The symbols in the plots refer to the belonging hemisphere of each region: black triangles indicate the region belongs to the right hemisphere, while grey dots are for regions belonging to the left hemisphere. It is immediate to notice that the use of a Gaussian distribution for the latent coordinates produces a space which clearly separates left and right hemispheres, and this partition is recovered for all the subjects; also, the shape of the coordinates resembles an horizontal projection of the brain on a 2D surface. Although some of the ROIs' estimated positions do not match with the classical Euclidean projection, our findings are mostly coherent with the physical observed topology of brain regions, as the estimated latent coordinates resemble those of the Desikan atlas. In order to quantify the latter finding, we inspected the Procrustes correlation<sup>2</sup> between the estimated latent spaces and

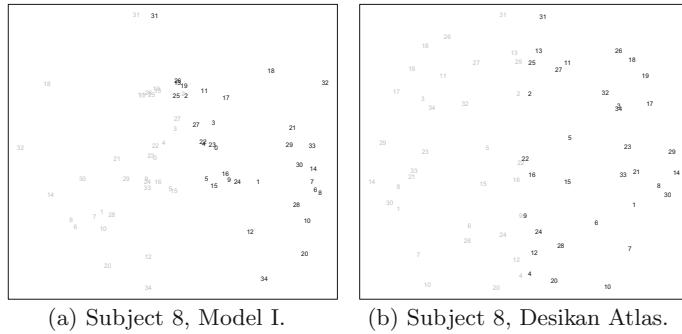
---

<sup>2</sup>Procrustes correlation,  $\rho(S_1, S_2)$ , is a measure of similarity among two spaces,  $S_1, S_2$ . In particular, it measures up to which degree space  $S_2$  was generated by a transformation (rotation, translation or scaling) of space  $S_1$ . It is bounded in  $[0, 1]$ .



(a) Subject 8, Model I. (b) Subject 9, Model I. (c) Subject 24, Model I.

**Fig. 5** Estimated latent positions for the brain regions of subjects 8, 9 and 24. Grey dots indicate left hemisphere regions; black triangles right hemisphere regions. The latent spaces for subjects 9 and 24 are rotated to match the estimated latent space of subject 8, to allow comparison of the different representations. The plots are all drawn on the same scale



(a) Subject 8, Model I.

(b) Subject 8, Desikan Atlas.

**Fig. 6** Brain regions coordinates according to latent space Model I and the Desikan atlas, for subject 8. The Desikan regions are coded according to the numbering in Table 6, reported in the Appendix

the atlas representation. We found an average value of 0.9, and the representations recovered for the different subjects are quite similar: pairs of estimated latent spaces exhibit a 0.95 Procrustes correlation on average. The high values of the correlation confirm the estimated latent structure to be similar between all the subjects in the study, and show that this latent space model provides a parsimonious representation of brain regions, as different subjects can be described by a similar space, a space highly correlated with the physical one but not identical. This suggests that a latent space representation could be meaningfully used to jointly describe data coming from different subjects. Figure 6 provides comparison of the estimated latent space for subject 8 and the general Desikan atlas representation of brain regions.

The estimated posterior means for the  $\beta$  coefficients of subjects 8, 9 and 24 are reported in Table 2, together with the standard deviations and the lower ( $q_{2.5}$ ) and upper bounds ( $q_{97.5}$ ) of the corresponding 95% credible intervals. In Appendix 7 we report Markov chain Monte Carlo trace plots and histograms of the estimated posterior distributions of the intercept parameters, for subjects 8, 9, and 24. The trace plots indicate good mixing of the MCMC sampler.

**Table 2** Estimated posterior summaries of the  $\beta$  coefficients for subjects 8, 9, and 24

Subject	$q_{2.5}$	$q_{97.5}$	Posterior mean	Posterior s.d.
8	7.79	8.98	8.36	0.30
9	6.29	7.15	6.71	0.22
24	6.53	7.54	7.03	0.25

## 4 Time-Varying Dynamic Bayesian Networks for the fMRI Dataset

The second batch of data is a collection of  $n$  time series  $\mathbf{y}_i = (y_i^1, \dots, y_i^t, \dots, y_i^T)$ ,  $i = 1, \dots, n$ , each of length  $T = 404$ , representing the BOLD technique signal recovered through resting state fMRI scans. The vector  $\mathbf{y}^t = (y_1^t, \dots, y_i^t, \dots, y_n^t)'$  is thus a snapshot of the activity levels for all  $n$  regions at time  $t$ . In the spirit of [28], we model the data with a time-varying dynamic Bayesian network (BN). First of all, we assume a first-order Markovian property that allows us to break down, with respect to time  $t$ , the joint probability of vectors  $\{\mathbf{y}^t\}_{t=1, \dots, T}$  into a product of conditional terms referring to each  $\mathbf{y}^t$ . Also, given the parameters of the model, we assume the  $n$  units to be independent from one another. We thus have:

$$P(\mathbf{y}^1, \dots, \mathbf{y}^t, \dots, \mathbf{y}^T | \Omega) = P(\mathbf{y}^1 | \Omega) \prod_{t=2}^T \prod_{i=1}^n P(y_i^t | \mathbf{y}^{t-1}, \Omega) \quad (3)$$

where the time-varying transition (conditional) probability  $P(\mathbf{y}^t | \mathbf{y}^{t-1}, \Omega)$  follows a time-varying linear dynamic model, where  $\Omega$  collects all the parameters of interest. Notice that the dependency on  $t$  is omitted in the notation. More formally:

$$\mathbf{y}^t = \boldsymbol{\Gamma}^{(t)} \mathbf{y}^{t-1} + \boldsymbol{\epsilon} \quad (4)$$

where:  $\boldsymbol{\epsilon}$  is a vector of idiosyncratic errors distributed as  $N(\mathbf{0}, \varsigma^2 I_n)$ , with  $\varsigma^2$  measuring the global noise level of the observed time series;  $\boldsymbol{\Gamma}^{(t)} = \{\gamma_i^t\}_{i=1}^n$  is an  $n \times n$  matrix of real-valued regression coefficients, where for each  $i, t$ ,  $\boldsymbol{\gamma}_i^t = (\gamma_{i,1}^t, \dots, \gamma_{i,j}^t, \dots, \gamma_{i,n}^t)$  is the vector of coefficients explaining the effect of all the units from the previous snapshot  $(y_1^{t-1}, \dots, y_n^{t-1})'$  on the current value of  $y_i^t$ . This means that, if  $\gamma_{i,j}^t \neq 0$  then  $y_j^{t-1}$  regulates  $y_i^t$ , with an edge being therefore present in the associated dynamic network at time  $t$ . In the original work, the authors cast the inferential problem in a penalized weighted regression framework. Following a frequentist approach, in order to estimate the coefficients  $\hat{\boldsymbol{\gamma}}_i^{t^*}$ , for each  $t^* = 1, \dots, T$ , [28] maximize the likelihood obtained from Eqs. 3 and 4, providing the following objective function to be optimized:

$$\hat{\gamma}_i^{t^*} = \operatorname{argmin} \left\{ \frac{1}{T} \sum_{t=1}^T w^{t^*}(t) [y_i^t - \gamma_i^{t^*} \mathbf{y}^{t-1}]^2 + \lambda \|\gamma_i^{t^*}\|_1 \right\}, \quad (5)$$

where  $\lambda$  is the penalization parameter. In Eq. 5, the weights of the weighted linear regression can be readily computed, as they are defined to be

$$w^{t^*}(t) = \frac{K_h(t - t^*)}{\sum_{t=1}^T K_h(t - t^*)},$$

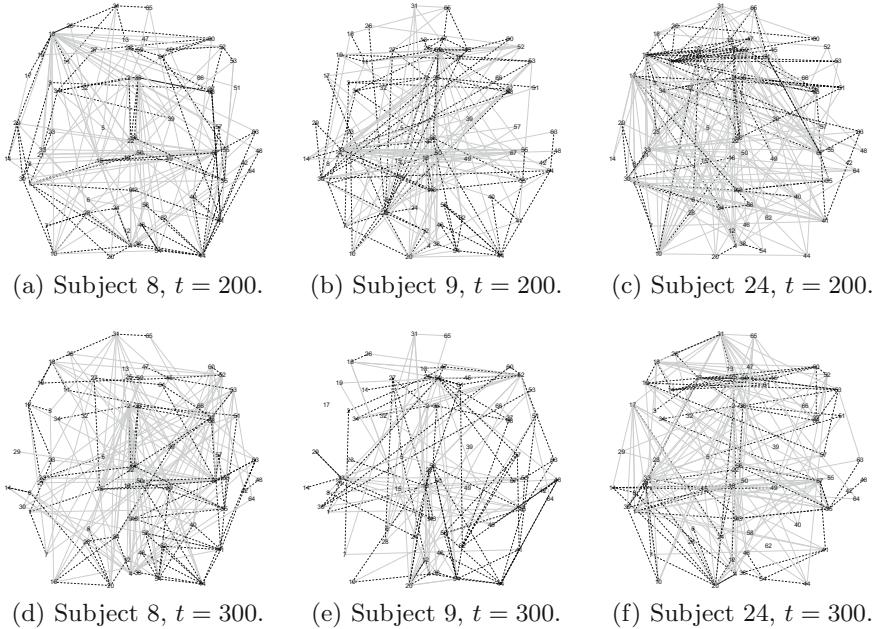
with the Gaussian kernel  $K_h(t - t^*) = \exp(-h^{-1}(t - t^*)^2)$ , where  $h$  is a tuning parameter (for a suggested criterion choice, we refer to [28]). We implemented the algorithm outlined in the original paper and some preliminary results on the data are presented in Sect. 4.1.

## 4.1 Results on the **fMRI** Dataset

We model brain activity via its BOLD proxy as described in Sect. 4. For each time point and each subject, we constructed a BN whose nodes are the ROIs. In this representation, an edge connecting node  $i$  to node  $j$  at time  $t$ , means that the activation of node  $i$  at time  $t - 1$  leads to the activation of node  $j$  at time  $t$ . The weights of each edge are given by the values of the coefficients  $\hat{\gamma}_{ij}^{t^*}$ , appearing in Eq. (5).

For each subject, we considered both scans, and constructed two BNs: as a matter of fact, the scans are repeated measures on the same subject and can thus be used to validate the estimates of the coefficients  $\hat{\gamma}_i^{t^*}$ . Looking at the time dynamic, the more often two regions are connected by an edge, the more we could expect the two to interact. The direction of the edge also indicates the causality (in a Granger sense) in the activation process. We consider the activation of region  $j$  by region  $i$  to be relevant only if the associated weights  $\hat{\gamma}_{ij}^{t^*}$  are positive in both the scan and re-scan BNs, and draw an edge whenever this happens. Applying this rule of thumb, we find that the BNs of each subject shared roughly 30% of their positive coefficients.

Figure 7 reports the dynamic BNs at times  $t = 200$  and  $t = 300$ , for subjects 8, 9, and 24. The edges are drawn according to the concordance between white matter fiber presence and high value of the activation coefficient among two regions, with dashed lines indicating concordance and continuous lines discordance. The oldest subject (labeled 24) had the largest number of connections per time, on average: 250, approximately 9% more than the other two younger participants (labeled 8 and 9). Discordant edges have been found to be more frequent than concordant ones in subjects 8 and 24, indicating the prevalence of interaction between regions not connected by white matter fibers. This tendency is reversed for subject 9, that has a larger count of concordant edges than of discordant ones. However, if these counts are compared to the overall observed number of white matter fibers, the BNs for all the subjects exhibit a higher relative frequency of concordant lines, hence of estimated

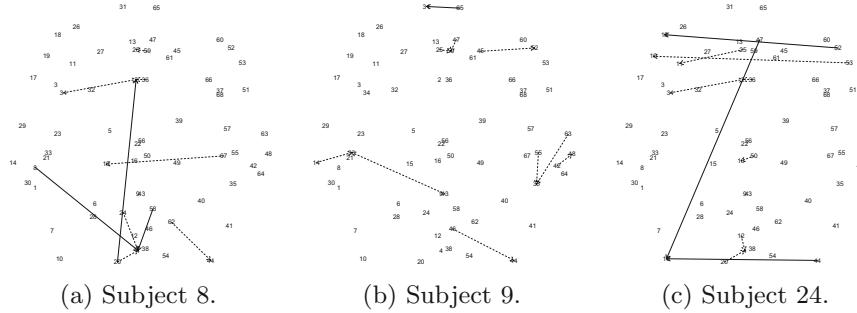


**Fig. 7** Frame of Dynamic Bayesian Networks at time  $t = 200$  and  $t = 300$ , for subjects 8, 9 and 24. The edges are drawn according to the concordance between white fiber matter presence and high value of the activation coefficient among two regions. Black (dashed line) indicates concordance, grey (continuous line) discordance

connections between regions that are already joined by white matter fibers. Figure 8 shows the ten most frequent connections for the same three subjects (8, 9, and 24). For these three individuals, connections within the same hemisphere were more frequent. Subject 8 and 24 share three of their ten most frequent connections:

- from the left insula (34) to the left caudal anterior cingulate (2): the first region is believed to play a role in consciousness and in functions linked to body's homeostasis and emotion [32]; the latter helps regulating blood pressure and heart rate [3], and it is also believed to be involved in the decision making process [1];
- from the left pericalcarine (20) to the left cuneus (4): cuneus' function is linked to visual processing [5], while the activating region (20) is where the primary visual cortex is concentrated;
- from the left caudal anterior cingulate (2) to the right caudal anterior cingulate (36).

In Appendix 6 we provide the names of all the Desikan regions, corresponding to the numbers in the plots. Figures 7 and 8 refer to the Desikan atlas representation of the brain. As already mentioned, only 68 of the 70 regions in the data are plotted and the two regions labeled as “unknown left” and “unknown right” are removed. This was done without loss of information in Fig. 8, because the strongest relations never involved such “unknown” regions.



**Fig. 8** Ten most frequent connections in the dynamic BNs for subjects 8, 9 and 24. The edges are drawn according to the concordance between white fiber matter presence and high value of the activation coefficient among two regions. Dashed lines indicate concordance, continuous discordance. The arrows point to the activated regions

For the benefit of science reproducibility, the R Script used for the analysis are available at <http://www.github.com/silviadangelo>.

## 5 Discussion

In this work, we analyzed both fMRI and DTI data, aiming at combining information from these sources, so as to learn possible patterns of dependencies between brain regions.

First, using the structural connectivity information of the DTI data, we estimated the latent positions of the brain regions, for each subject under study, with a latent space approach. Comparing the estimated latent space positions with the reference Desikan atlas, we found that—consistently for all subjects—the inferred latent space is highly correlated with the physical one, even if not completely identical. We are aware that the proposed latent space model is simplistic, being based on a dichotomized version of the data at hand: in fact, we only took into account whether a pair of regions was connected or not, loosing information regarding the strength of the physical connection. In future developments, we thus aim at considering this additional information, which amounts to modeling the entries of the count matrix. We also aim at generalizing the model to deal with over-dispersion and zero inflation of the data, thus moving from the Bernoulli as the distribution for the data to, for instance, a zero-inflated Negative Binomial distribution. A final direction we are interested in exploring is to incorporate into the analysis information regarding the volume of the brain regions. We indeed have the intuition that larger regions might inherently have a greater number of white matter fibers. The most immediate way to integrate the volume information would be to provide it to the model as a covariate or an offset term.

In a second moment, we studied Granger causality in the fMRI data, in order to learn patterns of activations between ROIs. In particular, we modeled the data of each subject separately with a dynamic Bayesian network approach as prescribed in [28], and fit it following a frequentist approach. The main finding of this analysis is that the presence of an edge in the inferred BN, measuring activation between pairs of regions, was tied to the physical connectivity measured by the DTI dataset. Regarding future developments for this second model, we plan to fit it following the Bayesian paradigm. The Bayesian approach would help in assessing coherently the uncertainty about the quantities of interest, and also aid in incorporating prior information into the statistical analysis.

Bridging the two models together would be an interesting and powerful way to combine all the information at our disposal in a single framework. One possible direction, in this sense, could be to use the information on white matter fibers, coming from the DTI dataset, as prior information for developing a more complex hierarchical structure to model the functional activity collected in the fMRI dataset: for example, additional latent layers, or inferred latent distances to be used as additional covariates of the weighted penalized regression. Another option would be to employ the latent space positions inferred from the DTI dataset as quantities to plug-in into the computation of the weights  $\{w^*(t)\}$  of the weighted penalized regression we perform the fMRI dataset. This approach would indeed smooth observations both across time and space, increasing the borrowing of information. Finally, we briefly report some considerations elaborated during and after this work. The first one regards the huge preprocessing the data underwent before being considered for analysis. For example, the data may have been corrected for slice timing, realignment, co-registration of structural and functional images, normalization and smoothing. So far, we did not have the opportunity to go deeper into this aspect of the data. We believe, however, this data preprocessing information to be potentially relevant in guiding us to choose a more appropriate modeling framework, since the preprocessing itself may have resulted in some artifacts, captured by the models we used, and/or altered the very same structure of the phenomenon.

On a related note, the second consideration involves the additional information a knowledge-domain expert would bring to the discussion table, allowing us to enhance the quality of the analysis and give insights about the obtained results. This is especially true when considering delicate modeling choices, as for example the threshold on the number of fibers we adopted for the first dataset. Also, on this matter, the potential impact of other covariates could prove to be an interesting direction for comparing groups of subjects: for example, gender has been shown in the literature to be correlated with white matter fibers' counts [19].

A last consideration concerns the fMRI data: given the nature of 'resting state' observations, information was only inspected from a correlation point of view because, without stimuli, the signal in the time-series appeared to be swamped by noise. An alternative, enriched, way to look at the data would involve fMRI recordings of subjects asked to perform different tasks, in order to effectively provide occasions to observe noticeable changes in the signal.

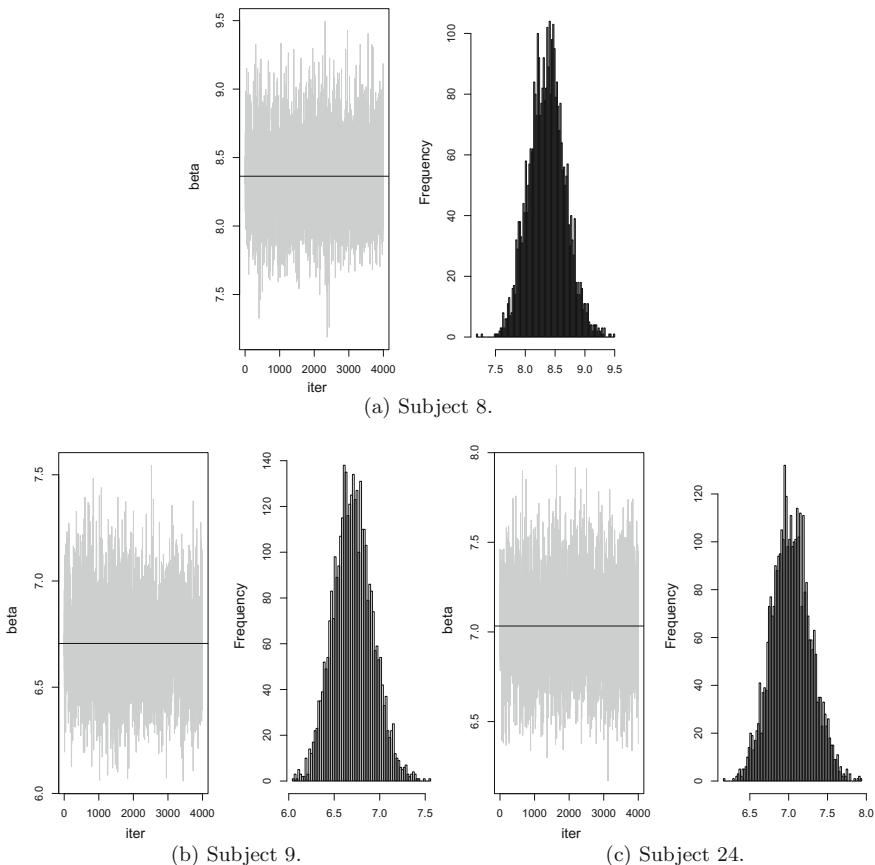
**Acknowledgements** Data were provided by Greg Kiar and Eric Bridgeford from NeuroData at Johns Hopkins University, who graciously preprocessed the raw DTI and R-fMRI imaging data available at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html). We would like to thank Ritabrata Dutta for initial discussions during ‘StartUp Research’ and for comments to the final version of the manuscript. Also, we would like to thank the organizers of ‘StartUp Research’ event, [www.congressi.unisi.it/startupresearch/](http://www.congressi.unisi.it/startupresearch/), for creating the opportunity for this research contribution and the other working groups present at the event for fruitful discussions.

## 6 A. Desikan Atlas Codes

Code	Left/Right	Code	Left/Right	Region name
1	L	35	R	bank of the superior temporal sulcus
2	L	36	R	caudal anterior cingulate
3	L	37	R	caudal middle frontal gyrus
4	L	38	R	cuneus
5	L	39	R	entorhinal
6	L	40	R	fusiform
7	L	41	R	inferior parietal lobule
8	L	42	R	inferior temporal gyrus
9	L	43	R	isthmus cingulate cortex
10	L	44	R	lateral occipital gyrus
11	L	45	R	lateral orbitofrontal
12	L	46	R	lingual
13	L	47	R	medial orbitofrontal
14	L	48	R	middle temporal gyrus
15	L	49	R	parahippocampal
16	L	50	R	paracentral
17	L	51	R	pars opercularis
18	L	52	R	pars orbitalis
19	L	53	R	pars triangularis
20	L	54	R	pericalcarine
21	L	55	R	postcentral
22	L	56	R	posterior cingulate cortex
23	L	57	R	precentral
24	L	58	R	precuneus
25	L	59	R	rostral anterior cingulate cortex
26	L	60	R	rostral middle frontal gyrus
27	L	61	R	superior frontal gyrus
28	L	62	R	superior parietal lobule
29	L	63	R	superior temporal gyrus
30	L	64	R	supramarginal gyrus
31	L	65	R	frontal pole
32	L	66	R	temporal pole
33	L	67	R	transverse temporal
34	L	68	R	insula

## 7 B. MCMC Diagnostics of Intercept Parameters of the Latent Space Model

See Fig. 9.



**Fig. 9** Trace plots and histograms of the posterior distributions of the intercept parameters  $\beta$  in the latent space model, for subjects 8, 9 and 24

## References

1. Allman, J.M., Hakeem, A., Erwin, J.M., Nimchinsky, E., Hof, P.: The anterior cingulate cortex. *Ann. N. Y. Acad. Sci.* **935**(1), 107–117 (2001)
2. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Rev. Neurosci.* **10**(3), 186–198 (2009)

3. Cersosimo, M.G., Benarroch, E.E.: Chapter 5—central control of autonomic function and involvement in neurodegenerative disorders. *Handbook Clin. Neurol.* **117**, 45–57 (2013)
4. Craddock, R.C., Jbabdi, S., Yan, C.G., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C., Heberlein, K., Colecombe, S., Milham, M.P.: Imaging human connectomes at the macroscale. *Nature Methods* **10**(6), 524–539 (2013)
5. Crockford, D.N., Goodyear, B., Edwards, J., Quickfall, J., el Guebaly, N.: Cue-induced brain activity in pathological gamblers. *Biological Psychiatry* **58**(10), 787–795 (2005)
6. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**(3), 968–980 (2006)
7. Durante, D., Dunson, D.B.: Bayesian inference and testing of group differences in brain networks. *Bayesian Anal.* **13**, 29–58 (2018)
8. Durante, D., Dunson, D.B., Vogelstein, J.T.: Nonparametric bayes modeling of populations of networks. *J. Am. Statist. Assoc.* **112**(520), 1516–1530 (2017)
9. Granger, C.W.: Investigating causal relations by econometric models and cross-spectral methods. *Econom. J. Econom. Soci.*, 424–438 (1969)
10. Granger, C.W.: Testing for causality: a personal viewpoint. *J. Econ. Dyn. Control* **2**, 329–352 (1980)
11. Hall, E.C., Raskutti, G., Willett, R.: Inference of high-dimensional autoregressive generalized linear models (2016). [arXiv:1605.02693](https://arxiv.org/abs/1605.02693)
12. Han, F., Lu, H., Liu, H.: A direct estimation of high dimensional stationary vector autoregressions. *J. Mach. Learn. Res.* **16**, 3115–3150 (2015)
13. Handcock, M.S., Raftery, A.E., Tantrum, J.M.: Model-based clustering for social networks. *J. Royal Statist. Soc. Ser. A* **170**(2), 1–22 (2007)
14. Hoff, P.: Bilinear mixed-effects models for dyadic data. *J. Am. Statist. Assoc.* **100**(469), 286–295 (2005)
15. Hoff, P.D., Raftery, A.E., Handcock, M.S.: Latent space approaches to social network analysis. *J. Am. Statist. Assoc.* **97**(460), 1090–1098 (2002)
16. Krivitsky, P.N., Handcock, M.S.: Fitting position latent cluster models for social networks with latentnet. *J. Statist. Softw.* **24**(5) (2008)
17. Liu, Y., Niculescu-Mizil, A., Lozano, A.C., Lu, Y.: Learning temporal causal graphs for relational time-series analysis. In: Proceedings of the 27th International Conference on Machine Learning (ICML2010), pp. 687–694 (2010)
18. Ma, L., Hasan, K.M., Steinberg, J.L., Narayana, P.A., Lane, S.D., Zuniga, E.A., Kramer, L.A., Moeller, F.G.: Diffusion tensor imaging in cocaine dependence: regional effects of cocaine on corpus callosum and effect of cocaine administration route. *Drug Alcohol Depend.* **104**(3), 262–267 (2009)
19. Menzler, K., Belke, M., Wehrmann, E., Krakow, K., Lengler, U., Jansen, A., Hamer, H., Oertel, W., Rosenow, F., Knake, S.: Men and women are different: diffusion tensor imaging reveals sexual dimorphism in the microstructure of the thalamus, corpus callosum and cingulum. *Neuroimage* **54**(4), 2557–2562 (2011)
20. Monnig, M.A., Caprihan, A., Yeo, R.A., Gasparovic, C., Ruhl, D.A., Lysne, P., Bogenschutz, M.P., Hutchison, K.E., Thoma, R.J.: Diffusion tensor imaging of white matter networks in individuals with current and remitted alcohol use disorders and comorbid conditions. *Psychol. Addict. Behav.* **27**(2), 455 (2013)
21. Nowicki, K., Snijders, T.A.B.: Estimation and prediction of stochastic blockstructures. *J. Am. Statist. Assoc.* **96**(455), 1077–1087 (2001)
22. Ramsey, J.D., Hanson, S.J., Hanson, C., Halchenko, Y.O., Poldrack, R.A., Glymour, C.: Six problems for causal inference from fMRI. *Neuroimage* **49**(2), 1545–1558 (2010)
23. Rosenbloom, M., Sullivan, E.V., Pfefferbaum, A., et al.: Using magnetic resonance imaging and diffusion tensor imaging to assess brain damage in alcoholics. *Alcohol Res. Health* **27**(2), 146–152 (2003)

24. Shojaie, A., Michailidis, G.: Discovering graphical granger causality using the truncating lasso penalty. *Bioinformatics* **26**(18), i517–i523 (2010)
25. Simpson, S.L., Bowman, F.D., Laurienti, P.J.: Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Statist. Surv.* **7**, 1 (2013)
26. Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W.: Network modelling methods for fmri. *Neuroimage* **54**(2), 875–891 (2011)
27. Snijders, T.A.B., Nowicki, K.: Estimation and prediction for stochastic blockmodels for graphs with latent block structure. *J. Classif.* **14**(1), 75–100 (1997)
28. Song, L., Kolar, M., Xing, E.P.: Time-varying dynamic Bayesian networks. In: Advances in neural information processing systems, pp. 1732–1740 (2009)
29. Sporns, O.: Structure and function of complex brain networks. *Dialogues Clin. Neurosci.* **15**(3), 247 (2013)
30. Stam, C.J.: Modern network science of neurological disorders. *Nature Rev. Neurosci.* **15**(10), 683–695 (2014)
31. Tank, A., Fox, E.B., Shojaie, A.: Granger causality networks for categorical time series (2017). [arXiv:1706.02781](https://arxiv.org/abs/1706.02781)
32. Xue, G., Lu, Z., Levin, I.P., Bechara, A.: The impact of prior risk experiences on subsequent risky decision-making: the role of the insula. *Neuroimage* **50**(2), 709–716 (2010)

# Hierarchical Graphical Model for Learning Functional Network Determinants



Emanuele Aliverti, Laura Forastiere, Tullia Padellini,  
Sally Paganin and Ernst Wit

**Abstract** Analysis of brain functionality is a stimulating research topic from both a neuroscientific and statistical perspective. Although several works have improved our comprehension of the relationship between subject-specific information and brain architecture, many questions remain open. The aim of this paper is to relate functional connectivity patterns with subject-specific features and brain constraints, such as age and mental illness of the subject and lobes membership for brain regions, and illustrate whether these phenotypes affect the neurophysiological dynamics. To address such goal we consider a modular approach that allows to remove noise from the fMRI data, estimate the functional dependency structure and relate functional architecture with structural and phenotypical information.

**Keywords** Functional connectivity · Gaussian graphical models  
Hierarchical models · Modular estimation

---

E. Aliverti · S. Paganin  
Department of Statistical Sciences, University of Padova, Padua, Italy  
e-mail: [aliverti@stat.unipd.it](mailto:aliverti@stat.unipd.it)

S. Paganin  
e-mail: [paganin@stat.unipd.it](mailto:paganin@stat.unipd.it)

L. Forastiere  
Department of Statistics, Informatics and Applications, University of Florence, Florence, Italy  
e-mail: [forastiere@disia.unifi.it](mailto:forastiere@disia.unifi.it)

T. Padellini (✉)  
Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy  
e-mail: [tullia.padellini@uniroma1.it](mailto:tullia.padellini@uniroma1.it)

E. Wit  
Faculty of Informatics, University of Italian Switzerland, Lugano, Switzerland  
e-mail: [ernst.jan.camiel.wit@usi.ch](mailto:ernst.jan.camiel.wit@usi.ch)

## 1 Introduction

In recent years, neuroscience has been a great source of inspiration in statistical methodology (e.g., [5, 19, 32]). The reason behind this interest, beside the obvious fascination with the quest for insights on how the brain works, is that neuroimaging modeling is at the crossroad between spatial statistics, time series, network analysis and high dimensional inference, thus allowing for an exciting interplay between different branches of statistics and other sciences. An area that is increasingly growing is the analysis of functional connectivity, which seeks to identify brain areas that behave similarly, potentially despite their spatial proximity or their membership to the same lobes and hemisphere.

The focus of this work is on estimating the relation between phenotypes and anatomical structure with functional brain behavior, employing functional magnetic resonance imaging (fMRI) as a measure of brain activity.

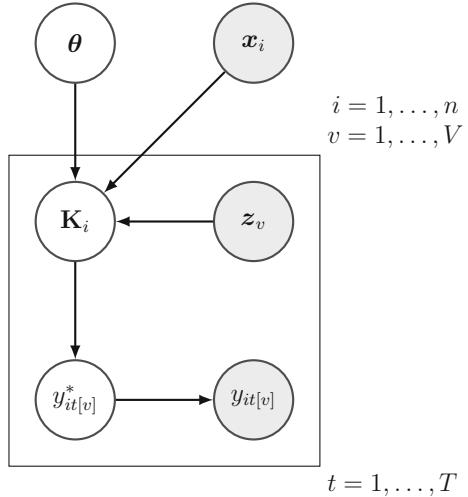
There is a rich literature related to the statistical study of functional connectivity patterns within the brain. Several approaches focus on representing the functional relationship among brain regions by means of a network, whose edges connect areas of the brain sharing similar behaviors in terms of functional properties. Nodes of the network are usually defined as regions of interest (ROIs), typically provided by experts in neuroscience (e.g., [7, 15]). Alternatively, ROIs can be identified with data-driven approaches [12] recovering lower dimensional structures in the high-dimensional fMRI data, such as Principal Component Analysis [3] or Independent Component Analysis.

A common approach to determine the functional edges interconnecting brain regions consists in thresholding the empirical correlations between fMRI series. The functional connectivity among subjects is then analyzed by assessing network properties (e.g. small-world, scale free connectivity) and comparisons are made using network summary statistics; see [10, 25] and references mentioned therein for a general review of these methods. A naive correlation-based approach, however, provides an incomplete representation of the brain's functional connectivity, since it does not take into account covariates and has been shown to produce nonzero estimates for the correlation of independent brain regions [32]. Furthermore, when the number of brain regions is relatively big with respect to the lengths of the fMRI series, the empirical estimator of the correlation matrix may exhibit poor performance, especially if the covariance matrix is close to singularity.

Several alternative approaches have been investigated to obtain more reliable representations and robust descriptions of the functional networks, such as wavelet based correlation analysis [1] and graphical models [14], along with a broad discussion about properties of the resulting networks. Nevertheless, these approaches still fail to acknowledge the impact of covariates, and more in general, little work has been done in assessing the relation between such networks and brain structure or subject-specific covariates.

We address such issue by proposing a sequential hierarchical approach, which estimates the functional connectivity from denoised signals and then relates it to

**Fig. 1** Hierarchical model representing the assumed probabilistic generative mechanism. Observed quantities are colored in light grey, unobservable in white



observed phenotypes. Although we build on hierarchical models in defining the probabilistic representation of the available quantities, we bypass the joint estimation procedure in order to provide a fast exploratory method, able to assess the relationship between phenotypes, brain constraints and neurophysiological dynamics. For the model fitting we adopt a modular strategy that leverages available methods in the literature. The modularization procedure consists of decomposing the hierarchical model in three sub-models: (i) a smoothing procedure to remove noise from the fMRI signal, (ii) a graphical model which encodes the functional brain connectivity and (iii) a regression model investigating the relation between phenotypes and functional connectivity patterns. Our approach retains ease of interpretation while accounting for functional relations across all the subjects; moreover, the robustness of inferential conclusions is assessed by means of a multiscale analysis.

The rest of the paper is organized as follows. In the following Sect. 2, we introduce the notation and define the general hierarchical specification of our modular approach. In Sect. 3 we detail the methods used in each module, along with the application to the data. Finally, Sect. 4 is dedicated to final remarks and our conclusions.

## 2 Hierarchical Model

Our motivating application is drawn from the NKI1 pilot study, part of the “Enhanced Nathan Kline Institute-Rockland Sample project” conducted over 24 healthy subjects; the dataset used in this application was kindly provided by Greg Kiar and Eric Bridgeford (NeuroData—Johns Hopkins University). The resting state fMRI raw measurement have been preprocessed using the ndmg pipeline [24] and the C-PAC software; for additional details on this procedure, see

<https://fcp-indi.github.io/>. Two subjects were removed from the analysis due to missing data in several features, and the final sample size for this application is equal to  $n = 22$  subjects.

For each subject  $i = 1, \dots, 22$ , fMRI signals referred to  $v = 1, \dots, 70$  regions of interest (ROI) of the brain were collected at  $t = 1, \dots, 404$  equally spaced times, with a time span between measurements of 1400 ms. Let  $Y_{it} = (y_{it[1]}, \dots, y_{it[70]})$  denote the vector of length 70 encoding the fMRI measurement for subject  $i$  at time  $t$ , for all the ROIs considered jointly, with generic element  $y_{it[v]}$  referred to the  $v$ -th ROI. Along with fMRI data, some additional features are available for every subject, such as age, mental status and handedness, which comprise the vector  $\mathbf{x}_i$  for each  $i = 1, \dots, 22$ . Some features related to the brain architecture, such as the lobe membership of each ROI, are also provided; these covariates are denoted as  $\mathbf{z}_v$ , for  $v = 1, \dots, 70$ . Although each subject was scanned twice, we decided not to use data from the second scan, as it was not available for every subjects.

In order to study the presence and the type of relation between the measured brain signals and the available features, we consider a global generative mechanism for the observed quantities, summarized in Fig. 1. We assume that the fMRI data stems from a generative process in which subject-specific features and brain anatomy affect the functional brain behavior, and such characteristics are associated with a set of parameters  $\boldsymbol{\theta} = \{\theta_x, \theta_z\}$  with elements referring respectively to the observed subject-specific features and ROI-specific properties. Furthermore, we suppose that the observed covariates affect the dependence structure among the functional time series, which we characterize by a graphical model or, equivalently, by its associated adjacency matrix  $\mathbf{K}_i$ . In the neuroscientific literature,  $\mathbf{K}_i$  covers a central role, since it characterizes the functional network among brain regions (e.g., [10]). In our specific setting, each node of the functional network—or, equivalently, each row and column of the associated adjacency matrix—represents one of the 70 regions of interest. The edges summarize dependence among ROIs in a functional perspective; if two nodes are connected, the corresponding brain regions will mutually influence their functional activity, resulting in cross-correlated measurements of the clean signal, that we denote with  $Y_{it}^*$ . If we suppose that the true signal can be accurately identified removing accidental noise from the observed data  $Y_{it}$ , the crucial aim of this application is to estimate properly the set of parameters  $\boldsymbol{\theta}$ , since those quantities measure the effect of phenotypical variation on the neurophysiological dynamics.

A joint model specification for all the quantities involved in Fig. 1 might be fairly complicated, since it requires the specification of a joint likelihood for the observed series  $Y_{it}$  as a function of all the unknown quantities and observed covariates; the inclusion of subject-specific information within the estimation of the dependency structure of the functional network is particularly challenging. The same conclusion holds for a potential joint estimation of the cross-sectional dependencies among the signal. In this application, we will consider a modular approach for estimating the model in Fig. 1, in order to provide preliminary insights about the phenotypical effect on brain functional dynamics, and potentially guide further investigations.

The statistical model in Fig. 1 can be decomposed in stages or “modules”, with each component specifying a single model for one or more variables at time. For

every module, several strategies of analysis are feasible, each of which has been extensively investigated and employed in the neuroscientific literature. We will consider then a separate approach in the estimation process, fitting each module and plugging-in the results from the previous step in the subsequent procedure. This plug-in approach, often called modularization [28] or two-step estimation [30], allows to build a complete model by combining different methods sequentially, with the output of a former stage used as input for the latter. Notable examples of application of modular approaches can be found in causal inference area with propensity score [31], pharmacology [6] and meta-analysis [27].

### 3 Modular Estimation Using Connectome Data

Modularization leads to two noticeable advantages in the estimation process. The first one is computational: since blocks are estimated disjointly, the parameter space to be explored in every module is small, and thus we can rely on relatively quickly estimation routines. This also allows for the possibility to conduct analysis under different settings in order to validate robustness of the results. The second benefit is that modularization reduces the effect of model misspecification, since fitting each step separately mitigates the propagation of error among consecutive steps and, potentially, reduces the impact of severe errors.

Our approach is particularly general and enables the inclusion of several techniques within each separate module; in the following we describe in details the modeling strategies adopted in every step along with their application to the data under investigation. For the ease of illustration, the hierarchical model in Fig. 1 was discussed from top to bottom, i.e. starting from what inference will focus on and describing how those quantities relate to the observed data; estimation, instead, will proceed in the opposite direction, using observed raw data as input to make inference on the parameters of interest.

#### 3.1 Denoising

We firstly focus on obtaining the signal component from the observed time series data. Despite the elaborate preprocessing procedures, neuroimaging data are typically corrupted by noise that masks the true signal; especially with fMRI data, it is common to filter them before the analysis to increase the signal to noise ratio and hence the reliability of the results. Recall that  $Y_{it}$ ,  $t = 1, \dots, 404$ , denotes the multivariate time series referred to the  $i$ -th subject for  $i = 1, \dots, 22$ , encoding the fMRI signal recorded over time. It is reasonable to assume that the path of the series over time domain is contaminated by some additive random noise that masks the original properties of the series itself; hence we assume that, at each time  $t$ , the observed fMRI signal for the  $i$ -th subject can be decomposed as

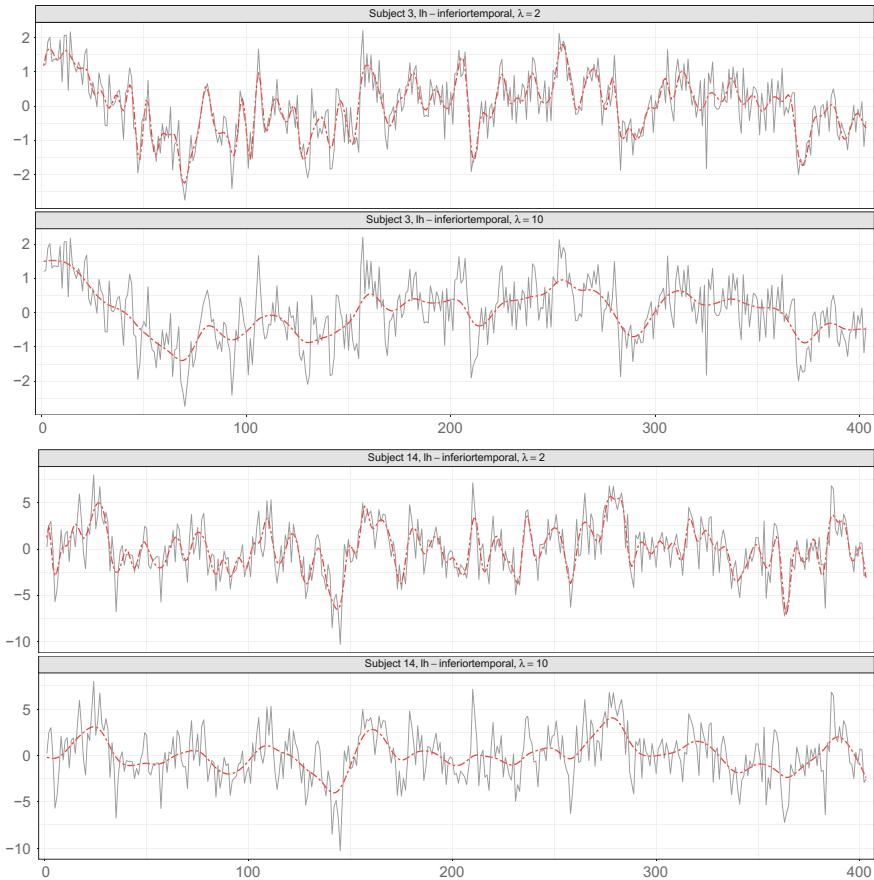
$$Y_{it} = Y_{it}^* + \varepsilon_{it} \quad i = 1, \dots, 22 \quad t = 1, \dots, 404 \quad (1)$$

where  $Y_{it}^*$  is the clean signal and  $\varepsilon_{it}$  represents the noise component. Noise correction is a crucial step of mapping resting state signal fluctuations, however which method is the most appropriate to remove noise from such signal is still an open question, since it is not clear what the “ground truth” signal consists of when the subject is not focused on well identified activities [8]. Several methods can be employed to perform this denoising, for example smoothing splines or total variation (e.g., [16, Chapter 6]). We opt for a smoothing approach to denoising, and to estimate the clean signal  $Y_{it}^*$ , as denoted in Eq. (1), by means of smoothing splines (e.g., [4]). Let  $y_{it[v]}$  denote the univariate time series for ROI  $v$  in subject  $i$ , with  $v = 1, \dots, 70$  and  $i = 1, \dots, 22$ , let  $y_{it[v]}^*$  denote its smoothed counterpart. The smoothed time series is the solution to the following minimization problem:

$$\underset{y_{it[v]}^*}{\operatorname{argmin}} \left\{ \sum_{t=1}^T (y_{it[v]} - y_{it[v]}^*)^2 + \lambda \int \left( \frac{\partial^2}{\partial t^2} y_{it[v]}^* \right)^2 dt \right\}, \quad (2)$$

where  $y_{i[v]}^* = (y_{i1[v]}^*, \dots, y_{i404[v]}^*)$ . Smoothing the signal from each ROI separately, we neglect the spatial dimension of the fMRI data; however, since our aim is not focused on modelling the effect of spatial constraints, we did not include such information on purpose. This strategy also avoids the potential issues involved with spatial smoothing, for example changes in the correlation structure of the data and strengthening of spurious spatial dependency [2].

The parameter  $\lambda$  in Eq. 2 controls the trade-off between complexity and goodness-of-fit of the smoothed series, and its choice determines implicitly the amount of noise we wish to remove. Existing methods for selecting the tuning parameters take into account the temporal structure of the data, however they are built for noisier fMRI signals and tend to oversmooth in the case of resting state fMRI [13]. Although it is reasonable to tune this parameter with automated methods such as Generalized Cross Validation, we considered conducting a sensitivity analysis with respect to the choice of this parameter, and evaluate whether inferential conclusions are stable when the smoothed series capture different trends. In Fig. 2 we reported, for two subjects, original and smoothed fMRI data referred to a region in the inferiotemporal lobes of the left hemisphere. Smoothed series are reported with two different levels of smoothing, respectively  $\lambda = 2$  and  $\lambda = 10$ . Figure 2 suggests that when the value of  $\lambda$  is increased, the estimated series become smoother and highlight the large scale variability, while when  $\lambda$  is fixed to a small value the estimated series tend to follow the accidental fluctuation.



**Fig. 2** An example of the original time series  $Y_{it}$  (solid line) and denoised estimates  $Y_{it}^*$  (dashed line), for subjects 3 and 14 with two different levels of the smoothing parameter  $\lambda$

### 3.2 Estimation of the Graphical Model

The dependence structure among the signal measured at different ROI is a key quantity in our model, since it connects the brain constraints and subject-specific features to the observed fMRI series, and describes the synchronization in brain activity for each pair of brain regions in each subject. Neuroscientific literature commonly refers to such structure as functional network, and several methods have been employed to provide a reasonable estimator for such quantity. A typical approach consists in representing functional connectivity by means of graphical models; in particular, Gaussian graphical models are becoming increasingly popular in neuroimaging (e.g., [14]), since they are able to capture conditional dependencies between brain regions with fast estimation routines and robust guarantees [17].

In order to estimate the functional network among brain regions, we first centered each smoothed time series with respect to its empirical mean. Assuming that for the  $i$ -th subject, at each time  $t = 1, \dots, 404$ , we observe a realization of a 70-variate Normal distribution with mean vector zero and precision matrix  $\Omega_i$ , conditional independence can be assessed estimating the precision matrix  $\Omega_i$ . Note that, even if the normality assumption is violated,  $\Omega_i$  still provides a measure of the association between the functional series for the  $i$ -th subject. A popular and reasonable approach to estimate a graphical model induces sparsity in the estimation of the precision matrix  $\Omega_i$  through an  $\ell_1$  penalty, favouring some elements of the estimated matrix to be shrunk toward zero and providing a well defined estimator when the covariance matrix is singular [17].

The problem solves, in its general form,

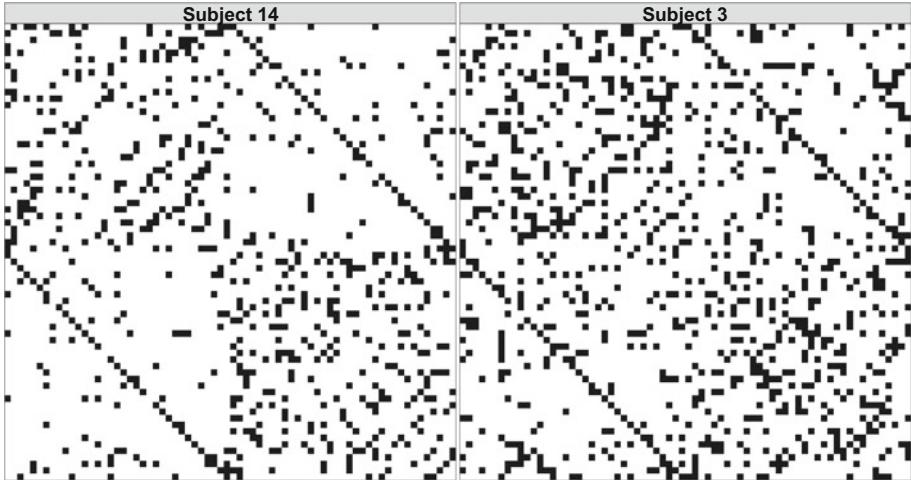
$$\hat{\Omega}_i = \operatorname{argmax}_{\Omega_i \in \mathcal{G}_{70}} \left\{ \log |\Omega_i| - \operatorname{Tr}(\Omega_i^\top S_i) - \xi_i \|\Omega_i\|_1 \right\}, \quad (3)$$

where  $\mathcal{G}_k$  is the manifold of positive defined matrices of dimension  $k$ ,  $S_i$  is the sample covariance matrix,  $\xi_i$  is a penalization parameter,  $|\cdot|$  indicates the matrix determinant while  $\|\cdot\|_1$  the  $\ell_1$ -norm; see [11, 18] for detailed information on this particular optimization problem. Let  $\mathbf{K}_i$  denote the binary version of  $\Omega_i$ , with generic element  $\mathbf{k}_{i[u,v]} = \mathbb{I}(\Omega_{i[u,v]} \neq 0)$ . Every  $\mathbf{K}_i$  can be interpreted as the adjacency matrix of the functional network for subject  $i$ , and the generic element  $\mathbf{k}_{i[u,v]}$  indicates whether, for subject  $i$ , region  $u$  and region  $v$  are connected, for subjects  $i = 1, \dots, n$  and brain regions  $u = 2, \dots, 70$  and  $v = 1, \dots, u-1$ .

The parameters  $\xi_i$  in Eq. 3 control the sparsity of the resulting matrix, and can be selected with several information criteria or stability principles [34]. Since we are assuming that the graphical models stem from the same generative process, we fix the value of  $\xi_i = \xi$  across subjects. Moreover, the choice of the smoothing level in the previous module has an important role in determining the characteristics of the resulting estimated graph, and since we aim to compare inferential conclusions at different level of the smoothed series, we opted for a fixed procedure in the choice of  $\xi$ .

In choosing the global penalization value, however, standard criteria often selected over-sparse solutions. Although extra sparsity does not constitute a serious issue in high-dimensional graphical models, when interest is on describing the functional networks more conservative configuration are preferred [10]. We restricted the range of the penalization parameter  $\xi$  indirectly, by placing constraints on the resulting minimum value of the functional networks density, measured as proportion of non-zero entries of the network's adjacency matrix. Different values for the minimum density were tried, ranging in the interval (0.05–0.20), with resulting estimates robust against different choices of the parameter.

In Fig. 3 we reported the estimated functional network for the same subjects reported in Fig. 2, using  $\lambda = 10$  and with a constraint on the functional networks density to values greater or equal to 0.10. We will use this setting for the remaining of the discussion, unless explicitly specified. Both functional networks report interesting patterns, for example a block structure that recalls hemisphere division. However,



**Fig. 3** Estimated functional networks for subjects 3 and 14. Black tiles correspond to edges, white to non-edges

there are also substantial differences between the two networks, that justify the further step of our procedure.

### 3.3 Regression with Covariates

The investigation of the relations between functional connectivity patterns and observed phenotypes is motivated by the subject-specific differences observed in the estimated graphs. The inclusion of covariates into the analysis of functional connectivity patterns aims to identify whether brain activity relates with personal features and behaviours and whether subject-specific information can provide insights on observed differences. Recent studies highlighted the relation among connectivity patterns and, among many others, diseases [33], violent behaviours [9, 29], or gender [20]. Functional networks, as opposed to structural information, contain important information regarding dynamical patterns of the brain architecture, and there is a promising extent of agreement between studies based either on functional or structural networks (e.g., [10]).

We investigate the relation among functional networks and covariates exploiting a simple model that encourages the interpretation of its coefficients and is able to provide interpretable insights on the effect of phenotypes over the structural network. Differently from standard models for network data—such as ERGM [23] or latent space models [21]—we want to focus on modeling multiple adjacency matrices  $\mathbf{K}_1, \dots, \mathbf{K}_n$ , instead of a single one.

We assume that the probability of a connection between each pair  $l = (u, v)$  of brain regions, with  $u = 2, \dots, 70$  and  $v = 1, \dots, u - 1$  in the network  $\mathbf{K}_i$  can be modeled using an exponential family, with natural parameters as function of phenotypical information, such as age, mental status, handedness, and brain-region specific information, such as lobes membership.

More formally, let  $\Pr(k_{il} = 1) = \pi_{il}$  define the vectorised probability to observe a connection for subject  $i$  in the pair of brain regions  $l$ , with  $i = 1, \dots, 22$  and  $l = 1, \dots, 2415 = (70 \times 69)/2$ . We model the logit of the connection probability as a function of phenotypical and brain-region information as follow:

$$\text{logit}(\pi_{il}) = \alpha + \theta_x^T \mathbf{x}_i + \theta_z^T \mathbf{z}_l$$

In particular we considered the following variables:

- subject covariates  $\mathbf{x}_i$ : age of the subject, mental health indicating the presence/absence/unknown status of a mental problem (absence used as reference class), handedness with three categories for left/right-handed and ambidextrous (ambidextrous as reference class).
- edge covariates  $\mathbf{z}_l$ : lobe membership, indicating whether the pair  $l = (u, v)$  of brain regions is in the same lobe (not belonging to the same lobe is taken as reference class).

The resulting estimates, for a value of the smoothing parameter  $\lambda = 10$ , are reported in Table 1.

Our empirical findings suggest a strong tendency for brain regions located in the same lobe to create more connections in the functional network. Moreover, subjects with a positive mental diagnosis report, on average, a lower probability to observe connected brain regions, with respect to healthy subjects and given the effect of the remaining covariates. Individuals whose mental status is not known report, instead, a higher probability to observe a connection. Handedness of the subjects under investigation is not resulted to be a determinant of functional network. Lastly, the age of the subjects in this study seems to have an effect in the determination of the connections

**Table 1** Estimated coefficients for the GLM model,  $\lambda = 10$

	Estimate	Std. Err.	z value	Pr (>  z )
(Intercept)	-1.805	0.056	-32.225	0.000
Age	-0.002	0.001	-2.209	0.027
Hand L	-0.005	0.054	-0.084	0.933
Hand R	0.004	0.044	0.087	0.931
Diagnosis YES	-0.128	0.034	-3.784	0.000
Diagnosis UNK	0.140	0.032	4.366	0.000
Lobes	0.774	0.026	29.676	0.000

of the functional network, even though the magnitude of this effect is small enough to be negligible.

### 3.4 Multiscale Analysis

In order to assess the robustness of our empirical findings, we performed a multiscale sensitivity analysis under different settings. The core idea of the multiscale approach is that whenever a signal can be measured at multiple resolutions, such as different level of smoothing in our case, information can and should be drawn exploiting all this information jointly. The principle that there is not one “correct” resolution at which the analysis should be performed is especially soothing in our context. As in resting state fMRI, it is not clear how noise may look [8], and it is important to consider more than just one resolution, or, equivalently, to explore different noise assumptions.

In the multiscale analysis, we track the evolution of the regression coefficients as the smoothness level increases. In Table 2, we re-estimated the entire model for different values of  $\lambda$  and evaluate changes in the regression coefficients. Smoother series (greater value of  $\lambda$ ) correspond to sparser graphs; when the smoothness increases, in fact, the method is able to detect only large scale variations. Since low scale dependency are suppressed, the resulting graphical models tend to be more sparse. In general, results for the sensitivity analysis tend to validate findings presented in the previous section, and estimated coefficient in Table 2 seems coherent with what shown in Table 1.

In particular, the impact of lobes and diagnosis is quite stable across different smoothing levels, which can be interpreted as an indication of robustness with respect to different noise scenario. The handedness of the subject, on the other hand, seems to have a more erratic effect on the connectivity structure, but its contribution is not substantial in the cases analyzed. A noticeable change in such behavior can be observed for values of  $\lambda \geq 18$ , which we interpreted as a symptomatic effect of over-smoothing in the denoising step.

**Table 2** Results of the multiscale sensitivity analysis conducted over different levels of  $\lambda$ . Estimated coefficients are reported for some representative levels of  $\lambda$ , with bold coefficients indicating an associated p-value less than 0.05

$\lambda$	(Intercept)	Age	Hand L	Hand R	Diagnosis YES	Diagnosis UNK	Lobes YES
0	<b>-1.951</b>	-0.001	0.018	0.046	-0.008	<b>0.124</b>	<b>0.880</b>
4	<b>-1.838</b>	-0.002	0.014	-0.004	<b>-0.083</b>	<b>0.147</b>	<b>0.854</b>
8	<b>-1.836</b>	<b>-0.002</b>	0.025	0.015	<b>-0.116</b>	<b>0.146</b>	<b>0.818</b>
10	<b>-1.805</b>	<b>-0.002</b>	-0.005	0.004	<b>-0.128</b>	<b>0.140</b>	<b>0.774</b>
14	<b>-1.957</b>	<b>-0.003</b>	-0.098	-0.035	<b>-0.079</b>	<b>0.068</b>	<b>0.721</b>
18	<b>-2.156</b>	-0.002	<b>-0.171</b>	-0.057	-0.002	0.038	<b>0.695</b>

## 4 Discussion

The analysis of neuroimaging data is a stimulating application field that embraces several disciplines; statistics covers a determinant role in this context, since it can provide deep insights on the underlying wiring mechanisms. However, statistical modeling of multiple brain networks is still in its infancy, and the inclusion of subject-specific information within repeated networks is incomplete from a literature viewpoint.

The approach suggested in this work has guided some preliminary insights on the relationship among functional networks, brain constraints and subject-specific phenotypes. One of the main advantages of our approach is its generality; within the modular structure, each block can be as complex as data allows for, leaving room for more appropriate model when needed. We have shown that even with rather simple modules, our empirical findings seem to give reasonable insights on the covariates effect on the functional dependence structure, and the sensitivity analysis performed at different levels of smoothing of the raw data did not seem to provide contradicting results.

The use of the modular approach is motivated by the computational burden and possible model misspecification that would otherwise affect a joint model. However, a two stages approach does not take full advantage of the hierarchical structure of the model, precluding the possibility to treat all uncertainties simultaneously.

An interesting future direction consists in the inclusion of models specific for network data, capable to take into account heterogeneity within the brains architecture. This aim could be achieved including random effects pairs for each ROI of the functional network [26], or using a more appropriate model for multiway data, for example adapting [22].

**Acknowledgements** The authors are grateful to the organizing committee of StartUp Research Lucia Paci, Antonio Canale, Daniele Durante and Bruno Scarpa for giving them the opportunity to take on such an inspiring challenge in a stimulating environment. The authors also wish to thank Greg Kiar and Eric Bridgeford from NeuroData at Johns Hopkins University, for pre-processing and providing the raw DTI and R-fMRI, and the other participants to Start-Up Research for prolific discussions, both during and after the meeting.

## References

1. Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E.: A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* **26**(1), 63–72 (2006)
2. Alakörkkö, T., Saarimäki, H., Glerean, E., Saramäki, J., Korhonen, O.: Effects of spatial smoothing on functional brain networks. *Eur. J. Neurosci.* **46**(9), 2471–2480 (2017)
3. Andersen, A.H., Gash, D.M., Avison, M.J.: Principal component analysis of the dynamic response measured by fMRI a generalized linear systems framework. *Magn. Reson. Imaging* **17**(6), 795–815 (1999)
4. Ashby, F.G.: Statistical Analysis of fMRI Data. MIT Press (2011)

5. Beckmann, C.F., Smith, S.M.: Tensorial extensions of independent component analysis for multisubject fMRI analysis. *Neuroimage* **25**(1), 294–311 (2005)
6. Bennett, J., Wakefield, J.: Errors-in-variables in joint population pharmacokinetic/pharmacodynamic modeling. *Biometrics* **57**(3), 803–812 (2001)
7. Brett, M., Anton, J.L., Valabregue, R., Poline, J.B.: Region of interest analysis using the Mars-Bar toolbox for SPM 99. *Neuroimage* **16**(2), S497 (2002)
8. Bright, M.G., Murphy, K.: Is fMRI “noise” really noise? Resting state nuisance regressors remove variance with network structure. *NeuroImage* **114**, 158–169 (2015)
9. Bufkin, J.L., Luttrell, V.R.: Neuroimaging studies of aggressive and violent behavior current findings and implications for criminology and criminal justice. *Trauma Violence Abuse* **6**(2), 176–191 (2005)
10. Bullmore, E., Sporns, O.: Complex brain networks graph theoretical analysis of structural and functional systems. *Nature Rev. Neurosci.* **10**(3), 186–198 (2009)
11. Cai, T., Liu, W., Luo, X.: A constrained  $\ell_1$  minimization approach to sparse precision matrix estimation. *J. Am. Statist. Assoc.* **106**(494), 594–607 (2011)
12. Calhoun, V.D., Adali, T., McGinty, V., Pekar, J.J., Watson, T., Pearlson, G.: fMRI activation in a visual-perception task network of areas detected using the general linear model and independent components analysis. *NeuroImage* **14**(5), 1080–1088 (2001)
13. Carew, J.D., Wahba, G., Xie, X., Nordheim, E.V., Meyerand, M.E.: Optimal spline smoothing of fMRI time series by generalized cross-validation. *NeuroImage* **18**(4), 950–961 (2003)
14. Craddock, R.C., Jbabdi, S., Yan, C.G., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C., Heberlein, K., Colcombe, S., Milham, M.P.: Imaging human connectomes at the macroscale. *Nature Methods* **10**(6), 524–539 (2013)
15. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006)
16. Fan, J., Yao, Q.: Nonlinear Time Series Nonparametric and Parametric Methods. Springer Science & Business Media (2008)
17. Friedman, J., Hastie, T., Hfling, H., Tibshirani, R.: Pathwise coordinate optimization. *Ann. Appl. Statist.* **1**(2), 302–332 (2007)
18. Friedman, J., Hastie, T., Tibshirani, R.: Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9**(3), 432–441 (2008)
19. Genovese, C.R., Lazar, N.A., Nichols, T.: Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* **15**(4), 870–878 (2002)
20. Gong, G., He, Y., Evans, A.C.: Brain connectivity gender makes a difference. *Neuroscientist* **17**(5), 575–591 (2011)
21. Hoff, P., Raftery, A., Handcock, M.: Latent space approaches to social network analysis. *J. Am. Statist. Assoc.* **97**(460), 1091–1098 (2002)
22. Hoff, P.D.: Multilinear tensor regression for longitudinal relational data. *J. Am. Statist. Assoc.* **9**(3), 1169–1193 (2015)
23. Hunter, D.R., Handcock, M.S., Butts, C.T., Goodreau, S.M., Morris, M.: ergm A package to fit, simulate and diagnose exponential-family models for networks. *J. Statist. Softw.* **24**(3), nihpa 54860 (2008)
24. Kiar, G., Roncal, W., Mhembere, D., Bridgeford, E., Burns, R., Vogelstein, J.T.: ndmg Neuro-datas MRI graphs pipeline (2016). <https://doi.org/10.5281/zenodo.60206>
25. Lee, M.H., Smyser, C.D., Shimony, J.S.: Resting-state fMRI a review of methods and clinical applications. *Am. J. Neuroradiol.* **34**(10), 1866–1872 (2013). methods
26. Li, H., Loken, E.: A unified theory of statistical analysis and inference for variance component models for dyadic data. *Statist. Sin.* **12**, 519–535 (2002)
27. Lipsey, M.W., Wilson, D.B.: Practical Meta-Analysis. Sage Publications, Inc. (2001)
28. Liu, F., Bayarri, M., Berger, J.: Modularization in Bayesian analysis, with emphasis on analysis of computer models. *Bayesian Anal.* **4**(1), 119–150 (2009)

29. Mehta, P.H., Beer, J.: Neural mechanisms of the testosterone-aggression relation the role of orbitofrontal cortex. *J. Cognit. Neurosci.* **22**(10), 2357–2368 (2010)
30. Murphy, K.M., Topel, R.H.: Estimation and inference in two-step econometric models. *J. Bus. Econ. Statist.* **20**(1), 88–97 (2002)
31. Rosenbaum, P.R., Rubin, D.B.: The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1), 41–55 (1983)
32. Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.: Tract-based spatial statistics voxelwise analysis of multi-subject diffusion data. *Neuroimage* **31**(4), 1487–1505 (2006)
33. Stam, C.: Modern network science of neurological disorders. *Nature Rev. Neurosci.* **15**, 683–695 (2014)
34. Zhao, T., Liu, H., Roeder, K., Lafferty, J., Wasserman, L.: The huge package for high-dimensional undirected graph estimation in R. *J. Mach. Learn. Res.* **13**, 1059–1062 (2012)

# Three Testing Perspectives on Connectome Data



Alessandra Cabassi, Alessandro Casa, Matteo Fontana, Massimiliano Russo  
and Alessio Farcomeni

**Abstract** Guided by an application in the analysis of Magnetic Resonance Imaging (MRI) scans from the neuroimaging realm, we provide some perspectives on statistical techniques that are able to address issues commonly encountered when dealing with Magnetic Resonance images. The first section of the chapter is devoted to a bootstrap-based inferential tool to test for correlation between anatomy and functional activity. The second provides a Bayesian framework to improve estimation of fiber counts from Diffusion Tensor Imaging (DTI) scans. The third one introduces an object-oriented framework to explore and perform testing over network-valued data.

**Keywords** Hypothesis testing · Permutation tests · Object oriented data analysis  
Bootstrap inference · Bayesian statistics · Graphical lasso

## 1 Introduction

The exceptional improvement in medical imaging techniques has provided clinicians and data analysts with a plethora of data objects, that can be seen as “big” in several ways: many data points, many features per data point, as well as complex

---

A. Cabassi

MRC Biostatistics Unit, School of Clinical Medicine, University of Cambridge, Cambridge, UK  
e-mail: [ac2051@cam.ac.uk](mailto:ac2051@cam.ac.uk)

A. Casa · M. Russo

Department of Statistical Sciences, University of Padova, Padua, Italy  
e-mail: [casa@stat.unipd.it](mailto:casa@stat.unipd.it)

M. Russo

e-mail: [russo@stat.unipd.it](mailto:russo@stat.unipd.it)

M. Fontana (✉)

Department of Management, Economics and Industrial Engineering, DIG, Politecnico di Milano,  
Milano, Italy  
e-mail: [matteo.fontana@polimi.it](mailto:matteo.fontana@polimi.it)

A. Farcomeni

Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy  
e-mail: [alessio.farcomeni@uniroma1.it](mailto:alessio.farcomeni@uniroma1.it)

dependencies and patterns inside a given data point. In particular, *Magnetic Resonance Imaging (MRI)* techniques have proven to be extremely powerful in unveiling novel insights about human anatomy and physiology, in particular in neurology and neuroimaging. MRI scans of the brain provide a lot of information but, even if the actual acquisition technique is stable and widely used, we still lack recognized statistical methods that are able to deal with the sheer complexity of the data generated.

In this contribution we aim to provide insights about some of the issues a statistician has to deal with when working with MRI data. The work embraces three different aspects of the analysis of MRI scans. In Sect. 2 we develop a bootstrap-based inferential tool to test if the functional connectivity among different brain areas corresponds to their structural connectivity and anatomical characteristics. In Sect. 3 we provide a Bayesian framework to estimate the fiber-count number provided by *Diffusion Tensor Imaging (DTI)* data. Finally, in Sect. 4 we describe an object-oriented approach to exploratory data analysis and hypothesis testing for network data. We shall see how it is possible to explore network valued datasets, and develop tests for the equality of network data, making use of results from functional data analysis, object-oriented data analysis and permutation testing.

## 2 Testing Functional Correlations in Connectomic Maps

### 2.1 *Background and Motivation*

The correlation in the activity of brain regions is known as *functional connectivity*. A set of brain areas together with the connections among them is then called a *functional connectome* or a *functional network*. In the past decades there has been an increasing attention on the detection of patterns of connections and activity in the human brain [14]. In the neuroscience it is widely assumed that the functional connectome should reflect the underlying structural networks, i.e. the anatomical links among different brain regions [26]. However, as [17] points out, the nature of these relations among different types of connectivity is not completely clear and it is still worth to investigate on it. There are indeed several questions still waiting for an answer. For example, it is interesting to study if it is possible to infer structural connections from functional ones and to obtain some indications about how these networks and their relations vary across time. The work by [17] tries to answer to some of these questions, highlighting some interesting results.

The aim of our method turns out to be slightly different since we propose a test to check if the absence of white matter fibers connecting brain regions is reflected in their functional correlation. Thus, the underlying hypothesis is confirmative since is related to the idea that the absence of connections between two different brain areas in the structural network should be reflected also in the functional network. Furthermore, if the relation among structural and functional connectomes is taken

for granted, the proposed method could be seen as a way to assess the quality of the collected data.

Before presenting the proposed procedure, we briefly introduce the considered statistical background and some of the ideas this work is built on.

To analyze and construct functional and structural networks we resort to graphical models. These are probabilistic models where a graph is used to express the conditional dependence structure between sets of observed random variables. In the last few years we have witnessed an increasing interest in these models since they constitute a useful tool to obtain information and to understand the relationships among variables in an intuitive way.

When dealing with a  $n \times p$  data matrix  $X$  whose rows are independent and identically distributed as a Gaussian random variable  $N(0, \Sigma)$  the interest is shifted towards the estimation of the so called *precision matrix*  $\Theta = \Sigma^{-1}$ . Precision matrices turn out to be particularly relevant since their estimation can be thought of as a way to estimate a graphical model: if we associate a node in a graph to each observed variable, there is an edge among node  $i$  and node  $j$  if and only if  $\Theta_{ij} \neq 0$ . Thus the zero entries in  $\Theta$  give a clear indication about conditional independence among variables.

Given  $S = X^T X/n$ , the empirical covariance matrix, the maximum likelihood estimate for  $\Theta$  is obtained maximizing the following profile log-likelihood

$$\log |\Theta| - \text{tr}(S\Theta), \quad (1)$$

where  $|A|$  is the determinant and  $\text{tr}(A)$  the trace of a matrix  $A$ . The maximization in (1) leads to  $\hat{\Theta} = S^{-1}$ , generally not containing elements equal to zero; note that in high-dimensional cases, when  $p > n$ , the maximum likelihood estimate cannot be computed since  $S$  will be singular. To overcome such a drawback and to induce sparsity [21] propose to fit a lasso model to each variable using the others as predictors. Thus  $\hat{\Theta}_{ij}$  is estimated to be zero if both the coefficient of variable  $i$  on  $j$  and the coefficient of variable  $j$  on  $i$  are zero.

Alternatively a penalized log-likelihood has been proposed, where the estimate for the precision matrix results from maximizing

$$\log |\Theta| - \text{tr}(S\Theta) - \lambda \|\Theta\|_1,$$

where, in a *lasso*-type fashion,  $\lambda$  is a tuning parameter and  $\|A\|_1 = \sum_{i \neq j} |A_{ij}|$  is the component-wise  $L_1$  norm of the matrix  $A$  excluding the diagonal. This approach has been proposed by [13] and goes under the name of *graphical lasso*. It has a relevant advantage that an estimate for  $\Theta$  is given even when  $S$  is singular. Furthermore a sparse representation of the precision matrix, and consequently of dependence structures of data, is provided. As  $\lambda$  varies, the level of sparsity changes and so does the dependence among variables.

Given the interpretation of the entries of  $\Theta$ , inference procedures on these entries are crucial in assessing if the conditional dependence among two variables are statistically significant. One of the first attempts to derive rigorous inferential tools in

this setting can be found in [16]. The authors proposed a set of hypotheses and test statistics corresponding to a sequence of  $\lambda$  along the solution path of the graphical lasso. The approach consists in testing if the variables that should be connected are actually in the same connected component.

Another inferential method, closely related to Gaussian graphical models and graphical lasso, can be found in [4]. The goal of this work is to identify differences in Gaussian graphical models known to have similar structure. They aim to determine which edges are different between two populations for which we expect different brain activity and connections (e.g. autistic and non-autistic in their work, left-handed and right-handed or people diagnosed with mental disorder or not, in our data). Lastly, note that the authors suggest the use of Gaussian graphical models, thus the estimation of the precision matrix, as a way to estimate and construct the functional connectivity network.

Hence, even if the statistical tools considered are similar, the previously mentioned works try to answer to slightly different questions with respect to our procedure. In the next section the methodology will be introduced in detail and the results obtained on our data are shown.

## 2.2 Methodology and Application

The aim of our proposed method is to test if the anatomical connections among brain regions are reflected in their functional connections. It can be seen both as a method to confirm the widely assumed link among structural and functional network and as a way to assess the quality of the collected data, if the former assumption is taken to be true.

We propose a test based on a parametric bootstrap approach (for an intuitive sketch of the procedure see Pseudo-algorithm 1) comparing the value of the test statistic evaluated on the observed data with the bootstrap distribution obtained sampling from the null distribution. In this situation the null hypothesis is that the absence of white fibers matter connecting brain areas is reflected in the absence of a functional correlation among them. In a more synthetic way the test could be represented as

$$H_0 : \Omega = \Omega_0 \quad \text{versus} \quad H_1 : \Omega \neq \Omega_0,$$

where  $\Omega$  and  $\Omega_0$  are two correlation matrices with the second one constrained by external information, in our situation coming from structural network.

Assume we have a  $n \times p$  data matrix  $X$  containing observations about functional activity of  $p$  different brain regions recorded for  $n$  distinct subjects. From  $X$  we obtain the unconstrained sample covariance matrix  $C$ , evaluating the functional covariances between the areas. Then we consider the matrix  $D$  expressing the structural network for the individuals and we estimate, using the graphical lasso, a constrained covariance matrix  $C^*$  where the constraints are given by the zero entries in the matrix  $D$ . The aim is to obtain an estimate  $C^*$  such that its inverse has zero entries if and

**Algorithm 1** *Parametric bootstrap procedure*

Denote with  $X$  an  $n \times p$  matrix containing functional activity of  $p$  brain regions for  $n$  different subjects. Let  $D$  be a matrix representing the structural network, hence containing the number of fibers connecting each pair of brain areas.

---

**Input**  $X, D$ .

- 1: from  $X$ , obtain the covariance matrix  $C$ ;
- 2: using the graphical lasso estimate  $C^*$  such that  $(C^*)_{ij}^{-1} = 0$  iff  $D_{ij} = 0$ ;
- 3: obtain  $C_1^*, \dots, C_B^*$  matrices sampling from Wishart distribution with scale matrix  $C^*$ ;
- 4: from  $C$  and  $C_k^*$ ,  $k = 1, \dots, B$ , obtain the correlation matrices  $c$  and  $c_k^*$ ;
- 5:  $S(c)$ : sum of the squared correlations among not connected regions;
- 6: compare  $S(c)$  with the bootstrap distribution obtained from  $S(c_i^*)$  with  $i, \dots, B$ ;
- 7: compute bootstrap  $p$ -value  $p_{\text{obs}}$ .

**Output:**  $p_{\text{obs}}$ ,  $p$ -value of the bootstrap test.

---

only if we have zero entries in the corresponding elements of  $D$ , i.e.  $(C^*)_{ij}^{-1} = 0$  iff  $D_{ij} = 0$ . In this way we are estimating a functional connectome considering a constraint which consists in the absence of edges when two areas are not anatomically connected, hence reflecting the nature of the null hypothesis. The graphical lasso is thus used mainly as a tool to obtain the constrained estimate; in a sense the degree of sparsity is considered fixed and linked to the specific hypothesis that we are testing.

Once  $C^*$  is obtained, we get  $B$  matrices sampling from  $\mathcal{W}_p(C^*, n)$ , a Wishart distribution with  $n$  degrees of freedom and scale matrix  $C^*$ . Note that  $n$  has to be greater than  $(p - 1)$ : this could constitute a serious limitation in the specific considered neurological context but, as we will see, it turns out not to be a problem in our application. The choice of the Wishart distribution as the sampling one appears reasonable and could be motivated by standard distributional results for the sample covariance matrix when data comes from a multivariate normal distribution.

After constructing the correlation matrices  $c$  and  $c_k^*$ ,  $k = 1, \dots, B$  from the corresponding covariance matrices  $C$  and  $C_k^*$ , the test statistic is given by  $S(c) = \sum_U c_{ij}^2$  where  $U$  represents the set of anatomically unconnected brain regions. Therefore, as test statistic, we are simply considering the sum of the squared functional correlation values for the unconnected regions: if the null hypothesis is true this sum should be small, supporting the assumption about the duality between functional and structural connectome. The inferential indication is thus obtained comparing the value of  $S(c)$  with the bootstrap distribution of  $S(c_k^*)$  with  $k = 1, \dots, B$ .

In our application we have dynamical functional activity measured on 24 subjects for 70 different brain regions over 404 moments in time. Furthermore these measures are taken in two different scans; we have conducted our analysis considering both scans. Regarding the first scan, we ignored the last measured time and two different subjects due to the large amount of missing values. For the second scan the missing values issue is more serious and we have had to remove again the last measured time as well as 13 subjects. In both cases we are in a high-dimensional setting since  $n_1 = 22$  and  $n_2 = 11$ , with  $n_i$  referring to the sample size for the  $i$ th scan, and  $p = 70$ . Hence the sparsity implied by the graphical lasso is required to obtain an estimate

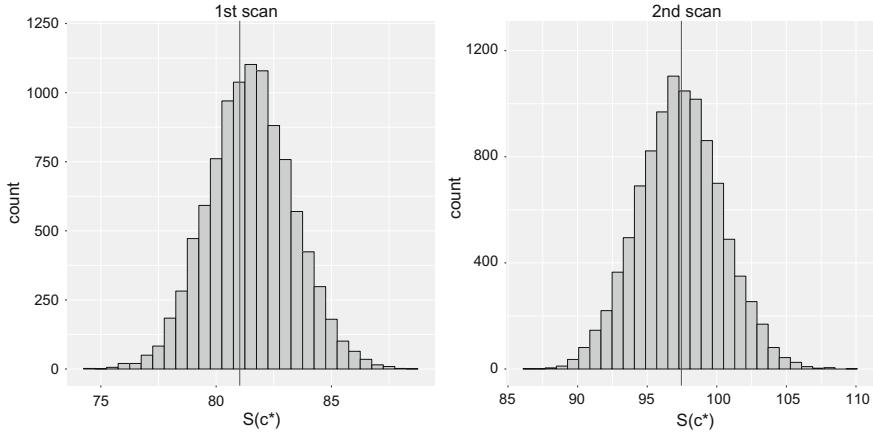
for the covariance matrix. Nonetheless, in principle we cannot consider the Wishart distribution as the sampling one since it would require  $n_i > p - 1$  with  $i = 1, 2$ . Nevertheless, we bypass this problem simply by stacking the temporal dimension thus obtaining the functional activities matrix with dimension  $n_i T \times p$ , with  $i = 1, 2$ ,  $p = 70$  and  $T = 403$ . Hence we are treating each observation in different moments of time as coming from a different subject, enlarging our effective sample size. Losing temporal dependencies among measurements could be harmful but, in our situation, some exploratory analyses have suggested that the informative content in the time domain is quite poor.

Note that different subjects could have different anatomical characteristics. In order to decide which elements of the precision matrix have to be estimated exactly equal to zero we have collapsed the structural network among individuals. In this way a zero entry in the precision matrix corresponds to two brain regions not being connected in any of the considered subjects; this choice could be somewhat arbitrary and could be relaxed. Among all the possible 2415 connections between brain areas, we have 816 anatomically unconnected regions in the first scan and 999 in the second one, indicating also a certain amount of variability between scans.

A graphical summary of the obtained results is shown in Fig. 1 where we compare the value of the test statistic computed from our data with the bootstrap distribution obtained by sampling from the null distribution. In both the first and the second scan we do not reject the null hypothesis, hence our data do not support the absence of a relation among the two networks. The obtained results hence appear to be consistent with the usual assumption made in the neuroscience community. For the sake of comparison, we consider also a likelihood ratio test comparing the constrained estimate of the covariance matrix with the unconstrained one and the obtained results are in agreement to the one mentioned above. We would like to highlight that these two tests are based on a different rationale, since the proposed test is not comparing directly the two matrices but is specifically focusing on that covariances between areas having a corresponding zero entry in the structural network.

In conclusion we proposed a simple, fast and easily implementable parametric bootstrap-based test that turns out to be useful to check the relation among the anatomical features and the functional connections among different brain regions.

Some possible directions for improvements and further research could require an accurate simulation study to check the statistical characteristics of the proposed test and to compare it with other possible solutions in the literature. As [4] point out, one of the main problems in this framework is indeed related to the statistical power of the tests used; a simulation study should help to shed light on it for the proposed method. Furthermore, it could be interesting to handle the temporal dimension more carefully than we did, trying to incorporate information about different moments of time. It is indeed well known that the functional connectivity reconfigures in less than seconds while the structural connections are more stable. This could lead to different conclusions depending on the time and on the specific activity required of the subject. Other useful information to integrate and to consider could be related to the distances among brain regions and to the specific condition of the considered individual (e.g. affected or not from some neurological disease).



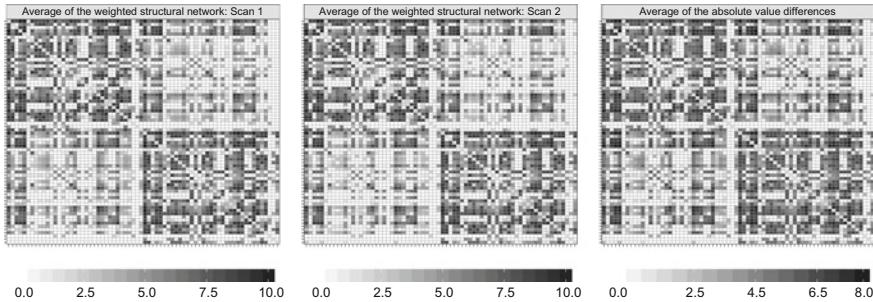
**Fig. 1** Bootstrap distribution of  $S(C^*)$  for the first scan (left) and the second one (right). The vertical lines denote the values of the observed  $S(c)$

### 3 A Bayesian Framework for Fiber Count Estimation

#### 3.1 Introduction

DTI is a method capable of mapping the fibers' architecture of tissues (e.g. nervous tissue, muscle) *in vivo*, and has been extensively used to detect the number of white matter fibers connecting areas of the brain. This technique was rapidly implemented by major MRI scanner companies, and due to the great availability of data, and to the plausibility of some results, DTI was viewed by imaging neuroscientists as a powerful and unique new tool for exploring the structural connectivity of the human brain. However, as noted in [18], DTI is a rather approximate technique which has frequently been given implausible interpretations leading to misleading results. Coherently with their findings, [5] pointed out that the reproducibility of the whole brain structural connectome, for the same subject and in similar experimental conditions, is affected by several external causes, including the quality of the scanner used, the method applied for calculating connectivity from DTI fiber tracking, and anatomical properties of each link. The effect of such inter-individual variability is evident in the considered data where we can observe white matter fiber counts for the same subject and brain's region that differ of several thousand units across the two considered scans.

To mitigate this effect related to high variability present in the available data we propose a hierarchical Bayesian model to estimate the effective unknown number of white matter fibers connecting each pair of brain regions. To accomplish this goal we leverage available information both at subject and brain region scale. We describe our approach with the aim of directly estimating the number of white matter fibers in each brain's region and for each subject, however the proposed procedure



**Fig. 2** Graphical representation of the weighted structural networks measuring fiber counts in  $\log(1+x)$  scale, averaged across subjects for Scan 1 (left), Scan 2 (center). The right panel represents instead the averaged absolute value of the differences across scans

should be considered as a building block for more refined Bayesian models, allowing additional uncertainty quantification on the number of fibers. Findings deriving from this procedure should in any case be compared with the ones obtained using raw white matter fiber counts highlighting the differences and assessing sensitivity of the results with respect to the specified model.

DTI brain data are available at different resolutions, going from single voxels to broader brain regions. Several classifications have been proposed, and for this contribution we focus on the one identified by the Desikan atlas parcellation [8], consisting of 70 interconnected areas. Our findings can, however, be applied to more refined and potentially more complicated brain structures. The available data consists of two brain scans for  $n = 24$  subjects. We discarded 4 of them because of the presence of missing data in both scans, hence information on the number of white matter fibers connecting each pair of brain regions is available via  $70 \times 70$  dimensional matrices  $D_{ki}$  for each subject  $i = 1, \dots, 20$  considered in the analysis, and scan  $k = 1, 2$ . The structural networks are sparse, with a moderate number of fibers counts being zero, and others having a wide range of variability both across subjects and across scans for the same subject. The average number of white matter fibers connecting all pair of regions for both scans is showed in Fig. 2, together with the absolute value of the difference of the two scan counts. We can notice that, although the average number of connections is quite stable across the two scans, individual variability is very high. To limit the effect of such variability, while seeking more robust findings, a broad section of research focused on binary structural networks measuring presence or absence of white matter fibers (e.g. [1]). Compared to raw counts, binary networks are in general more stable across scans, nonetheless they still retain a high inter individual variability. Moreover, this approach deeply reduces the amount of information used for modeling purposes.

### 3.2 Model Formulation

With the aim of estimating the effective number of white matter fibers connecting each pair of brain regions for each subject, we consider as data the raw counts  $d_{kij}$  for each  $i = 1, \dots, n = 20$  subject,  $k = 1, \dots, K = 2$  scan, and  $j = 1, \dots, J = 2415$  pairs of brain regions, obtained by stacking the elements of the lower triangular part of  $D_{ki}$ . Each of the two available scans is considered as a sample from a Binomial distribution with a common index  $M_{ij}$ . To take into account uncertainty in the value of  $M_{ij}$ , representing the unknown amount of white mater fibers connecting pair of regions  $j$  for subject  $i$ , we consider a Poisson distribution for this index, depending on both subject specific and area specific covariates. The considered model can be expressed in the following hierarchical form

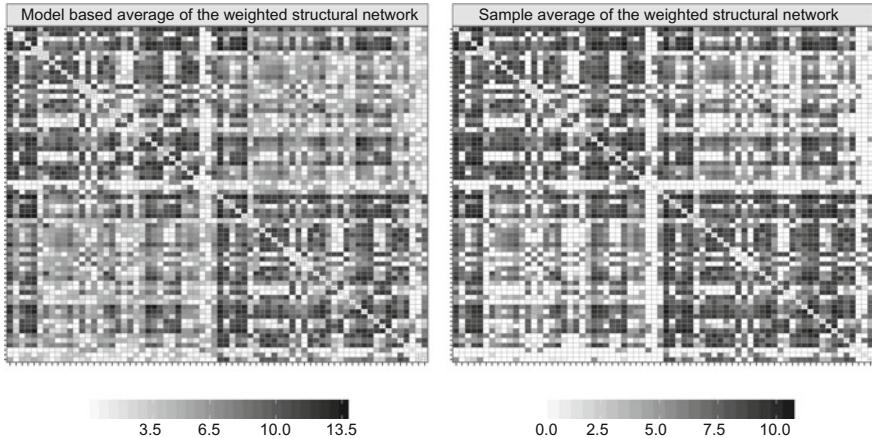
$$\begin{aligned} (\{d_{kij} : k = 1, \dots, K\}) &\sim \text{Bin}(M_{ij}, \pi_j), \\ \text{logit}(\pi_j) &= \alpha_j + \alpha \text{MatchHemisphere}_j, \\ M_{ij} &\sim \text{Pois}(\lambda_{ij}), \\ \log(\lambda_{ij}) &= \beta_i + \beta_j + \beta \text{age}_i, \end{aligned} \tag{2}$$

where the variable  $\text{age}_i$  indicates the age of the subject  $i$ , while  $\text{MatchHemisphere}_j$  is a dichotomous variable indicating if the pair of brain regions indexed by  $j$  share the same hemisphere. It is worth noting that according to model specification (2) the probability of observing a single connection in the pair  $j$ , denoted by  $\pi_j$ , does not depend upon individual scale parameters. This hypothesis reflects the fact that the probability of observing a connection might be influenced by region-specific covariates, such as shape, size, anatomical location of the considered brain areas, and, as in our case, if the regions share the same hemisphere. As noted in [5], these characteristics can in fact alter the sensitivity of the machine used to produce the data, which is not expected to change across subjects. We can also note that without additional constraints both individual specific and brain region specific parameters in model (2) are not identifiable, however this raises no concern when the inference is just targeted towards estimation of the white matter fibers connecting brain regions.

The posterior distribution for model (2) cannot be obtained in closed form, but an MCMC algorithm can be used to draw samples from it. To produce a better mixing we integrate the parameter  $M_{ij}$  with respect to its distribution. The integrated model can be easily obtained leveraging Proposition 1.

**Proposition 1** *Let  $N | M \sim \text{Bin}(M, \pi)$  and  $M \sim \text{Pois}(\lambda)$  then  $N \sim \text{Pois}(\lambda\pi)$ .*

The resulting integrated model is a special case of a Poisson regression with an unknown offset, where the values of  $\pi_j$  represents the population probability of a connection between the pair of regions  $j = 1, \dots, J$ . Alternatively, we can interpret the values  $\pi_j$  as information about the machine resolution for each pair of locations  $j$ .



**Fig. 3** Graphical representation of the structural networks measuring logarithm of 1 plus fiber counts, averaged across subjects and scans, using model (2) (right) and using the sample mean (left)

Similar offset estimation has been carried out in Poisson factor model using maximum likelihood estimation in [19] with the aim of normalizing Rna-seq counts derived from multiple experiments.

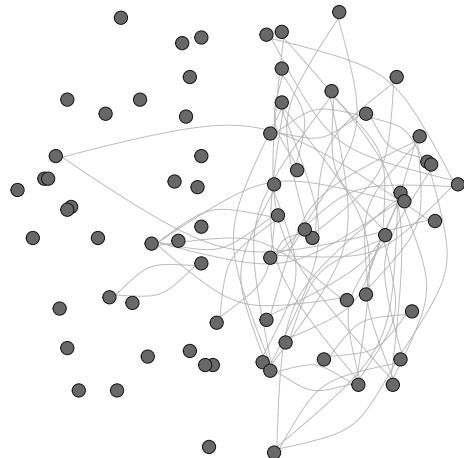
### 3.3 Application to DTI Data

We used non informative normal priors for all the coefficients in model (2), relying on Hamiltonian Monte Carlo implemented in the R packages [32] for the sampling. We run the algorithm for 5000 iterations, and trace plots suggested that convergence has been reached after a burn-in of 2500.

Figure 3 shows the estimated average number of counts for the brain regions, obtained from the posterior sample of model (2), and the sample average. We can notice that the connection pattern looks almost the same as the original one—we have the same active connections—while the model based estimate has, on average, a slightly higher fiber count compared to the naive estimate, obtained as sample average across subjects and scans. This finding agrees with the current literature which tells that some DTI scans might present an underestimate of the fiber counts due to several exogenous causes (e.g. [31]).

Additionally the  $\pi_j$ s indicate which areas have an higher connection probability according to the results of model (2). Provided the validity of model (2), these parameters can be interpreted as giving information on the regions in which is easier to measure connections. The first 60 highest  $\pi_j$ s are represented in Fig. 4, indicating

**Fig. 4** Network based representation of the brain where the edges indicates the 60 pairs having higher values of the posterior mean of  $\pi_j$  according to model (2)



that, according to (2), regions with higher probability to observe connections share the same hemisphere and are located in the right one.

In conclusion, DTI data represent a valid source of information for data analysis, however they should be treated with care, and information on how the data have been collected and preprocessed should always be considered during data analysis and interpretation. In our opinion the proposed approach can mitigate the effect of subject-specific variability, leading to more reliable estimates of the fiber counts that can help assessing sensitivity with respect to possible undercount effects.

## 4 Object-Oriented Nonparametric Exploration and Hypothesis Testing for Network Data

### 4.1 Introduction

In this section we develop an object-oriented nonparametric test for the equality of two or more groups of functional networks derived from the functional magnetic resonance imaging (fMRI) data. We work in the context of Object-Oriented Data Analysis [20, 33], the branch of statistics that treats data that are “complex”, in the sense that they do not live in the classical  $\mathbb{R}^n$  space. Using this perspective, we consider the whole network as the statistical unit of our analysis, and we aim to use techniques that analyze the data in the mathematical space in which they live.

Several other methods already exist in the literature to address this problem (see for example [6, 30]). Many of them proceed by reducing each observed network to a vector of summary statistics, as in [25, 28, 29]. However, using summary statistics usually leads to the loss of important information that can explain differences across groups [2]. Others have proposed univariate testing approaches that consider each

edge separately and then adjust the resulting  $p$ -values to control the false discovery rate [11] or the family-wise error rate [34] with thresholding procedures that take into account the network structure. [27] gain power in multiple testing by using auxiliary data such as spatial proximity to inform the posterior probability that specific pairs of nodes interact differently across groups or with respect to a baseline. [10] develop a Bayesian procedure for inference and testing of group differences in the network structure, which relies on a nonparametric representation for the conditional probability mass function associated with a network-valued random variable. [15] propose to test the equality of two groups of networks using the concept of Fréchet mean of networks and deriving a central limit theorem for sequences of network averages, using the Euclidean distance. Here we focus on inference on the average network, but we use distances that have been shown to perform better than the Euclidean distance in previous studies [9].

## 4.2 Metrics for Network Data

### 4.2.1 The Procrustes Size-and-Shape Distance

Let  $G_1, \dots, G_N$  denote the adjacency matrices of  $N$  graphs, each assumed to have the same number of vertices  $V$ . The  $G_i$ s are assumed to be positive semi-definite (PSD) and independent and identically distributed according to a distribution  $f$ .

Given a distance  $d(\cdot, \cdot)$  between PSD matrices and a probability distribution for a  $V \times V$  PSD matrix  $G$  on a Riemannian matrix space with probability distribution  $q(G)$ , we can define the *Fréchet mean* as

$$\Gamma = \arg \inf_{\Gamma} \frac{1}{2} \int d(G, \Gamma)^2 q(G) dG, \quad (3)$$

[12]. Moreover, given a sample  $G_1, \dots, G_N$  of independent and identically distributed observations, the *sample Fréchet mean* is

$$\hat{\Gamma} = \arg \inf_{\Gamma} \sum_{n=1}^N d(G_n, \Gamma)^2, \quad (4)$$

Contrary to [15], we use non-Euclidean metrics, that have been shown to perform better in practice. In particular, we consider the square root distance and the Procrustes size-and-shape distance, that have proved useful for positive-semidefinite matrices [9]. Given two PSD matrices  $G_1$  and  $G_2$ , the former is defined as

$$d_S(G_1, G_2) = \|G_1^{1/2} - G_2^{1/2}\|, \quad (5)$$

and the latter is

$$d_P(G_1, G_2) = \inf_{R \in O(V)} \|L_1 - L_2 R\|, \quad (6)$$

where  $L_i$  is a decomposition of  $G_i$  such that  $G_i = L_i L'_i$ ,  $i = 1, 2$ ,  $O(V)$  is the set of  $V \times V$  orthogonal matrices and  $\|\cdot\|$  denotes the Frobenius norm. However, any other distance between PSD matrices can be used, such as for example the log-Euclidean distance of [3].

#### 4.2.2 The Metric Approach and Gromov-Wasserstein Distances

The Procrustes size-and-shape distance is defined only for PSDs. While this condition is verified for fMRI connectomes (that are positive semidefinite by construction), this is in general not true for DTI connectomes, for which the element of the connectome adjacency matrix  $g_{ij}$  is the count of fibers that connect ROI  $i$  to ROI  $j$ . Moreover, while for this application we are considering a fixed number of ROIs, positioned on an atlas that is common for all subjects, this is actually not the case for more general applications of statistical methods for network-valued data objects. For this reason, we also propose as a viable network-oriented approach the one by [23], based on Gromov-Wasserstein (GW) distance.

The basic idea behind the use of the GW distance is to see data points as represented by metric spaces, then, looking for a sufficiently rich, abstract metric space  $Z$  that admits isometric copies of available data, and then computing some kind of distance between the isometric copies. The arbitrary nature of the procedure can be eliminated by optimizing over the choice of  $Z$ .

We now very briefly define Gromov-Hausdorff and Gromov-Wasserstein distances: for a full understanding of the theoretical properties and features of these distances, the reader should refer to [23], and references therein. The Gromov-Hausdorff distance between compact metric spaces  $(X, d_X)$  and  $(Y, d_Y)$  is:

$$d_{GH} = \frac{1}{2} \inf_R \sup_{(x,y),(x',y') \in R \times R} |d_X(x, x') - d_Y(y, y')|,$$

where  $R$  ranges over the set of all correspondences between  $X$  and  $Y$ , denoted by  $\mathcal{R}(X, Y)$ . The previous formulation is equivalent to:

$$d_{GH} = \frac{1}{2} \inf_R \|d_X(x, x') - d_Y(y, y')\|_{L^\infty_{R \times R}}.$$

The Gromov-Wasserstein distance is obtained by replacing the  $L^\infty$  norm with a  $L^p$  one, thus obtaining

$$d_{GH} = \frac{1}{2} \inf_R \|d_X(x, x') - d_Y(y, y')\|_{L^p_{R \times R}}.$$

It can be shown that this formulation becomes more tractable from a mathematical point of view, yielding a continuous (even if not convex) optimization problem, instead of a combinatoric one.

### 4.3 Hypothesis Testing

Let  $G_{11}, \dots, G_{N_1}$  and  $G_{12}, \dots, G_{N_2}$  be two groups of adjacency matrices where the observations in each group are independent and identically distributed samples from two random processes with mean  $\Gamma_1$  and  $\Gamma_2$  respectively. We would like to test the hypothesis

$$H_0 : \Gamma_1 = \Gamma_2 \quad \text{against} \quad H_1 : \Gamma_1 \neq \Gamma_2. \quad (7)$$

We can adopt a similar strategy to the one used in [24] for testing the equality of covariance operators of functional data, i.e. we reformulate the test as follows

$$H_0 : d(\Gamma_1, \Gamma_2) = 0 \quad \text{against} \quad H_1 : d(\Gamma_1, \Gamma_2) > 0 \quad (8)$$

and we use the distance between the sample Fréchet means of the two groups  $d(\hat{G}_1, \hat{G}_2)$  as a test statistic. Then we consider  $B$  permutations of the group labels and for each of them compute the distance between the sample means  $\hat{G}_i^*, i = 1, 2$  of the permuted samples. The  $p$ -value of the test is the proportion of the  $d(\hat{G}_1^*, \hat{G}_2^*)$  that are greater than or equal to  $d(\hat{G}_1, \hat{G}_2)$ . Similarly, the comparison of multiple samples can be done using synchronized permutations, as explained in [7].

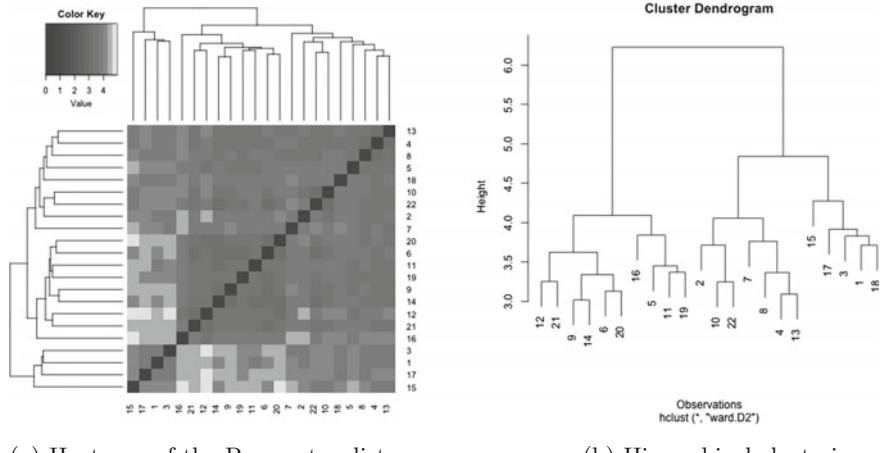
### 4.4 Results

From the imaging data, we construct a functional network, calculated as the Spearman correlation of the fMRI scans. We are interested in testing the equality of the mean network in different groups of patients. We are also looking at the morphologic connectome as obtained by DTI scans: the edges weights  $n_{ij}$  are the count of the number of fibers from ROI  $i$  to ROI  $j$ . Throughout this section we only consider data obtained from the first scan.

#### 4.4.1 Exploratory Analysis

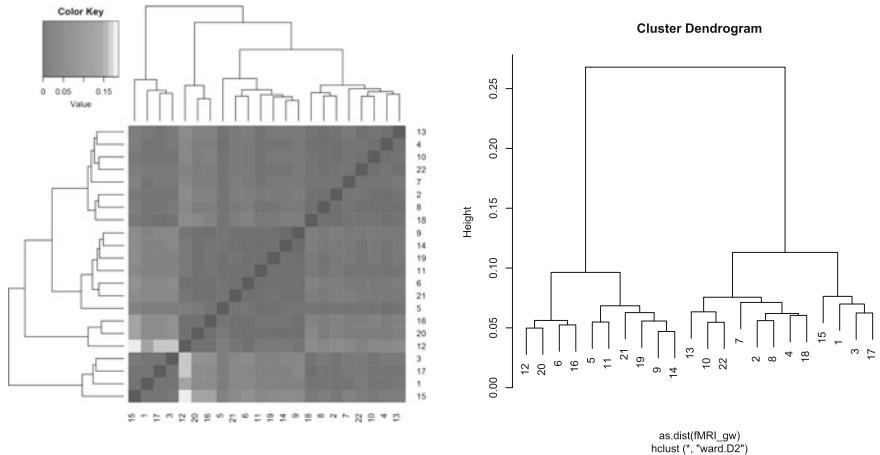
First, we want to demonstrate the use of the distances defined above in the more intuitive case of data exploration and clustering. In particular, here we are interested to see if there are evident differences between the observed brain networks. To this end, we utilize the Procrustes distances to apply standard clustering methods to the data. In Fig. 5a is shown a plot of the matrix containing the differences between each pair of individuals. The hierarchical clustering performed on that matrix of distances is shown in Fig. 5b. From this exploratory analysis, there seems to be two main clusters of people.

We also propose the same analysis using GW distances. In Fig. 6a we can see a heatmap of the matrix containing the distances between different pairs of individuals.



(a) Heatmap of the Procrustes distances.

(b) Hierarchical clustering.

**Fig. 5** Exploratory analysis of fMRI data with the Procrustes distance

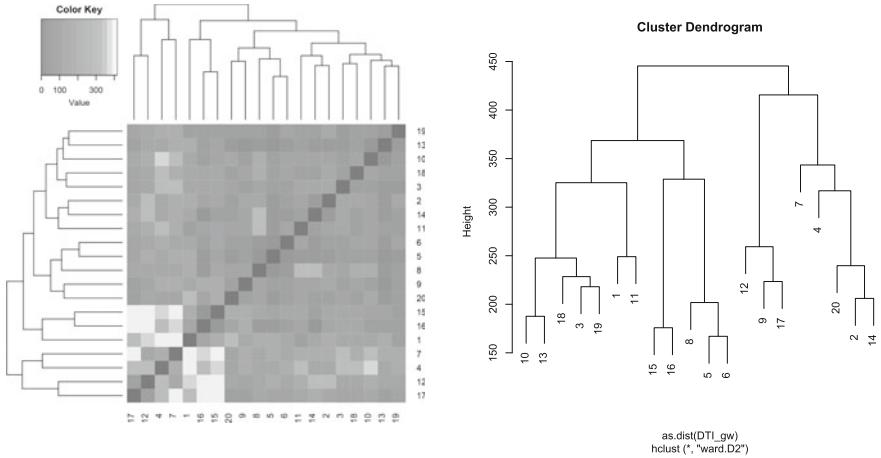
(a) Heatmap of the Gromov-Wasserstein distances.

(b) Hierarchical clustering.

**Fig. 6** Exploratory analysis of fMRI data with the Gromov-Wasserstein distance

In Fig. 6b the hierarchical clustering dendrogram performed using Ward linkage is shown. The results seem to confirm the presence of two clusters of subjects, and appear to be even sharper than the ones obtained using the Procrustes distance.

The flexibility of the Gromov-Wasserstein distance, that is defined for every matrix, and not only PSMs allows us to perform an exploratory data analysis also on the DTI data. The results are shown in Fig. 7. In Fig. 7a the usual distance heatmap



(a) Heatmap of the Gromov-Wasserstein distances.

(b) Hierarchical clustering.

**Fig. 7** Exploratory analysis of DTI data with the Gromov-Wasserstein distance

can be seen: again a bipartite nature of the dataset can be identified. This intuition is again supported by the analysis of the dendrogram in Fig. 7b.

#### 4.4.2 Hypothesis Testing

After exploring the data with the newly defined distances for networks, we can formulate and test hypotheses about them. In particular, we want to test if there exists a difference between the average functional networks in the following groups of people:

1. People who have or have had during their life a diagnosis of a mental disorder versus the others;
2. People who are less than 30 years old versus the others;
3. People who are less than 50 years old versus the others.

It is important to note that for the first test we are considering only 18 people, since the information about previous and current mental disorder diagnoses is not available for all the individuals considered in this study. The  $p$ -values of the tests of these hypotheses are reported in Table 1. We used a Bonferroni-Holm procedure to control the family-wise error rate.

Despite the very small sample size, the third test has a low  $p$ -value. This may indicate that there are some differences in the functional brain networks of people over and under 50 years old. Interestingly, all the individuals over 50 belong to the cluster on the left in Fig. 5b.

**Table 1** *p*-values of the tests

Test	<i>p</i> -value	Adjusted <i>p</i> -value
1. Mental disorder diagnosis	0.914	1
2. Under versus over 30	0.634	1
3. Under versus over 50	0.091	0.273

This approach can be in principle used also for the Gromov-Wasserstein distance. However, the extreme computational burden for the calculation of this distance makes the permutation approach unfeasible in this case.

## 4.5 Discussion

By employing an object-oriented framework, we developed a novel approach for performing exploratory data analysis for any kind of network valued data object. Subsequently one can simply test the equality of networks and all data objects that can be put into the form of symmetric and positive definite matrices. Since it does not require matrix inversion, our method avoids the numerical problems encountered by the state-of-the-art approach described in [15]. As grounds for future work we plan to devise an approach to perform a post-hoc analysis, aimed at identifying which pairs of nodes are different among the groups, similarly to what is done in [15]. The testing method used in this work is easily generalized to Gromov-Wasserstein distances: further developments are in any case needed to speed up the computation of the distance, and thus be able to use computational intensive methods such as permutation testing.

We have used our proposed approach to analyze a complex data set with little information (admittedly, a pilot study). We thus have demonstrated how our approach works in practice, and that it can identify groups of subjects that may have significantly different functional brain networks even with a very coarse measurement network. However, given the complexity of the considered data objects, a larger sample size is required to obtain sufficient statistical evidence.

**Acknowledgements** The authors are very grateful to Greg Kiar and Eric Bridgeford from Neuro-Data at Johns Hopkins University, who graciously pre-processed the raw DTI and R-fMRI imaging data available at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html), using the pipelines ndmg and C-PAC. Moreover, the authors would like to thank the organizing committee of StartUp Research for the splendid management of such a beautiful event. Alessandra Cabassi and Matteo Fontana wish to thank Dr. Davide Pigoli and Prof. Piercesare Secchi for the fruitful discussions.

## References

1. Agosta, F., Sala, S., Valsasina, P., Meani, A., Canu, E., Magnani, G., Cappa, S.F., Scola, E., Quatto, P., Horsfield, M.A., Falini, A., Comi, G., Filippi, M.: Brain network connectivity assessed using graph theory in frontotemporal dementia. *Neurology* **81**(2), 134–143 (2013)
2. Arden, R., Chavez, R.S., Grazioplene, R., Jung, R.E.: Neuroimaging creativity: a psychometric view. *Behav. Brain Res.* **214**(2), 143–156 (2010)
3. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log Euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.* **56**(2), 411–421 (2006)
4. Belilovsky, E., Varoquaux, G., Blaschko, M. B.: Testing for differences in Gaussian graphical models: applications to brain connectivity. In: *Advances in Neural Information Processing Systems*, pp. 595–60 (2016)
5. Bonilha, L., Gleichgerrcht, E., Fridriksson, J., Rorden, C., Breedlove, J.L., Nesland, T., Paulus, W., Helms, G., Focke, N.K.: Reproducibility of the structural brain connectome derived from diffusion tensor imaging. *PloS one* **10**(9), e0135247 (2015)
6. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**(3), 186–198 (2009)
7. Cabassi, A., Pigoli, D., Secchi, P., Carter, P.A.: Permutation tests for the equality of covariance operators of functional data with applications to evolutionary biology. *Electron. J. Stat.* **11**(2), 3815–3840 (2017). <https://doi.org/10.1214/17-EJS1347>
8. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, **31**(3), 968–980 (2006)
9. Dryden, I.L., Koloydenko, A., Zhou, D.: Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Ann. Appl. Stat.* **3**(3), 1102–1123 (2009)
10. Durante, D., Dunson, D.B.: Bayesian inference and testing of group differences in brain networks. *Bayesian Anal.* **13**(1), 29–58 (2018)
11. Fornito, A., Zalesky, A., Breakspear, M.: Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage* **80**, 426–444 (2013)
12. Fréchet, M.: Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. l’Institut Henri Poincaré* **10**(3), 215–310 (1948)
13. Friedman, J., Hastie, T., Tibshirani, R.: Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9**(3), 432–441 (2008)
14. Friston, K.J.: Functional and effective connectivity in neuroimaging: a synthesis. *Hum. Brain Mapp.* **2**(1–2), 56–78 (1994)
15. Ginestet, C.E., Li, J., Balachandran, P., Rosenberg, S., Kolaczyk, E.D.: Hypothesis testing for network data in functional neuroimaging. *Ann. Appl. Stat.* **11**(2), 725–750 (2017)
16. GSell, M.G., Taylor, J., Tibshirani, R.: Adaptive testing for the graphical lasso. arXiv preprint (2013). [arXiv:1307.4765](https://arxiv.org/abs/1307.4765)
17. Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P.: Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci.* **106**(6), 2035–2040 (2009)
18. Jones, D.K., Knösche, T.R., Turner, R.: White matter integrity, fiber count, and other fallacies: the do’s and don’ts of diffusion MRI. *Neuroimage* **73**, 239–254 (2013)
19. Lee, S., Chugh, P.E., Shen, H., Eberle, R., Dittmer, D.P.: Poisson factor models with applications to non-normalized microRNA profiling. *Bioinformatics* **29**(9), 1105–1111 (2013)
20. Marron, J.S., Alonso, A.M.: Overview of object oriented data analysis. *Biometrical J.* **56**, 732–753 (2014)
21. Meinshausen, N., Bühlmann, P.: High-dimensional graphs and variable selection with the lasso. *Ann. Stat.* **34**(3), 1436–1462 (2006)
22. Mémoli, F.: The Gromov-Wasserstein distance: a brief overview. *Axioms* **3**(3), 335–341 (2014)
23. Mémoli, F.: Gromov-Wasserstein distances and the metric approach to object matching. *Found. Comput. Math.* **11**(4), 417–487 (2011)

24. Pigoli, D., Aston, J.A., Dryden, I.L., Secchi, P.: Distances and inference for covariance operators. *Biometrika* **101**(2), 409–422 (2014)
25. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**(3), 1059–1069 (2010)
26. Rykhlevskaia, E., Gratton, G., Fabiani, M.: Combining structural and functional neuroimaging data for studying brain connectivity: a review. *Psychophysiology* **45**(2), 173–187 (2008)
27. Scott, J.G., Kelly, R.C., Smith, M.A., Zhou, P., Kass, R.E.: False discovery rate regression: an application to neural synchrony detection in primary visual cortex. *J. Am. Stat. Assoc.* **110**(510), 459471 (2015)
28. Simpson, S.L., Hayasaka, S., Laurienti, P.J.: Exponential random graph modeling for complex brain networks. *PLoS one* **6**(5), e20039 (2011)
29. Simpson, S.L., Bowman, F.D., Laurienti, P.J.: Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Stat. Surv.* **7**, 1 (2013)
30. Stam, C.J.: Modern network science of neurological disorders. *Nat. Rev. Neurosci.* **15**(10), 683–695 (2014)
31. Stippich, C.: Clinical Functional MRI: Presurgical Functional Neuroimaging. Springer, Heidelberg (2015)
32. Stan Development Team. RStan: the R interface to Stan. R package version 2.17.2 (2017). <http://mc-stan.org/>
33. Wang, H., Marron, J.S.: Object oriented data analysis: sets of trees. *Ann. Stat.* **35**(5), 1849–1873 (2007)
34. Zalesky, A., Fornito, A., Bullmore, E.T.: Network-based statistic: identifying differences in brain networks. *Neuroimage* **53**(4), 1197–1207 (2010)

# An Object Oriented Approach to Multimodal Imaging Data in Neuroscience



Andrea Cappozzo, Federico Ferraccioli, Marco Stefanucci  
and Piercesare Secchi

**Abstract** We propose a methodological framework for exploring complex multimodal imaging data from a neuroscience study with the aim of identifying a data-driven group structure in the patients sample, possibly connected with the presence/absence of lifetime mental disorder. The functional covariances of fMRI signals are first considered as data objects. Appropriate clustering procedures and low dimensional representations are proposed. For inference, a Frechet estimator of both the covariance operator itself and the average covariance operator is used. A permutation procedure to test the equality of the covariance operators between two groups is also considered. We finally propose a method to incorporate spatial dependencies between different brain regions, merging the information from both the Structural Networks and the Dynamic functional activity.

**Keywords** Data objects · Functional data analysis · Principal components  
Multimodal Imaging · Neuroscience

## 1 Introduction

The following work arises from the StartUp Research experience, a workshop held at Certosa di Pontignano on June 25–27 2017. Seven groups formed by early-career

---

A. Cappozzo  
Department of Statistics and Quantitative Methods,  
University of Milano-Bicocca, Milan, Italy  
e-mail: [andrea.cappozzo@unimib.it](mailto:andrea.cappozzo@unimib.it)

F. Ferraccioli (✉)  
Department of Statistical Sciences, University of Padova, Padua, Italy  
e-mail: [ferraccioli@stat.unipd.it](mailto:ferraccioli@stat.unipd.it)

M. Stefanucci  
Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy  
e-mail: [marco.stefanucci@uniroma1.it](mailto:marco.stefanucci@uniroma1.it)

P. Secchi  
MOX Department of Mathematics, Politecnico di Milano, Milan, Italy  
e-mail: [piercesare.secchi@polimi.it](mailto:piercesare.secchi@polimi.it)

researchers and a senior mentor acting as group leader were challenged to develop novel methods for analysing a common dataset.

Both researchers and practitioners involved in the field of data analysis are nowadays increasingly challenged in confronting with data structures that lie outside the classical Euclidean framework. That is, thanks to the technological advancements of measurement machineries, not only datasets are becoming massive in terms of size (the way-too-exploited buzzword big data is a living proof of the concept) but also substantial in terms of data complexity. As a consequence, statisticians are encouraged to sharpen their mathematical and programming skills for tackling the enormous knowledge-discovery opportunities that lie within these complex datasets. Object oriented data analysis (OODA) is a framework, firstly introduced in [24], for approaching data challenges in which the *object* of the analysis (i.e., the observation or statistical unit) possesses distinctive features that would not be exploited by performing a classical multivariate analysis after data dimension reduction. Examples of data objects that are considered by OODA include (but are not limited to) curves, images, tree structured data and positive semi-definite matrices [14]. In such a context mathematics plays a fundamental role in rigorously defining the embedding space and properties of the objects under study, and consequently fostering the development of new statistical methodologies. Two central notions are the base-ground for understanding the conceptual framework of OODA:

- *Object Space*: is the set in which the mathematical representation of the data lie. For example, the employed object space for the dynamic functional activity (see Sect. 2) is the Hilbert space  $\mathbb{L}^2$  of square-integrable functions.
- *Feature Space*: is the set of features that numerically represent the data object. The feature space for the scan-rescan dynamic functional activity of the 24 subjects in the study (see Sect. 2) is a digitized  $70 \times 404 \times 24 \times 2$  array.

The OODA framework is particularly appropriate when applied to neuroscience, where the large use of Magnetic Resonance Imaging (MRI) in the study of brain connectivity and activity has recently created new challenges for statisticians. The nature and complexity of data coming from electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and diffusion tensor imaging (DTI) have favoured the development of ad-hoc methodologies greatly expanding the statistical neuroscience literature [8, 18]. During the StartUp Research workshop our group attempted to analyse the provided dataset employing mathematical tools coming from OODA, with the aim of exploring the connectivity structure within subject brains and across groups of subjects with different traits in order to identify possible meaningful and significant patterns.

The data comes from a pilot study of the Enhanced Nathan Kline Institute-Rockland Sample project; it comprises multimodal imaging data and subject-specific covariates for  $n = 24$  subjects, for 12 of which 2 scan-rescan imaging sessions are available. A detailed description of the project, scopes, and technical aspects can be found at [http://fcon\\_1000.projects.nitrc.org/indi/enhanced/](http://fcon_1000.projects.nitrc.org/indi/enhanced/). The pilot study includes three data sources:

- **Structural networks:** These data measure the anatomical interconnections—made by white matter fibers—among brain regions of interest, and are collected from DTI.
- **Dynamic functional activity:** These data measure the dynamic activity of each brain region through changes in the blood-oxygen-level dependent (BOLD) signal during resting state fMRI (R-fMRI) sessions.
- **Functional networks:** These data measure synchronization in brain activity for each pair of brain regions, and are obtained from the correlation in dynamic functional activity.

Some missing data are present in the dataset: the Dynamic functional activity for 2 subjects and the Structural networks for 4 subjects were not collected. Additionally, subject-specific information related to age, whether she/he is left-handed, right-handed or ambidextrous and her/his current and lifetime mental disorder were available only for 20 samples, impacting the performance evaluation of the method proposed in Sect. 7.

In Sect. 2 the necessary framework is introduced and Functional Data Analysis methods [19] are employed for obtaining the main data object of our analysis: a set of 22 functional networks numerically represented as correlation matrices. Subsequently, a proper distance metric for the aforementioned objects is considered for performing cluster analysis, as reported in Sect. 3. Section 4 considers a low dimensional representation of the data objects, and comparison with the results obtained by the clustering method is addressed. Section 5 reports a formal permutation procedure to test the equality of the mean functional networks between the two groups determined in Sect. 3. In order to identify possible different sources of variation a thorough study of the eigenstructure for the two mean functional networks is reported in Sect. 6. Section 7 considers a possible solution to account for spatial dependence between Dynamic functional activity of different regions, performing data fusion for the subset of subjects for which both Structural networks and Dynamic functional activity are available.

## 2 Curves and Correlation Matrices as Data Objects

Let us first consider the fMRI signal from the first scan. The data consist of 70 signals for each of the 24 subjects, corresponding to the BOLD activity of the 70 brain regions described by the Desikan Atlas [5]. Over the past decades the number of fMRI studies has increased exponentially [20], fostering the development of several methods for the analysis and interpretation of resting-state fMRI data, such as seed-based correlation analysis, independent component analysis and network-based models [4]. We propose to employ a Functional Data Analysis approach for performing the analysis, considering each signal as a realization of a stochastic process  $X(t_i)$  sampled at times  $t_i$ , where  $i = 1, \dots, 403$ ; the last instant of time was not recorded for several patients and therefore it was not considered in the analysis. Subjects are

sampled at the same time schedule, so that registration is not deemed to be necessary [19]. Two subjects are not considered in the following analysis because of missing data, namely patient with ID 1 and patient with ID 21.

When dealing with functional data the usual starting point is to represent the data observed on a finite grid of points as functions. This part of the analysis is called smoothing, and there are several approaches to do it. Two important classes of smoothers are represented by kernel smoothing and orthogonal basis [19]. Both approaches sharing the idea of filtering out the short-time variation while keeping the global shape of the signal: we employ the latter for pre-processing the fMRI data. Orthogonal basis smoothing relies on the fact that, given an orthogonal basis for the space of interest, every function can be represented as an infinite linear combination of bases. A truncated version of the infinite sum provides a continuous representation of the discrete signal and reduces the dimensionality of the problem. The following analyses are based on a Fourier expansion, a standard choice in signal processing literature, with 100 bases. As it happens, it is not clear whether the short-time oscillations can be treated as noise or they might be related to some specific conditions of the brain. Future work might consider more appropriate bases such as wavelets [13] or Hierarchical Component Analysis [23]. An example of a smoothed function and its residuals for a given brain area and subject is reported in Fig. 1. A first interesting question that arises from the smoothing process would be to understand whether the residuals of the smoothing have some kind of clinical interpretation. We now have 70 functions for each of the 22 subjects, each function related to a different brain region. We used these functions to construct a correlation matrix between regions for each of the subjects. More in detail, we can compute the correlation between pairs of functions for every subject

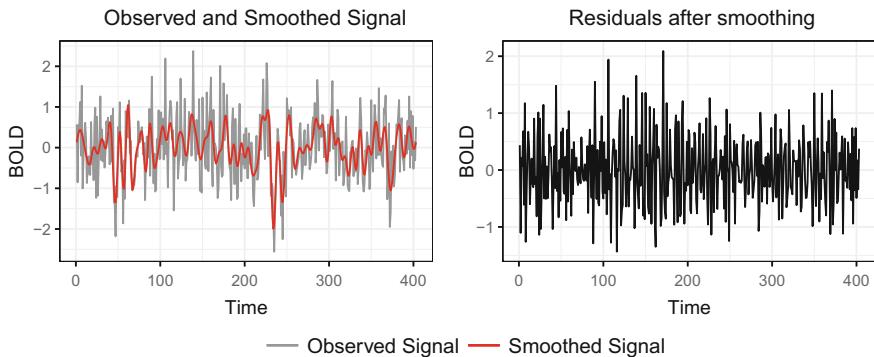
$$\text{Cor}(f_i, f_j) = \frac{\langle f_i, f_j \rangle}{\|f_i\| \|f_j\|} = \cos(\theta_{ij}), \quad \text{for } i, j = 1, \dots, 70. \quad (1)$$

Here  $\langle \cdot, \cdot \rangle$  denotes the inner product

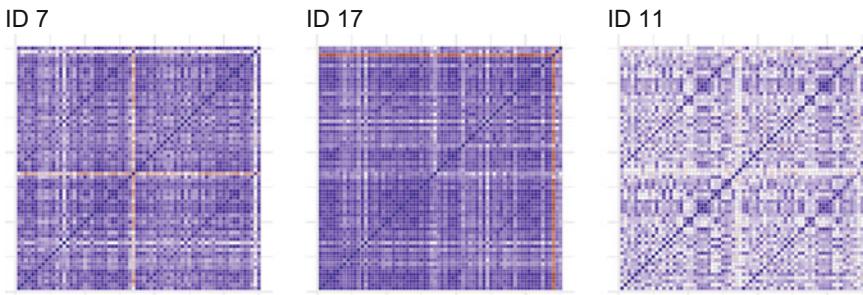
$$\langle f_i, f_j \rangle = \int_{\Omega} f_i f_j d\mu \quad (2)$$

on the Hilbert space  $\mathbb{L}^2(\Omega, \mathcal{B}, \mu)$ , where  $\Omega = [0, 403]$ ,  $\mathcal{B}$  is the Borel  $\sigma$ -algebra of  $[0, 403]$  and  $\mu$  the Lebesgue measure. The norm  $\|\cdot\| = \langle \cdot, \cdot \rangle^{1/2}$  is induced by the inner product in (2). For a more detailed treatment of the underlying Hilbert space theory for functional data analysis, see [9, 10]. Processing the functional signal through the operator defined in (1) results in a  $70 \times 70$  correlation matrix for each of the 22 subjects in the study. In Fig. 2 a subset of the so computed correlations matrices are graphically represented as heatmaps.

There is a clear difference in terms of correlation magnitude amongst subjects. Particularly, it seems that a subgroup of patients (ID 7, 10, 15, 17, 20, 22 and 23) present a much higher positive correlation between brain regions than the ones recorded for the rest of the subjects, visible by the overall darker blue areas in the correlation



**Fig. 1** Observed and smoothed signals (left plot) and residuals after the Fourier basis approximations (right plot) for a given subject and brain area



**Fig. 2** Heatmaps of the first scan fMRI signal correlation matrices for 3 patients in the study, each belonging to a different subgroup identified by visual exploration of the magnitudes in the matrices

plots. Another interesting pattern visible in some patients (ID 6, 12, 14, 17 and 19 primarily) is given by the presence of a specific brain region, namely *rh-frontalpole*, that is negatively correlated to the remaining areas identified by the Desikan Atlas. A third behaviour that emerges from the visual exploration of the plots in Fig. 2 is the mild negative correlation and/or almost absence of correlation for two brain areas with the others for some subjects (ID 7, 8, 11, 14, 15, and 23). Particularly, these two areas are *lh-frontalpole* and *lh-temporalpole*. Lastly, there are two subjects (ID 2 and 23) that present almost individual patterns in the correlation structure between brain regions.

### 3 Clustering of Functional Networks

It is of interest to verify the presence of groups of subjects with similar brain activity, employing appropriate statistical methods given the complex structure of the objects under analysis. That is, the aim is to define a suitable distance concept in order to

characterize proximity amongst objects and subsequently perform cluster analysis according to the provided metric.

Given the considered context we cannot embed our objects of interest, i.e., the aforementioned correlation matrices, in a classical Euclidean space. Particularly, the correlation matrices represented in Fig. 2 are finite-dimensional approximations of a rescaled covariance operator for functional random processes, and therefore a suitable inference framework must be considered. Given a random function  $f$  taking values in  $\mathbb{L}^2(\Omega)$  we define the covariance operator  $C_f$  for  $g \in L^2(\Omega)$ :

$$C_f g(t) = \int_{\Omega} E([f(t') - E(f(t'))][f(t) - E(f(t))]g(t')dt'. \quad (3)$$

For a review of definitions and theoretical properties of operators on  $L^2(\Omega)$  see [2].

Denote with  $PD(p)$  the space of positive semi-definite symmetric matrices of dimension  $p$ , that is the set of real symmetric matrices having non-negative eigenvalues [1]. We recall that  $PD(p)$  is not a vector space and an inner product is not defined; it is however a Riemannian manifold in which we can define a distance. For a detailed list of non-Euclidean distances for covariance matrices, see for example [6]. However, in a context of functional data, infinite dimensional extension of metrics for positive-semidefinite matrices must be used. Employing the inferential framework for covariance operators introduced in [17] we are able to extend the matrix-based distances to the functional case.

With the aim of measuring synchronization in brain activity and their respective dissimilarities among patients we consider the functional extension of the square root distance between variance covariance matrices, firstly defined in [6]. That is, given two covariance operators  $S_1$  and  $S_2$  their square root distance is defined as

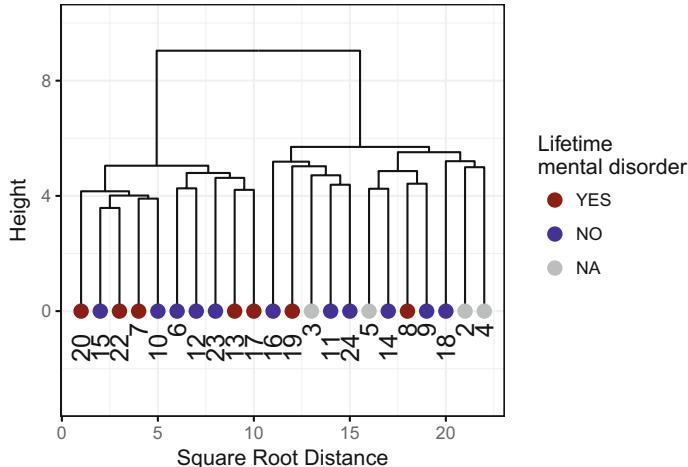
$$d_R(S_1, S_2) = ||S_1^{1/2} - S_2^{1/2}||_{HS} \quad (4)$$

where  $|| \cdot ||_{HS}$  denotes the Hilbert-Schmidt norm, generalization of the Frobenius norm for finite-dimensional matrices. Among the available matrix-based distances extendable to the functional case we decided to consider (4) since it takes into account the full eigenstructure of the covariance operator [17]. The definition of a proper distance is directly linked to the introduction of a mean value concept, given the chosen distance. Particularly, letting  $S_1, \dots, S_n$  be a sample of independent covariance operators we define its sample *Fréchet mean* based on the square root distance (4) as

$$\hat{\Sigma} = \hat{\Delta} \hat{\Delta}^\top \quad (5)$$

where

$$\hat{\Delta} = \underset{\Delta}{\operatorname{arginf}} \left\{ \sum_{i=1}^n ||S_i^{1/2} - \Delta||_{HS}^2 \right\} = \frac{1}{n} \sum_{i=1}^n S_i^{1/2}. \quad (6)$$



**Fig. 3** Dendrogram from hierarchical clustering with Ward agglomeration method to the functional network data. The dendrogram highlights the presence of two main clusters not seemingly related to patients' mental health status

For proofs and discussion related to the consistency of the sample Fréchet mean based on the square root distance, refer to [11].

Making use of the square root distance defined in (4) we proceed in trying to identify possible presence of groups amongst the data objects employing a distance-based clustering algorithm. Particularly, the analysis was carried out considering hierarchical clustering with Ward agglomeration method [15].

The result of the clustering algorithm is graphically presented in Fig. 3: the dendrogram clearly highlights the presence of two different clusters in our sample of patients. The groups however do not seem to be separated along the additional information on the subjects provided in the study. Therefore, even though the difference between the mean correlation matrices of the two groups results to be statistically significant (see Sect. 5), interpretation explaining the groupings remains still unclear. A clinician assessment, together with a thorough consideration of the medical history of each patient involved in the study would provide insight on groups interpretability and classification.

In the upcoming section, the problem of finding homogeneous groups amongst functional networks is differently tackled employing a non-linear dimensionality reduction technique. Both methodologies agree in terms of identified number of groups and groups structure.

## 4 Low Dimensional Representation

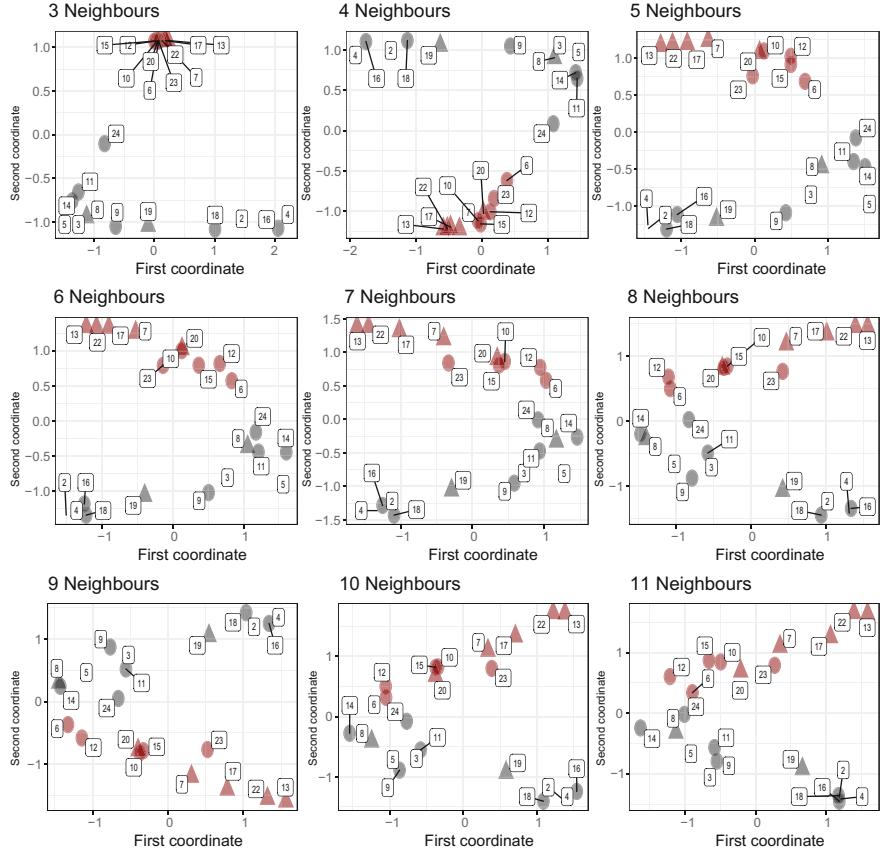
In order to obtain a low dimensional representation of the correlation matrices a Local Linear Embedding (LLE) algorithm [21] is considered. This method is based on a simple geometric intuition. Suppose the data consist of  $N$  real-valued vectors  $X_i$ , each of dimensionality  $D$ , sampled from some smooth underlying manifold. We expect each data point and its neighbours to lie on or close to a locally linear patch of the manifold. We can characterize the local geometry of these patches by linear coefficients that reconstruct each data point from its neighbours. In the first step of the algorithm one identifies  $K$  nearest neighbours per data point, as measured by Euclidean distance. In the second step the weights  $W_{ij}$  that best reconstruct each data point  $X_i$  from its neighbours are computed, minimizing

$$\sum_{i=1}^n \left( X_i - \sum_{j=1}^K W_{ij} X_j \right)^2.$$

The weights  $W_{ij}$  summarize the contribution of the  $j$ -th data point to the  $i$ -th reconstruction. Finally we can compute the vectors  $Y_i$  of low dimensional coordinates,  $d < n$ , best reconstructed by the weights  $W_{ij}$ , minimizing

$$\sum_{i=1}^d \left( Y_i - \sum_{j=1}^K W_{ij} X_j \right)^2.$$

This cost function—like the previous one—is based on locally linear reconstruction errors, but here we fix the weights  $W_{ij}$  while optimizing the coordinates  $Y_i$ . In Fig. 4 we can see the two dimensional representations ( $d = 2$ ), for different number of neighbours (from 3 to 11, starting from the left upper corner). The triangles represent the patients with lifetime disease, while the colour represents the groups identified by hierarchical clustering. We firstly note that the algorithm is robust with respect to the choice of the hyper-parameter  $K$ . Secondly, and more relevant for the scope of our analysis, we recognize that in all the considered representations 5 out of the 7 subjects with lifetime disease are in the red group and the remaining 2 in the black group (these are patients labelled with ID 8 and 19 respectively). We can also note that the low dimensional representation preserves the structure of the original space and the separation performed by the hierarchical clustering is still clearly visible: an average Adjusted Rand Index of 0.96 between the groupings found with the two methods, varying  $K$  from 3 to 11 in the LLE, is obtained. A formal permutation test for statistically assessing the significant difference between the two sub-populations identified by both hierarchical clustering and LLE is developed in the upcoming section.



**Fig. 4** 2-dimensional representation of the correlation matrices through Local Linear Embedding algorithm for different number of neighbours  $K$ . Triangles represent patients with lifetime disease; colours represent the groups identified by hierarchical clustering of Sect. 3

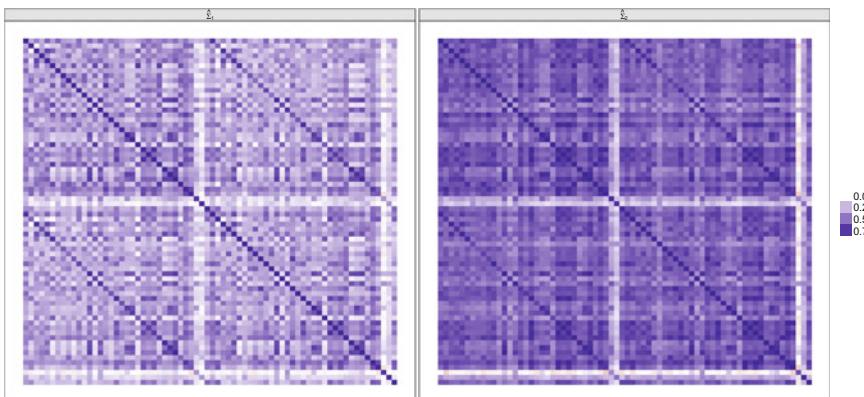
## 5 Hypothesis Testing for Correlation Structures

Let us consider the two groups of patients identified in the previous Sections. We want to verify whether the functional activity, recorded in terms of  $70 \times 70$  correlation matrices for each of the 22 subjects, is significantly different in the 2 groups. We assume that our two samples are such that  $S_1^{(1)} \dots S_{n_1}^{(1)}$  are random  $PD(p)$  matrices with expectation  $E(S_i) = \Sigma_1$ ,  $i = 1, \dots, n_1$  and  $S_1^{(2)} \dots S_{n_2}^{(2)}$  are random  $PD(p)$  matrices with expectation  $E(S_j) = \Sigma_2$ ,  $j = 1, \dots, n_2$ .  $S_1^{(1)} \dots S_{n_1}^{(1)}$  and  $S_1^{(2)} \dots S_{n_2}^{(2)}$  are the sample correlation matrices belonging to the first and second group respectively. Particularly, in our context  $n_1 = 12$  and  $n_2 = 10$  with patients

(2, 3, 4, 5, 8, 9, 11, 14, 16, 18, 19, 24) belonging to the first group and patients (6, 7, 10, 12, 13, 15, 17, 20, 22, 23) to the second group (see Fig. 3). We would like to test

$$H_0 : \Sigma_1 = \Sigma_2 \text{ versus } H_1 : \Sigma_1 \neq \Sigma_2.$$

To test these hypotheses we follow a permutational approach along the methods advanced in [16, 17]. We reformulate the test in terms of square root distances between covariance objects: the considered test statistic is  $d(\hat{\Sigma}_1, \hat{\Sigma}_2)$  where  $\hat{\Sigma}_1$  and  $\hat{\Sigma}_2$  denotes the Fréchet mean as defined in (6) for the samples in the two groups.  $H_0$  is rejected for large values of  $d(\hat{\Sigma}_1, \hat{\Sigma}_2)$ . The test is simply a two way ANOVA, but equipped with a proper metric and consequently with a proper definition of sample mean. If  $H_0$  is true, complete exchangeability of the random variables generating the sample observations holds and therefore, in order to approximate the distribution of the test statistic under  $H_0$ , the two samples are pooled together and randomly assigned to the two groups preserving sample sizes. The test consists in a comparison of  $d(\hat{\Sigma}_1, \hat{\Sigma}_2)$  with  $M$  random permutations computed via Monte Carlo of  $d(\hat{\Sigma}_1^{(m)}, \hat{\Sigma}_2^{(m)})$ ,  $m = 1, \dots, M$ ; where  $\hat{\Sigma}_i^{(m)}$  is the sample mean correlation matrix for group  $i$  in permutation  $m$ . The  $p$ -value with  $M = 100$  permutations is less than 0.01, with a difference of the two sample means of 2.41. Thus, we conclude that the two sub-populations have statistically different correlation matrices, confirming and validating the results previously highlighted by the clustering and LLE methods. The same permutation test had been initially applied to groups clustered by subjects characteristics; notwithstanding, none of the additional information available for the subjects under study (age, handedness, current/lifetime mental disorder) have been proved significant in distinguishing different groups. Figure 5 shows the heatmaps of the sample mean correlation matrices of the two considered groups. The difference between the two is clear, with higher correlation values in the second group.



**Fig. 5** Heatmap of the sample mean correlation matrices in the two identified groups

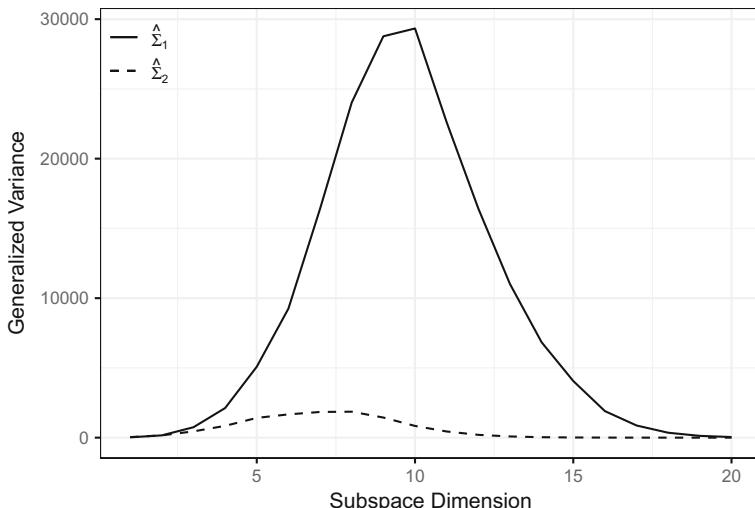
## 6 Eingenstructure of the Mean Correlation Matrices

In this Section a comprehensive analysis of the main sources of variability for the two sample mean correlation matrices identified in Sect. 3 will be addressed. Figure 6 displays the generalized variance in subspaces of increasing dimension for the two correlation matrices, defined as the cumulative product of their eigenvalues [7].

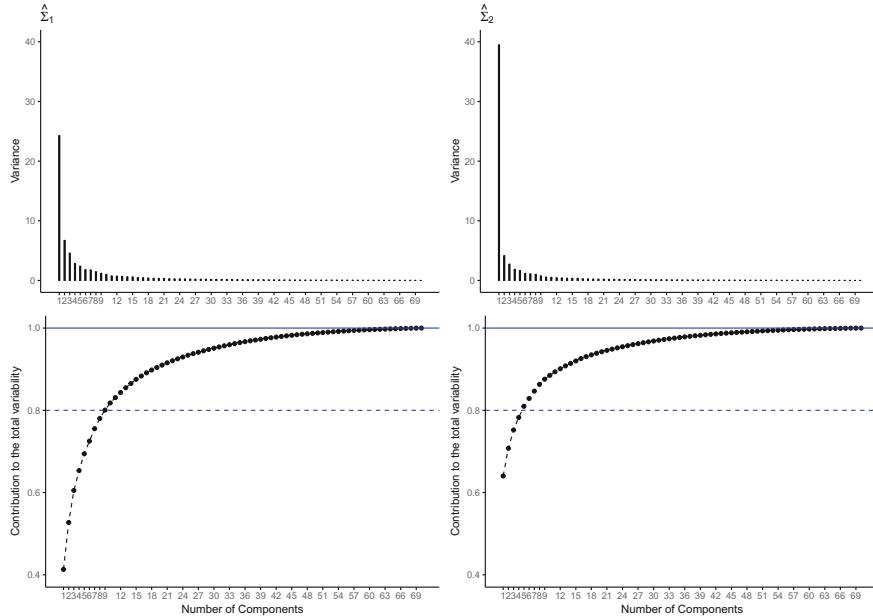
The generalized variance is proportional to the square of the volumes of the hyper-ellipsoids projected onto the principal components subspaces. It is clearly visible that the first group is characterized by a much larger generalized variance, supporting the significant difference between the two groups highlighted by the permutation test.

A spectral decomposition of the two sample mean correlation matrices is reported in Fig. 7. Since we are considering correlation matrices, the employed terminology comes from the Principal Components Analysis literature [12]. Particularly, the variance denotes the magnitude of the different eigenvalues whereas the contribution to the total variability is calculated dividing the cumulative sum of the eigenvalues by their total. The magnitude of the eigenvalues in the first group decreases more slowly than in the second group, as it was already apparent in Fig. 6. Five components account for 80% of the total variability in  $\hat{\Sigma}_2$ , whereas for  $\hat{\Sigma}_1$  nine components are necessary for achieving the same contribution.

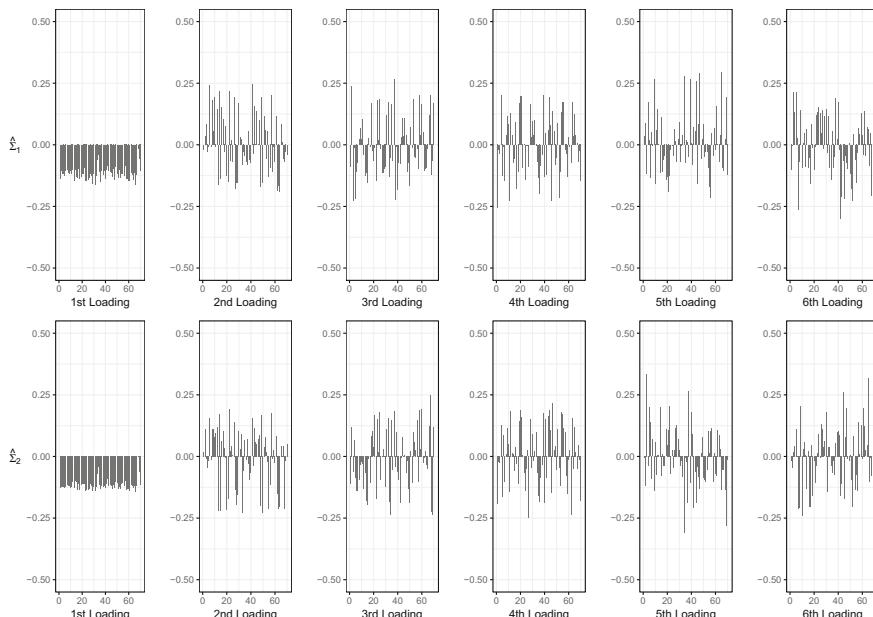
In order to check whether the source of variability is different in the two groups, the components (loadings) of the first six eigenvectors for the two sample mean covariance matrices are plotted in Fig. 8. As it can be seen from the graphs, the source of variability seems different, especially considering the first three loadings. This is further highlighted by the graphical representation of the 3-D spatial coordinates for



**Fig. 6** Generalized variance for the two sample mean correlation matrices, defined as the cumulative product of their eigenvalues



**Fig. 7** Eigenvalues and relative cumulative sum of eigenvalues for the mean correlation matrices of the two groups identified in the patients sample



**Fig. 8** Entries of the first 6 loadings vectors (eigenvectors) for the two sample mean correlation matrices of the two groups identified in the patients sample

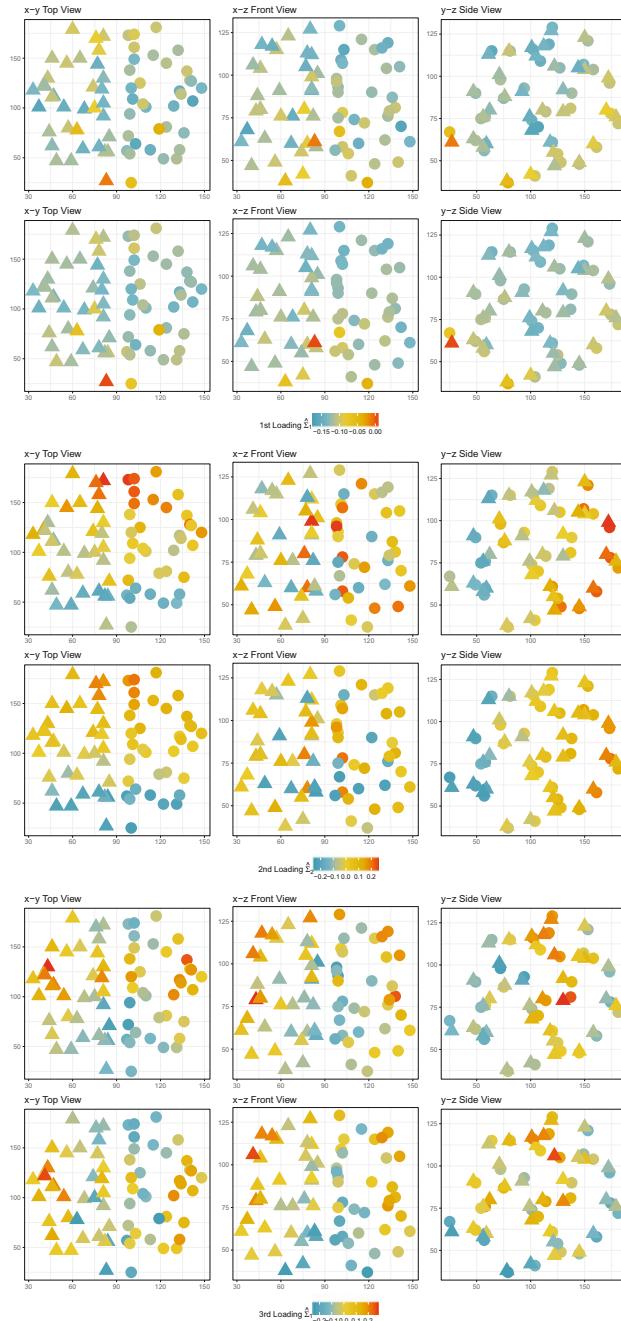
the centroids of the brain regions reported in Fig. 9, where centroids are coloured according to the first, second and third loading vectors entries respectively. A spatial pattern seems to be present in the matrices eigenstructure. The present analysis motivates and justifies the novel approach introduced in the upcoming section, where we attempt to account for the spatial dependence employing a re-weighted version of the functional network. Particularly, the structural network (i.e., the count of number of white fibers that connect each brain region) is interpreted as a measure of proximity between the brain regions.

## 7 Spatial Dependence for Functional Networks

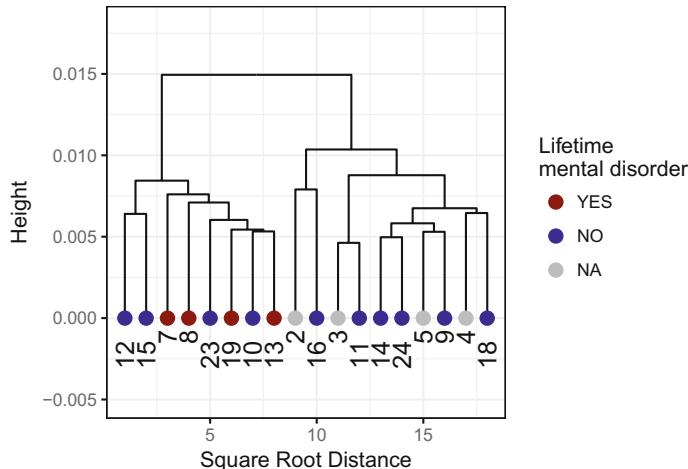
So far, we have only considered the fMRI data corresponding to the 70 regions of the Desikan atlas as independent. Nonetheless, the spatial dependence has not been filtered out during pre-processing and it is therefore reasonable to suppose that some sort of spatial dependence is still present in the registered signals, as it can be graphically seen in Fig. 9. A possible procedure for incorporating the spatial dependence within our analysis framework would be to exploit the information contained in the structural networks available for each patient. The structural networks contain the total number of white matter fibers connecting each pair of brain regions for each subject. The aforementioned structure can be interpreted as an adjacency matrix: the intuition behind this definition is that the more white fibers connecting a pair of brain regions the closer the two brain regions can be considered. Particularly, the functional networks (identified by the correlation matrices employed in the previous Sections) can be re-weighted according to the magnitude enclosed in the structural networks, subject-wise. Considering only the first scan, indicate with  $d_{uv}$  the count of how many white matter fibers are found to connect brain regions  $u$  and  $v$  for a specific subject. We define the symmetric  $70 \times 70$  weight matrix  $W$  induced by the structural network for each subject having entries as follows:

$$w_{uv} = \begin{cases} 1 & u = v \\ d_{uv} / \left( \sum_{u=1}^{70} \sum_{v=1}^{70} d_{uv} \right) & u \neq v \end{cases} \quad (7)$$

Subsequently, we define the *re-weighted functional networks*  $R$  as the Hadamard product between  $W$  and the functional networks computed in (1). For obtaining  $R$  both structural and functional networks must be available, therefore it was possible to calculate the re-weighted functional networks only for 18 out of 24 patients present in the study. Notice that  $R$  is still a symmetric and positive semi-definite matrix thanks to Schur product theorem [22]. Employing the same methodology described in Sect. 3 we perform hierarchical clustering on the re-weighted functional networks: the dendrogram of the clustering procedure is reported in Fig. 10. Likewise in the previous analysis the dendrogram highlights the presence of two different clusters, with a significant difference in their mean correlation matrices (the test in Sect. 5 was



**Fig. 9** Graphical representation of the 3-D spatial coordinates for the centroids of the brain regions, under different 2-D views. Colour intensity is associated with the entries of the first (top), second (middle) and third (bottom) loading of  $\hat{\Sigma}_1$  (first rows) and  $\hat{\Sigma}_2$  (second rows) respectively. Shapes describe hemisphere membership



**Fig. 10** Dendrogram from hierarchical clustering with Ward agglomeration method to the re-weighted functional network data. The dendrogram highlights the presence of two main clusters, partially related to patients' mental health status

repeated and the null hypothesis rejected). In addition, the two identified groups seem at least partially related to the presence or absence of lifetime mental disorder for the available set of patients. Although the sample size is very small, we empirically evaluate the main source of dissimilarity between the two groups of patients with and without lifetime mental disorder considering their mean re-weighted functional networks, computed using (5). The preeminent differences are due to the higher weighted correlations found for patients with lifetime mental disorder between areas *lh - posteriorcingulate* and *lh - corpuscallosum*, *rh - posteriorcingulate* and *lh - posteriorcingulate*, *rh - superiorfrontal* and *rh - caudalmiddlefrontal*, *rh - corpuscallosum* and *lh - posteriorcingulate*, compared to the weighted correlations in these areas for patients with absence of lifetime mental disorder.

## 8 Conclusions and Future Research Directions

The present work is the result of a 48 h workshop during which the authors, guided by their senior group leader Piercesare Secchi, were asked to propose original statistical methods for data analysis in neuroscience [3]. We applied several techniques from Object Oriented Data Analysis literature for exploring data coming from the Enhanced Nathan Kline Institute-Rockland Sample project. Three different clustering methods are proposed for the fMRI data, with the last and most promising one involving the processing of both structural and functional network for each patient.

Our approach began with the identification of two clusters in the space of the correlation matrices of the smoothed fMRI signals. These two groups corresponded only partially to the labelling concerning the presence/absence of mental disease. A non-linear dimensional reduction technique helped us to visualize the clusters: the two sub-populations structure is clear in the identified subspace. The difference between the two groups, formalized through a statistical test in which the null hypothesis was the equality of the two mean correlation matrices, is highly significant. A deeper analysis of the eigenstructure of the two mean correlation matrices highlighted the differences in the sources of variability in the two groups, together with a possible spatial dependence in the data objects. Lastly, an attempt at performing data fusion weighting the functional networks with the structural networks is addressed: promising initial results seem to have been achieved. In particular, employing re-weighted functional networks, subjects with confirmed presence of mental disease are more clearly separated from patients with absence of mental disease, fostering the employment of the aforementioned procedure whenever functional and structural networks are available. Nevertheless, both a larger sample size as well as knowledge domain would be necessary for establishing and interpreting the described discoveries.

The StartUp Research workshop has been a challenging yet enriching and unforgettable experience, in which we had the chance to meet, connect and learn from our peers, colleagues and senior mentors. We early-career researchers had a direct experience on the essential importance of interaction and knowledge-sharing which, ultimately, lead to knowledge creation.

**Acknowledgements** We acknowledge Greg Kiar and Eric Bridgeford from NeuroData at Johns Hopkins University, who pre-processed the raw DTI and R-fMRI imaging data available at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html). We would like to deeply thank the StartUp Research Scientific Committee for efficiently and flawlessly organizing such a motivating experience. We thank Professor Francesca Greselin and Doctor Mauro Ceroni for their support and help throughout the drafting of this manuscript.

## References

1. Amari, S.I.: Differential-geometrical methods in statistics. Lecture Notes in Statistics, vol. 28. Springer, New York (1985)
2. Bosq, D.: Linear Processes in Function Spaces. Lecture Notes in Statistics. Springer, New York (2000)
3. Canale, A., Durante, D., Paci, L., Scarpa, B.: Connecting statistical brains. *Significance* **15**(1), 38–40 (2018)
4. Cole, D.M., Smith, S.M., Beckmann, C.F.: Advances and pitfalls in the analysis and interpretation of resting-state fMRI data. *Front. Syst. Neurosci.* **4**, 8 (2010)
5. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**(3), 968–980 (2006)
6. Dryden, I.L., Koloydenko, A., Zhou, D.: Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Ann. Appl. Stat.* **3**(3), 1102–1123 (2009)

7. Friendly, M., Monette, G., Fox, J.: Elliptical insights: understanding statistical methods through elliptical geometry. *Stat. Sci.* **28**(1), 1–39 (2013)
8. Heuvel, M.P.V.D., Pol, H.E.H.: Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* **20**(8), 519–534 (2010)
9. Horváth, L., Kokoszka, P.: Inference for Functional Data with Applications. Springer Series in Statistics. Springer, New York (2012)
10. Hsing, T., Eubank, R.: Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators. Wiley Series in Probability and Statistics. Wiley, Chichester, UK (2015)
11. Huckemann, S.: Intrinsic inference on the mean geodesic of planar shapes and tree discrimination by leaf growth. *Ann. Stat.* **39**(2), 1098–1124 (2010)
12. Jolliffe, I.T.: Principal Component Analysis and Factor Analysis, pp. 115–128. Springer, New York (1986)
13. Mallat, S.: A Wavelet Tour of Signal Processing. Academic Press (1999)
14. Marron, J.S., Alonso, A.M.: Overview of object oriented data analysis. *Biom. J.* **56**(5), 732–753 (2014)
15. Murtagh, F., Legendre, P.: Ward’s hierarchical clustering method: clustering criterion and agglomerative algorithm. *J. Classif.* **31**(3), 274–295 (2011)
16. Pesarin, F., Salmaso, L.: Permutation Tests for Complex Data: Theory, Applications and Software. Wiley (2010)
17. Pigoli, D., Aston, J.A.D., Dryden, I.L., Secchi, P.: Distances and inference for covariance operators. *Biometrika* **101**(2), 409–422 (2014)
18. Plis, S., Meinecke, F.C., Eicheler, T.: Analysis of multimodal neuroimaging data. *IEEE Rev. Biomed. Eng.* **4**, 26–58 (2011)
19. Ramsay, J., Silverman, B.W.: Functional Data Analysis. Springer Series in Statistics. Springer, New York (2005)
20. Roalf, D., Gur, R.: Functional brain imaging in neuropsychology over the past 25 years. *Neuropsychology* **31**(8), 954–971 (2017)
21. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. *Science* **290**(5500), 2323–2326 (2000)
22. Schur, J.: Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. *J. Reine Angew. Math.* **140**, 1–28 (1911)
23. Secchi, P., Vantini, S., Zanini, P.: Hierarchical independent component analysis: a multi-resolution non-orthogonal data-driven basis. *Comput. Stat. Data Anal.* **95**, 133–149 (2016)
24. Wang, H., Marron, J.S.: Object oriented data analysis: sets of trees. *Ann. Stat.* **35**(5), 1849–1873 (2007)

# Curve Clustering for Brain Functional Activity and Synchronization



**Gaia Bertarelli, Alice Corbella, Jacopo Di Iorio, Anastasia Gorshechnikova and Marian Scott**

**Abstract** Functional Magnetic Resonance Imaging (fMRI) has become one of the leading methods for brain mapping in neuroscience and it is an important tool in modern neuroscience investigation. Moreover, the recent advances in fMRI analysis are widely used to define the default state of brain activity, functional connectivity and basal activity. Signal processing schemes have been suggested to analyze the resting state Blood-Oxygenation-Level-Dependent (BOLD) signal from simple correlations to spectral decomposition. Our goal is to determine which brain areas behave similarly in the time domain. To address this question, we apply functional curve clustering methods. We carry out an exploratory study using classical functional clustering of fMRI time series. The analysis confirms the hypothesis of a possible spatial influence on the results and therefore suggests the development of spatial curve clustering methods for brain data.

**Keywords** Curve clustering · fMRI · Functional boxplot · Smoothing

---

G. Bertarelli (✉)

Department of Economics and Management, University of Pisa, Pisa, Italy  
e-mail: [gaia.bertarelli@ec.unipi.it](mailto:gaia.bertarelli@ec.unipi.it)

A. Corbella

MRC Biostatistics Unit, School of Clinical Medicine,  
University of Cambridge, Cambridge, UK  
e-mail: [alice.corbella@mrc-bsu.cam.ac.uk](mailto:alice.corbella@mrc-bsu.cam.ac.uk)

J. Di Iorio

MOX, Department of Mathematics, Politecnico di Milano, Milan, Italy  
e-mail: [jacopo.diiorio@polimi.it](mailto:jacopo.diiorio@polimi.it)

A. Gorshechnikova

Department of Statistical Sciences, University of Padova, Padua, Italy  
e-mail: [gorshechnikova@stat.unipd.it](mailto:gorshechnikova@stat.unipd.it)

M. Scott

School of Mathematics and Statistics, University of Glasgow, Glasgow, UK  
e-mail: [Marian.Scott@glasgow.ac.uk](mailto:Marian.Scott@glasgow.ac.uk)

## 1 Introduction

Over the past two decades the study of human cognition has greatly benefited from innovations in magnetic resonance imaging, such as the development of techniques that aim at detecting physiological markers of neural activity. Functional Magnetic Resonance Imaging (fMRI) is a tool for studying brain function, that is the relation between a neural substrate and a particular behaviour or condition. The fMRI measurement is related to the amount of de-oxygenated haemoglobin in the blood and it measures the so-called Blood-Oxygenation-Level-Dependent (BOLD) signal. While dependent upon many other factors, the BOLD signal gives an indication of the amount of blood that flows to specific locations of the brain over time. Data might be more or less granular according to the measurement discretization on time and space (at specific voxels). After neurons in a small area become more active, the blood flows to that area is increased to supply the metabolic demand of the neural activity. The more intense the neural activation is, the more the blood flow increases, but the relationship is not necessarily linear or straightforward [11]. However, the mechanisms according to which the neural activity is related to cerebral blood volume, flow and oxygenation are still not completely clear [9] and this fact restricts the interpretation of fMRI studies [17].

Different types of analyses can be performed on fMRI data. For example, the characterization of the neural substrate of behaviour in terms of the location in the brain and the magnitude of the response may be clinically relevant for an individual to make a diagnosis, assess the course of disease or arrange a treatment and monitor its impact. Furthermore, it is also relevant to study how the location, the intensity and size of the neural activation differ between sick and healthy individuals in the population.

One of the main problems of fMRI data is that they are affected by multiple sources of noise. Hence, the true time evolution of brain-functional activity is often masked by systematic noise (e.g., scanner instability or changes); individual errors (e.g., anatomical variability, head motion, spin history, heartbeat, respiration, subject wakefulness) [11] as well as by other sources of noise. In approaching fMRI data we should firstly perform exploratory analysis to assess the quality of the data used in subsequent modelling. This may highlight aberrant features of specific regions, particular individuals or unexpected observations. Clustering of the fMRI time series has emerged in recent years as a possible alternative to parametric modelling approaches [10, 12], going beyond the traditional analysis of fMRI data which typically evaluate the level of activation of individual voxels. The study of the possible aggregation of different active voxels has the potential to identify biological relationships between the different functional areas of the brain. This motivated our investigation of fMRI clustering where we aim at grouping the regions with similar time series [25]. In this paper, we focus on a preliminary investigation of  $k$ -means clustering applied to a small subset of the existing data as a proof of concept.

The paper is organized as follows. Section 2 is dedicated to examining the available dataset. In Sect. 3 the  $k$ -means clustering is introduced and the proposed functional

methodology for the descriptive analysis of our data is described, whereas Sect. 4 reports our results. Section 5 contains discussion and possible future directions of our preliminary work.

## 2 Data Description

Investigation of brain structure and function has been much enhanced by recent innovations in both instruments (e.g., fMRI) and data analytic tools. This has raised new questions on brain function and it has motivated statistical developments to handle complex data structures [2, 17, 20, 21] and infer non-trivial phenomena.

The multimodal imaging dataset provided for this study comes from a pilot study of the Enhanced Nathan Kline Institute-Rockland Sample project. This project aims at providing a large cross-sectional sample of publicly shared multimodal neuroimaging data and psychological information to support and motivate researchers to explore the mechanisms underlying the complex brain system. A detailed description of the project, scope, and technical aspects can be found at [http://fcon\\_1000.projects.nitrc.org/indi/enhanced/](http://fcon_1000.projects.nitrc.org/indi/enhanced/).

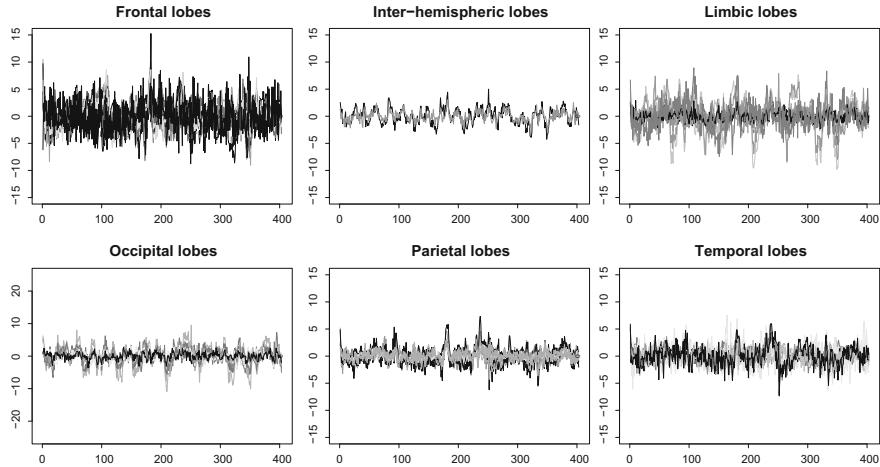
The specific data come from the pilot NKI1 study and include multimodal imaging data and subject-specific covariates for 24 subjects in resting state. In the NKI1 study, the subjects are simply asked to stay awake with their eyes open. Detailed information can be found at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html). The main variable of interest in our study is the dynamic functional activity: these data measure the dynamic activity of each brain region through changes in the BOLD signal during resting state fMRI (R-fMRI) sessions.

The anatomical brain regions on which such data are collected, are always the same across the subjects and they are based on the Desikan atlas [5]. For these  $e = 1, \dots, 70$  regions we have additional information on  $3 - D$  spatial locations, hemisphere and lobe membership. In addition, for the subjects in the study, we have age, handedness and psychological traits (whether diagnosed with a current or lifetime illness).

The dataset comprises an array of dimensions  $70 \times 404 \times 24 \times 2$  comprising the  $70 \times 404$  multivariate time-series activity data collected for the 24 subjects in each of the 2 scan-rescan imaging sessions. In particular,  $Y[:, , m, v]$  is a  $70 \times 404$  matrix whose rows contain the dynamic activity data of the brain regions, collected at  $l = 1, \dots, 404$  equally spaced times, for subject  $m$ , monitored during scan  $v$ , for every  $m = 1, \dots, 24$  and  $v = 1, 2$ .

### 2.1 Data Selection

Data are highly affected by missingness: if we consider the first scan, the time series over all regions are fully missing for two patients; regarding instead data on the second scan, more than half of the subjects have fully missing time series. This was



**Fig. 1** Time series of the BOLD from scan 1 for Patient 2: each plot groups all the measurements taken in a specific lobe

a first reason that motivated the selection of the time series of only one patient as our main dataset.

Furthermore, data are collected while patients are in resting state, and therefore there is no reason to believe that observations at a given time point should be similar across patients. Hence, we decided to focus part of our analysis only on one individual and we selected Patient 2 since he/she has very few missing values. In selecting only one patient we remove the main source of variability (subject) at the expense of obtaining results that will not be generalizable to the whole population.

The selected patient is a 52-years old, right handed person. If we group the BOLD time series by lobe, focussing only on the first scan, different patterns are identifiable, as shown in Fig. 1. Time series of some regions belonging to the same lobe show similarities, e.g., the time series of the inter-hemispheric, parietal and occipital lobes. Nevertheless, other lobe groups contain highly heterogeneous time series. This fact motivates us to apply  $k$ -means clustering to explore potentially unexpected spatial aggregation of BOLD time series.

### 3 Methodology

#### 3.1 *k*-Means Clustering

In recent years many studies have addressed the analysis of fMRI time series [10]. The aim of our contribution is to perform and evaluate  $k$ -means clustering of the available fMRI time series. To do that, we smooth the time series by  $b$ -splines and partition the estimated curves using  $k$ -means algorithm as in [1].

$k$ -means [16] is one of the simplest unsupervised clustering algorithms which is used when data are unlabelled (i.e., data without defined categories or groups). The procedure classifies units of a given dataset into  $k$  clusters, where  $k$ , the number of clusters, is fixed a priori. The implementation of this methods starts with the definition of  $k$  centroids, one for each cluster. Then, each unit is associated to the nearest centroid.  $k$ -means algorithm is a simple method with a fast convergence but, on the other hand, the results could severely depend on the number of clusters and distribution of the data. In fact, the algorithm relies on the parametric assumption that the data distribution is a mixture of  $k$  components. Furthermore,  $k$ -means algorithm does not necessarily find the optimal configuration, corresponding to the global objective function minimum.

To choose the number of clusters  $k$ , we use the well-known gap statistic [23]

$$gap(k) = \frac{1}{U} \sum_u \log(W_u^*(k)) - \log(W(k)) \quad (1)$$

which compares the observed within-cluster dispersion,  $W(k)$ , with the dispersion expected under the null hypothesis  $W_u^*(k)$  (i.e., no obvious clustering). This measure is calculated on  $U$  different uniform datasets, each with the same range as the original data, which are generated and grouped in  $k$  clusters. We computed the value of the gap statistic for several  $k$  and we selected the smallest number of clusters  $k$  such that

$$gap(k) \geq gap(k+1) - s_{k+1}$$

where  $s_k$  is the standard deviation of  $\log(W_u^*(k))$ .

### 3.2 Smoothing Procedure

Curves are usually observed at discrete observation points. For this reason, when working with functional data it is necessary to reconstruct the functional form of data by smoothing the time series. Smoothing is a very common procedure in the field of functional data analysis (FDA) as it estimates a smooth and manageable function able to capture relevant patterns in the original data while leaving out noise.

To perform the smoothing procedure we decided to use  $b$ -splines [4]. The selection of these techniques is motivated by the ease of their use and high flexibility. To do this, the abscissa-axis is divided into a number of intervals, where the endpoints of each interval are called breakpoints. These breakpoints are then converted to knots by imposing continuity and smoothness conditions at each interface. Given a non-decreasing knot vector  $\mathbf{t} = \{t_0, t_1, \dots, t_{n+o-1}\}$ , the  $n$  bases splines of order  $o$  are defined by

$$B_{h,1}(x) = \begin{cases} 1 & t_h \leq x \leq t_{h+1} \\ 0 & \text{otherwise} \end{cases}$$

$$B_{h,o}(x) = \frac{x - t_h}{t_{h+o-1} - t_h} B_{h,o-1}(x) + \frac{t_{h+o} - x}{t_{h+o} - t_{h+1}} B_{h+1,o-1}(x)$$

for  $h = 0, \dots, n - 1$ . When appropriate, knots are chosen on an interval, then the  $b$ -splines knots form a complete set on that interval and we can write the smoothing function as

$$f(x) = \sum_{h=0}^{n-1} c_h B_{h,o}(x) \quad (2)$$

where the coefficients  $c_h$  can be readily obtained from a least-squares fit [4]. In order to use  $b$ -splines, some parameters have to be tuned: specifically, we need to set the order of the spline and the number of basis functions (i.e., the knots) used for interpolation. Given the high level of noise that affects our time series, the choice of the number of basis functions is not straightforward as there is the risk of under-smoothing by choosing too many bases. The information on the data-collection scheme provided is not useful in determining the level of smoothing we should adopt. Therefore it is challenging to decide how much to smooth the original data since, possibly, turbulent and instantaneous phenomena, such as spikes, are not to be interpreted as pure noise, but as significant events that should not be smoothed.

We base our choice of the number of bases both on graphical evaluation and on cross-validation methods, while recognising that in a correlated data context cross-validation is known for not being optimal. The basic idea behind cross-validation is to divide the data into  $p$  subsets. Then the model that has to be validated is fitted  $p$  times, such that each time, one of the  $p$  subsets is used as the test set/validation set and the other  $p - 1$  subsets are put together to form a training set. Taking this procedure to its extreme, one can select a validation sample composed of one single datum, fit the data to the whole data except this one and then estimate the fitted value for the left out data value. This procedure is repeated for each observation in turn and for each parameter  $\lambda$ , such as the number of basis functions belonging to a user defined range. A criterion is recorded every time and, by minimizing it, the best  $\lambda$  is identified.

In the spline smoothing literature the most famous cross-validation technique is Generalized Cross-Validation (GCV) developed by Craven and Wahba in 1978 [3]. It has been proven to be more conservative than other cross-validation methods in avoiding over-smoothing. The criterion to minimize can be expressed in the following formula:

$$GCV(\lambda) = \left( \frac{n}{n - df(\lambda)} \right) \left( \frac{SSE}{n - df(\lambda)} \right)$$

where the  $SSE$ , the residual sum of squares, is first discounted by  $n - df(\lambda)$ , the number of training data minus an equivalent degrees of freedom measure, and then is further discounted by  $n/(n - df(\lambda))$ , the left factor. The minimization of GCV with respect to  $\lambda$  will inevitably involve a large number of values of  $\lambda$ , where grid-search or a numerical optimization algorithm is used. However it is possible to speed up the computation by performing a preliminary generalized eigenanalysis.

In this analysis we used the Generalized Cross-Validation available in the R package `fda` [18].

### 3.3 Functional Boxplot

The boxplot, firstly introduced by Tukey [24], is a graphical method used for summarizing the distribution of a dataset through the 5-number summary, which are the minimum and the maximum range values, the upper and the lower quartiles and the median. One of the advantages of this technique is the ability to identify observations classified as outliers in the data distribution. This exploration tool is principally based on data ordering which in the functional setting requires the introduction of a new depth measure indicating how much an observation is deep (central) or outlying.

In the case of functional data, Lopez and Romo [15] introduced the notion of Simplicial Band Depth (SBD) to order a set of real functions  $y_i(q)$  with  $i = 1, \dots, s$  and  $q \in I$ , where  $I$  is an interval in  $\mathbb{R}^2$ . Precisely, denoting  $y_{[i]}(q)$  as the sample curve associated with the  $i$ -th largest band depth value,  $y_{[1]}(q), y_{[2]}(q), \dots, y_{[s]}(q)$  can be viewed as order statistics where  $y_{[1]}(q)$  is the deepest (most central) curve, i.e. the median curve.

Given the curves  $y_1, \dots, y_s$ ,

$$D(y_1, \dots, y_s) = \{(q, x(q)) : q \in I, \min_{i=1, \dots, s} y_i(q) \leq x(q) \leq \max_{i=1, \dots, s} y_i(q)\}$$

is the band in  $\mathbb{R}^2$  composed by these curves. The band depth for a given curve  $x(q)$  is obtained by computing the fraction of the bands determined by  $j$  different curves containing the whole trajectory of the curve  $x(q)$ . The bigger the value of band depth, the more central position the curve has. In order to give a more flexible definition of the band depth, a Modified version of the Band Depth (MBD) was introduced by measuring the proportion of times that a curve  $x(q)$  is in the band. Taking inspiration from the classical boxplot, Sun and Genton [22] used these depth measures to define the functional boxplot.

In the classical boxplot, the middle 50% of the data are represented by the box itself. This same idea is still used in the functional boxplot: the  $\alpha$  central region of the plot delimits the  $\alpha$  proportion of deepest curves. For instance the 50% central region is

$$C_{0.5} = \{(q, x(q)) : \min_{i=1, \dots, [s/2]} y_{[i]}(q) \leq x(q) \leq \max_{i=1, \dots, [s/2]} y_{[i]}(q)\}$$

where  $[s/2]$  identifies the smallest integer not less than  $s/2$ . The borders of the box of the classical boxplot are the borders of the central region of the functional boxplot. In addition, the central region, which is a robust range, contains the median  $y_{[1]}(q)$  or the curve having the largest band depth value. In order to identify outlying curves the whiskers must also be defined. The whiskers are vertical lines of the plot extending from the box and indicating the maximum envelope of the dataset except the outliers. These fences are obtained, following the classical outlier criterion, by increasing the 50% central region by 1.5 times the range of the 50% central region. Any curve outside the whiskers can be considered potential outliers. More useful information about a functional dataset can be revealed by looking at the position, size, length and shape of box, whiskers and median.

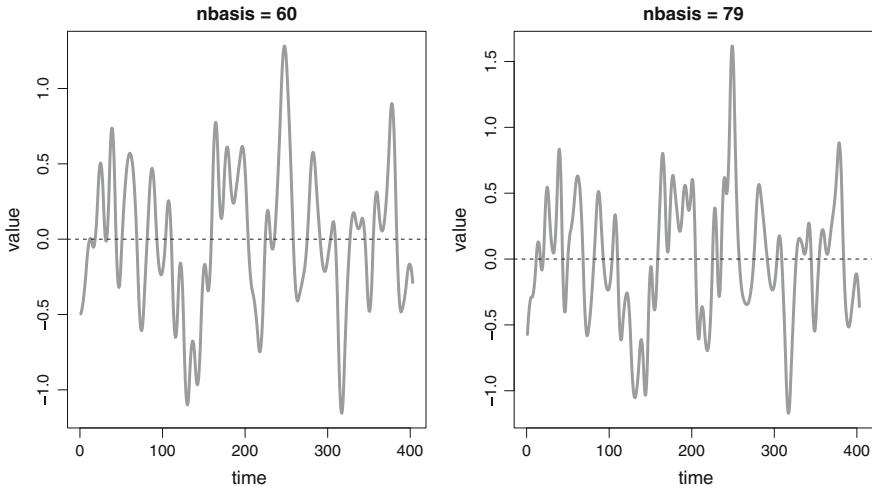
## 4 Results

This section reports the results of the analyses introduced in Sect. 3. As mentioned in Sects. 2.1 and 3.1, the clustering technique was applied to the data of Patient 2. From Sect. 4.1 onwards, data are treated as functional data and therefore they are smoothed. We then use an exploratory method for functional data analysis: the functional boxplot (Sect. 4.2). Finally, the results of  $k$ -means clustering are reported in Sect. 4.3.

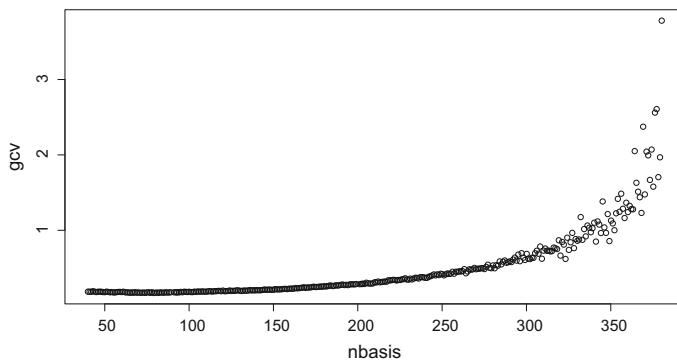
### 4.1 Smoothing Procedure

According to the methods exposed in Sect. 3.2, we smoothed the time series using splines of order 3 with 60 basis functions in order to give the data a more regular surface without losing possible differences in the signal between the time points. To further evaluate our chosen bases number, we verify the goodness of our choice with GCV. Using the `gcv` function, available in the R package `fda` [18], the number of basis functions, minimizing the aforementioned criterion, was 79. Despite this result, we stick to our initial choice of 60 basis, because of the graphical evaluation of Fig. 2 and of Fig. 3. The former shows that two smoothed functions, obtained by using 60 and 79 bases functions, differ only in some minor peaks. The latter demonstrates no visually-evident difference between selecting 60 or 79 basis functions in terms of the GCV criterion, rather than having a unique minimum, the GCV, is relatively small in the range 50-90, among these values we chose one of the smaller to minimize model complexity. The results of the smoothing procedure is shown in Fig. 4.

We repeated the same analysis on the data from scan 2 and we selected 150 bases.



**Fig. 2** Same smoothed curve using 60 or 79 basis functions

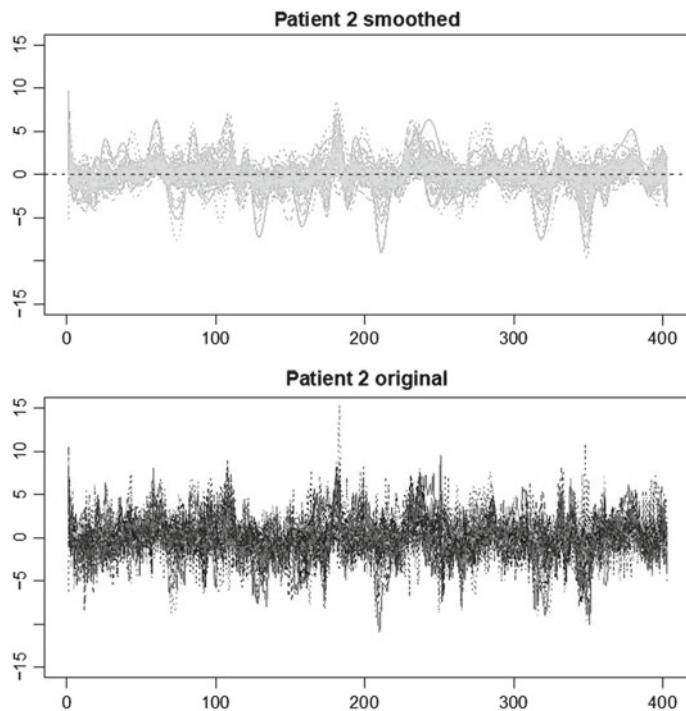


**Fig. 3** GCV value for numbers of basis functions in the case of the first scan of Patient 2

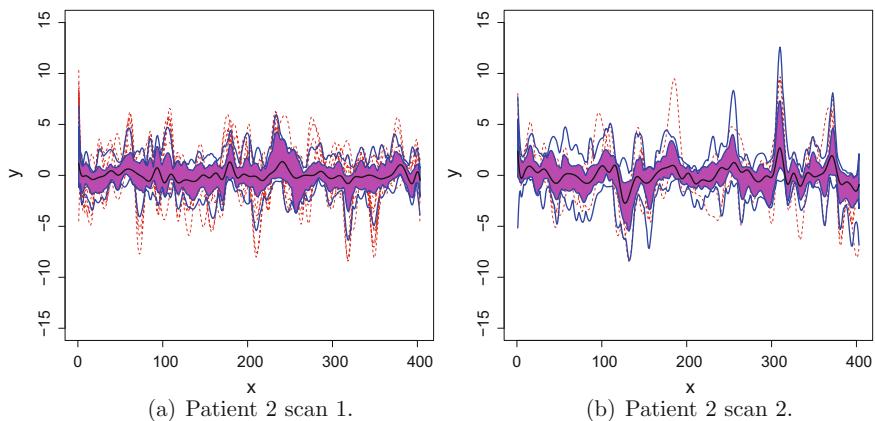
#### 4.2 Functional Boxplot

The functional boxplot was used to identify outlying curves in the smoothed dataset. Focusing on Patient 2, it is possible to notice how, passing from scan 1 to scan 2, the detected outliers differ. Figure 5a, b show these differences. Precisely, blue solid curves denote envelopes, black curves represent the median curve and the red dashed curves are the outlier candidates which are those curves out of the boundaries identified by the whiskers. The magenta area indicates the 50% central region.

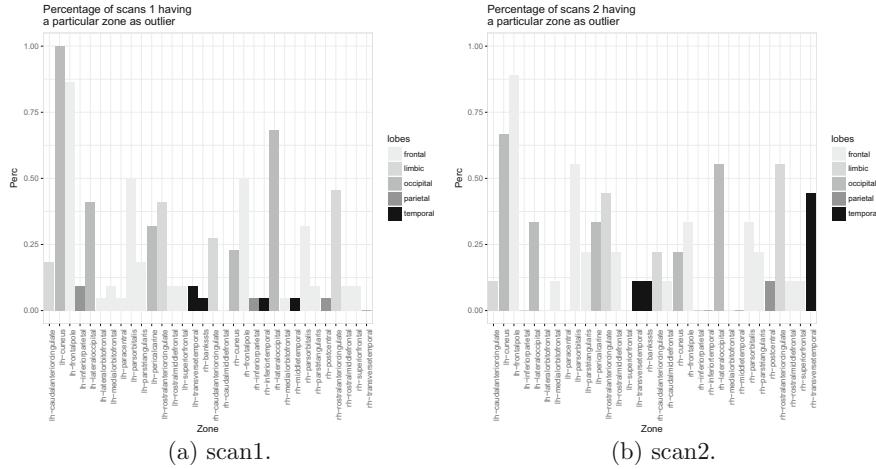
As one can see in Fig. 5a, scan 1 presents 10 curves detected as outliers. These are the lh-caudalanteriorcingulate, the lh-cuneus, the lh-parsorbitalis, the lh-parstriangularis, the lh-superiorfrontal, the lh-frontalpole, the rh-caudalanteriorcingulate, the rh-rostralanteriorcingulate, the rh-superiorfrontal, and



**Fig. 4** Smoothed data with  $b$ -splines of order 3 and 60 knots



**Fig. 5** Functional Boxplots



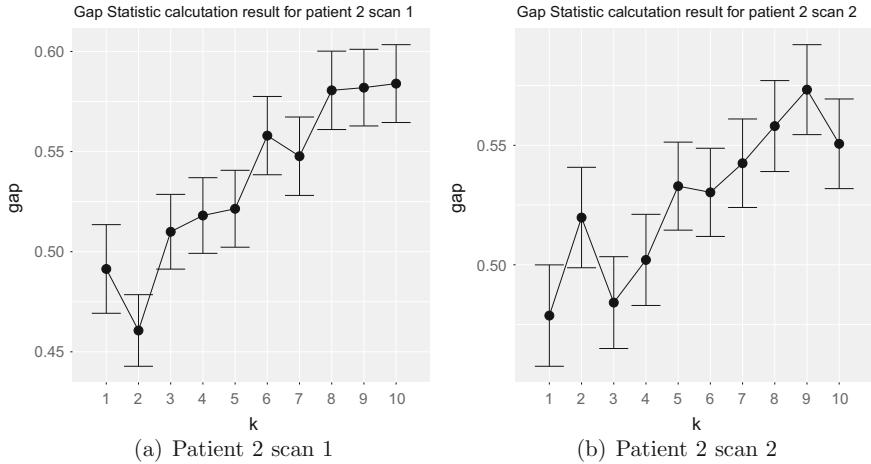
**Fig. 6** Percentage of scans 1 and 2 having a particular zone as outlier

the rh-frontalpole. Instead there are 4 outlying functions shown in Fig. 5b: the lh-lateraloccipital, the lh-parstriangularis, the lh-frontalpole and the rh-transversetemporal. Hence, only the lh-parstriangularis and the lh-frontalpole are coherently detected as outliers in both the scans.

Using the functional boxplot applied to every scan of every patient we could assess whether some brain regions are detected as outlying more often than others. Figure 6 shows the percentage of patients having a particular curve as outlier in scan 1 and scan 2, respectively. Figure 6a, shows that the lh-cuneus is an outlier curve in all the patients. Repeating the same analysis on scan 2, (Fig. 6b) lh-cuneus is identified as the second most frequent outlier region below the lh-frontalpole. Based on the comparison between scan 1 and scan 2 of the patients having both scans, it is evident that the curves detected as outliers usually differ from one scan to another. This means that there is no patient having the same outlying functions according to both scan 1 and scan 2.

### 4.3 *k*-Means Clustering

Functional data analysis focuses on the development of statistical tools for analysing samples of curves. Classical multivariate techniques can be applied to functional datasets, but they do not take advantage of the additional information contained in the functions and its derivatives. Using the smoothed data it is possible to find clusters [14]. We apply  $k$ -means clustering to smoothed data of Patient 2, for scan 1 and scan 2 separately.



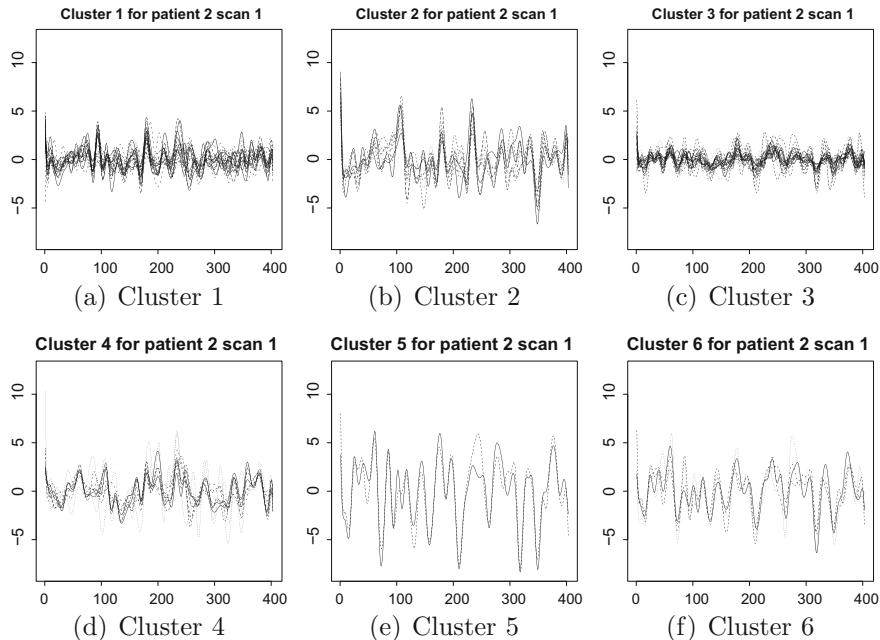
**Fig. 7** gap Statistic for  $k$ -means clusters algorithm ( $k = 1, \dots, 10$ ) in Patient 2 in scan 1 (a) and scan 2 (b) for the smoothed data

**Table 1** Clusters' cardinality in  $k$ -means algorithm with  $k = 6$

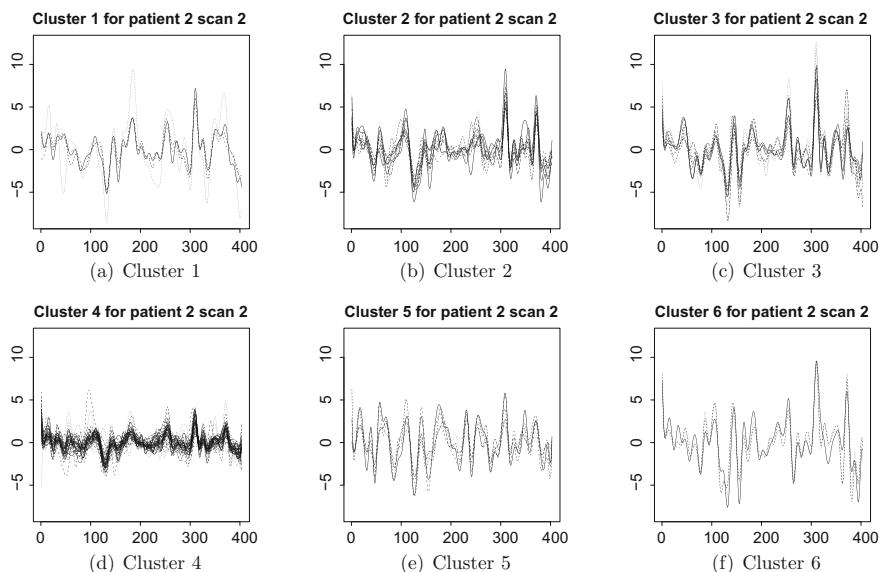
Scan	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
1	21	9	25	9	2	4
2	4	11	8	39	5	3

Regarding the choice of  $k$ , the gap statistic does not uniquely identify the number of clusters  $k$ , i.e., there is no clear point at which the rate of increase of the gap statistic reaches a plateau (Fig. 7). The number of selected clusters seems to be higher than  $k = 10$  and this is not desirable in our context. However, the tendency of the gap statistic to overestimate the number of clusters has been demonstrated by Dudoit and Fridlyand [6]. One possible reason for the deficiency of the gap method is that the statistic summarizing the within-clusters homogeneity, is not suitable for measuring the clustering adequately. Another reason might be the excessive noise of the observations over time despite the smoothing procedure. Given that the gap statistic curve does not give a clear result, we adopted a more pragmatic criterion to chose the number of clusters  $k$ . Since human brain is divided in six lobes, this could be a criterion for the choice of the number of clusters. Therefore, we select  $k = 6$  clusters.

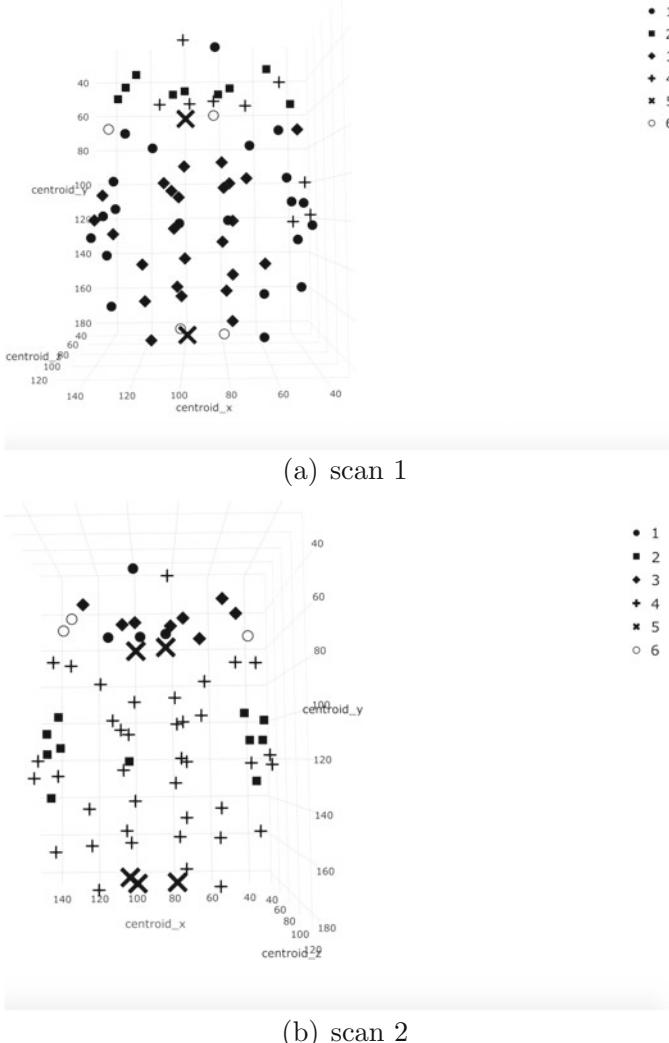
In Table 1 we report the cluster cardinalities of the  $k = 6$  clusters resulting from the analysis applied to data from scan 1 and scan 2: results are not consistent across scans. Hence, we test the robustness of the classification by the adjusted Rand Index [13, 19] using R package mclust [7, 8]. The resulting adjusted Rand index is equal to 0.1 and it confirms the hypothesis of weak agreement between the results of  $k$ -means alignment cluster applied to scan 1 and scan 2.



**Fig. 8** Clusters for  $k$ -means algorithm with  $k = 6$  in Patient 2 in scan 1



**Fig. 9** Clusters for  $k$ -means algorithm with  $k = 6$  in Patient 2 in scan 2



**Fig. 10** Clustering on human brain based on  $k$ -mean clustering with  $k = 6$

Figures 8 and 9 show the curves inside each clusters in scan 1 and scan 2. With regard to scan 1, we can observe that some clusters have very high peaks while others (e.g., cluster 3) have more smooth curves. This does not happen in scan 2 where all the curves are characterized by high peaks.

The brain is neuro-anatomically divided into two hemispheres and the clustering results, which are shown in Fig. 10, support the statistically defined clusters.

## 5 Discussion and Future Directions

In this paper we have explored a dataset of resting state measurements of BOLD over time and over brain regions for 24 patients. These data are by nature very complex, being spatio-temporal and with repeated measures. Firstly, there are no specific reasons for systematic changes over time that a non-resting state fMRI experiment could imply, with reactions to external stimulus. Secondly, regions are characterized by their location in the brain. However, due to the physical structure of the grey matter, geometric distance on the 3D space might not be an adequate measurement for the actual distance across regions. The consequence of these features are detected in our exploratory analysis that concludes that the raw-data are dominated by noise and by variability across the subjects.

The aim of this paper is to classify the regions of the brain according to the temporal pattern of the BOLD. We have focused on the data from one patient. Clearly this could be extended to include the set of patients and the repeated scans.

$k$ -means is one of the simplest unsupervised clustering algorithm but, despite its computational tractability, it is a stochastic algorithm, and therefore it strongly depends on the initialization of the centroids.

To perform functional clustering we firstly must detect the signal underlying the noisy.  $b$ -splines smoothing has been used to approximate the time series. The choice of the number of knots and basis functions to use was made according to graphical evaluation and cross validation.

Then  $k$ -means clustering was performed on the smoothed data with the number of groups  $k = 6$ , same as the number of lobes in the brain.

Brain data are very complex and noisy, traditional clustering methods might not be adequate to classify the regions of the brain. Advanced spatial functional clustering leads to the identification of areas behaving similarly, even in resting states time series. While our work has focussed on a very limited subset of the data, it would be interesting to analyse how selecting different smoothing procedures and settings could affect the clustering results and to explore a more detailed model for all patients and both scans. To do this, the joint work of statisticians, neurologists and technicians who take the measurements seems to be essential.

**Acknowledgements** We are grateful to Greg Kiar and Eric Bridgeford from NeuroData at Johns Hopkins University, who graciously pre-processed the raw DTI and R-fMRI imaging data available at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html), using the pipelines ndmg and C-PAC and to the reviewers and The Scientific Committee of StartUp Research for all the suggestions aimed at improving this paper.

## References

1. Abraham, C., Cornillon, P.A., Matzner-Løber, E., Molinari, N.: Unsupervised curve clustering using b-splines. *Scand. J. Statist.* **30**(3), 581–595 (2003)
2. Craddock, R.C., Jbabdi, S., Yan, C.G., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C., Heberlein, K., Colcombe, S., Milham, M.P.: Imaging human connectomes at the macroscale. *Nature Methods* **10**(6), 524 (2013)
3. Craven, P., Wahba, G.: Smoothing noisy data with spline functions. *Numerische Mathematik* **31**(4), 377–403 (1978)
4. De Boor, C.: *A Practical Guide to Splines*. Springer, New York (1978)
5. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**(3), 968–980 (2006)
6. Dudoit, S., Fridlyand, J.: A prediction-based resampling method for estimating the number of clusters in a dataset. *Genome Biol.* **3**(7), 1–21 (2002)
7. Fraley, C., Raftery, A.E.: Model-based clustering, discriminant analysis and density estimation. *J. Am. Statist. Assoc.* **97**, 611–631 (2002)
8. Fraley, C., Raftery, A.E., Murphy, T.B., Scrucca, L.: mclust Version 4 for R: Normal Mixture Modeling for model-based clustering, classification, and density estimation (2012)
9. Goense, J., Bohrhus, Y., Logothetis, N.K.: fMRI at high spatial resolution: implications for BOLD-models. *Front. Computat. Neurosci.* **10**, 66 (2016)
10. Goutte, C., Toft, P., Rostrup, E., Nielsen, F.Å., Hansen, L.K.: On clustering fMRI time series. *NeuroImage* **9**(3), 298–310 (1999)
11. Greve, D.N., Brown, G.G., Mueller, B.A., Glover, G., Liu, T.T., et al.: A survey of the sources of noise in fMRI. *Psychometrika* **78**(3), 396–416 (2013)
12. Heller, R., Stanley, D., Yekutieli, D., Rubin, N., Benjamini, Y.: Cluster-based analysis of fMRI data. *NeuroImage* **33**(2), 599–608 (2006)
13. Hubert, L., Arabie, P.: Comparing partitions. *J. Classif.* **2**(1), 193–218 (1985)
14. Jacques, J., Preda, C.: Functional data clustering: a survey. *Adv. Data Anal. Classif.* **8**(3), 231–255 (2014)
15. López-Pintado, S., Romo, J.: On the concept of depth for functional data. *J. Am. Statist. Assoc.* **104**(486), 718–734 (2009)
16. MacQueen, J., et al.: Some methods for classification and analysis of multivariate observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Oakland, CA, USA vol. 1, pp. 281–297 (1967)
17. Monti, M.M.: Statistical analysis of fMRI time-series: a critical review of the GLM approach. *Front. Hum. Neurosci.* **5**, 28 (2011)
18. Ramsay, J.O., Wickham, H., Graves, S., Hooker, G.: fda: Functional data analysis (2017). <https://CRAN.R-project.org/package=fda/>
19. Rand, W.M.: Objective criteria for the evaluation of clustering methods. *J. Am. Statist. Assoc.* **66**(336), 846–850 (1971)
20. Sporns, O.: Structure and function of complex brain networks. *Dialogues Clin. Neurosci.* **15**(3), 247–262 (2013)
21. Stam, C.J.: Modern network science of neurological disorders. *Nature Rev. Neurosci.* **15**(10), 683–695 (2014)
22. Sun, Y., Genton, M.G.: Functional Boxplots. *J. Comput. Graph. Statist.* **20**(2), 316–334 (2011)
23. Tibshirani, R., Walther, G., Hastie, T.: Estimating the number of clusters in a data set via the gap statistic. *J. Royal Statist. Soc. Ser. B (Statist. Methodol.)* **63**(2), 411–423 (2001)
24. Tukey, J.W.: *Exploratory Data Analysis*, vol. 2. Reading, Mass (1977)
25. Yeo, B., Ou, W.: Clustering fMRI time series (2004). <http://people.csail.mit.edu/ythomas/unpublished/6867fMRI.pdf>

# Robust Methods for Detecting Spontaneous Activations in fMRI Data



Francesca Gasperoni and Alessandra Luati

**Abstract** Functional magnetic resonance imaging (fMRI) is a technique for measuring brain activity. The outcomes of fMRI measurements are complex data that can be interpreted as multivariate time series, recorded at different brain locations, usually across subjects. The literature has been mainly concerned with task-based fMRI analysis, which focuses on the response to controlled exogenous stimuli. Nevertheless, resting state fMRI (RfMRI) analysis, dealing with spontaneous brain activity, is considered the key to understand the neuronal organisation of the brain. The aim of this paper is to identify spontaneous neural activations and to estimate the brain response function in RfMRI data, called Hemodynamic Response Function (HRF). To this purpose, we apply an existing method based on a normality assumption for the data generating process and we consider a novel, more general method, based on robust filtering. Finally, we compare the neural activations and HRF estimates for two specific patients.

**Keywords** BOLD signal · Heavy tails · HRF estimation · Resting state  
Robust filtering · Spatial dependence

## 1 Introduction

Functional magnetic resonance imaging (fMRI) is a non invasive technique for collecting data on brain activity, with a good resolution in terms of space and time. Essentially, fMRI measures the increase in the oxygenation level at some specific

---

F. Gasperoni (✉)  
MOX, Department of Mathematics, Politecnico di Milano, Milan, Italy  
e-mail: [francesca.gasperoni@polimi.it](mailto:francesca.gasperoni@polimi.it)

A. Luati  
Department of Statistics, University of Bologna, Bologna, Italy  
e-mail: [alessandra.luati@unibo.it](mailto:alessandra.luati@unibo.it)

brain region, as long as an increase in blood flow occurs, due to some brain activity. The latent signal in the observed fMRI data is referred to as the blood oxygenation level dependent (BOLD) signal.

A large stream of the literature has dealt with the analysis of task based experiments, where the BOLD is measured in response to some stimulus or event. On the other hand, only in recent years the interest has been concentrating toward resting state fMRI (RfMRI), as the key to understand the neuronal organisation of the brain, through the investigation of the spatial and temporal structure of spontaneous neural activity. The earliest work that focused on RfMRI was the one of Biswal et al. [2], where it is shown that the same brain regions that were active and correlated during finger tapping were also correlated with the spontaneous BOLD fluctuations in the absence of any stimulus. Since then, growing attention has been devoted to RfMRI data analysis. In the review paper by Biswal [3], resting state fMRI is described as the candidate approach capable of addressing the core challenge in neuroimage, i.e., the development of common paradigms for interrogating the functional systems in the brain, without the constraint of a priori hypotheses.

When analysing fMRI series, major inferential issues arise, due to the complexity of the data. As a matter of fact, fMRI are recorded as high frequency time series, observed at different brain region of interests (ROI) or, on a finer scale, at different voxels, across individuals. Hence, besides the intrinsic dynamic nature of fMRI, the researcher has to take into account aspects like curse of dimensionality at the voxel level, presence of explanatory variables depending on the design of the experiment, spatial correlation, which is of interest for connectivity analysis, as well as multivariate aspects related to multisubject or group analysis.

Despite specifically depending on the aim of the analysis, the most common approach followed to analyse fMRI data consists in a sequence of analytical steps that eventually result in a general linear model accounting for time and space correlation. The basic assumption for the first step, that is a univariate time series analysis, consists in a simple decomposition of pre-processed fMRI data into an unobserved signal plus noise. The underlying hypothesis on the two latent variables are related to the evolution of the components in the time and with the dependence on some explanatory variables. Once the dynamic characteristics of the series are acknowledged, their interrelation across ROIs and subjects becomes the main object of the analysis, so that methods ranging from spatial modelling to large covariance matrix estimation are required.

Most of the models used for the time series analysis of fMRI assume a stationary Gaussian distribution for the noise term. However, the stationarity hypothesis seems to be not justifiable and the assumptions on the dynamics in fMRI are still controversial. Indeed, there is still a considerable debate on the dynamic properties of fMRI and AR( $p$ ) errors have been considered, see e.g. [23, 27], as well as fractional noise error processes [5], and recently change point methods [1] as an alternative to stationarity. Moreover, Lund et al. in [24] concluded that no commonly accepted model for noise in fMRI exists and that regressors may whiten the noise as well as high pass filters. In resting state studies, the dynamics of the noise component are even more relevant than in task based experiments, as recognised in [12], since, ideally,

no exogenous stimulus affects the underlying signal and the noise dynamics reflect the human brain resting activity.

In this paper, we aim at specifying a model with more general assumptions on the noise term, which is allowed to be non Gaussian, and on the signal, which is a possibly non stationary dynamic process that may depend on estimated hemodynamic response functions specifically designed for spontaneous neural activity.

The first contribution of our analysis is indeed the estimation of the hemodynamic response function, and consequently, of the potential explanatory variables for the signal of RfMRI data. In the case of resting state experiments, the challenge lies in the specification of an impulse response function that accounts for spontaneous neural activity in the absence of controlled stimulus for each ROI. Considering a general regression model with Gaussian stationary autocorrelated errors, Wu et al. in [29] designed a strategy for HRF reconstruction based on the identification of spontaneous activations as extreme values from a Gaussian distribution. As pointed out in [28], the problem of detecting spontaneous activations has much in common with outlier detection. Our idea is then to identify spontaneous activations as outliers of a Normal distribution or as extreme values of a heavy tailed distribution. Specifically, we assume a Student- $t$  distribution for the noise term and identify spontaneous activations as extreme values of residuals obtained from a robust procedure for signal extraction, as in [18]. Modeling the data under the assumption of a heavy tail distribution for the noise affecting the BOLD signal is the second contribution of the paper.

We shall estimate and compare the HRF obtained from the two methods and illustrate the results on two patients and four ROIs. Spatial dependence will be explored through the investigation of a proper similarity index for binary data. In particular, we compute and plot the similarity matrices related to the estimated spontaneous events across ROIs. Thanks to this approach, we are able to detect and study groups of ROIs with the highest similarity.

The paper is organised as follows. Section 2 describes the methods used for identifying the spontaneous activations and estimating the HRF. Section 3 illustrates the results of the analysis obtained in four specific ROIs of two patients. Section 4 concludes the paper with some directions for further research.

## 1.1 *Dataset Description*

The multimodal imaging dataset that we are considering comes from a pilot study of the Enhanced Nathan Kline Institute-Rockland Sample project. This project aims at providing a large cross-sectional sample of publicly shared multimodal neuroimaging data and psychological information to support and motivate researchers in the relevant scientific goal of understanding the mechanisms underlying the complex brain system. A detailed description of the project, scopes, and technical aspects can be found at [http://fcon\\_1000.projects.nitrc.org/indi/enhanced/](http://fcon_1000.projects.nitrc.org/indi/enhanced/). The pilot NKI1

study comprises multimodal imaging data and subject-specific covariates for  $n = 24$  subjects. Detailed information can be found at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html). An appealing aspect of this pilot study, compared to the whole dataset, is that for a wide set of subjects scan-rescan imaging data are available, thereby allowing validation and inference also on subject-specific variability in brain functions and structures.

For each subject several information are collected, both personal covariates (i.e., anxiety diagnosis, age, gender, handedness) and BOLD signals, which are recorded for all 70 ROIs according to Desikan atlas. For such  $V = 70$  regions we have additional information on 3-D spatial locations, hemisphere and lobe membership. One region, in the left and right hemisphere, is marked as unknown, and typical analyses consider only the 68 regions characterizing the Desikan atlas.

We have the following datasets:

- (1) **Structural networks**,  $70 \times 70 \times 24 \times 2$ , comprising the  $70 \times 70$  structural connectivity networks collected for the 24 subjects in each of the 2 scan-rescan imaging sessions.

Focusing on subject  $i$  and on scan  $k$ , where  $i = 1, \dots, 24$  and  $k = 1, 2$ , we have a  $70 \times 70$  symmetric adjacency matrix measuring the total number of white matter fibers connecting each pair of brain regions.

- (2) **Dynamic functional activity**,  $70 \times 404 \times 24 \times 2$ , comprising the  $70 \times 404$  multivariate time-series data collected for each subject in each of the 2 scan-rescan imaging sessions.

This imaging technology monitors brain functional activity at different regions via dynamic changes in blood flow creating a low frequency BOLD signal when the subject is not performing an explicit task during the imaging session. In the present NKI1 study, the subjects are simply asked to stay awake with eyes open. Focusing on subject  $i$  and on scan  $k$ , where  $i = 1, \dots, 24$  and  $k = 1, 2$ , we have  $70 \times 404$  matrix whose rows contain the dynamic activity data of the brain regions, collected at  $T = 404$  equally spaced times (time lag is 1400 ms).

- (3) **Functional networks**,  $70 \times 70 \times 24 \times 2$ , measuring synchronization in activity for each pair of brain regions (obtained through Pearson correlation of the previous dataset fixing the ROI, the subject and the scan).

Our major interest lays on the investigation of differences among subjects, based on the dynamic functional activity of their brains, so we focus on the second dataset.

Moreover, we have ROI-specific information. Not having voxel-specific data leads to small data, which means that we expect not to have computational issues and therefore we shall not need to apply dimension reduction techniques such as principal component analysis (PCA) or independent component analysis (ICA).

## 2 Modelling fMRI Data

The most common approach for analysing fMRI data is based on General Linear Models (GLM), where the observed fMRI is taken as the response variable, while the BOLD signal is modelled through a set of explanatory variables, usually taken as the HRF and its temporal and dispersion derivatives. The latter are the partial derivatives with respect to the two parameters that characterise the canonical HRF distribution: the time to peak and the width of the HRF at half the height (Chap. 14 in [20]). On one hand, the estimate of temporal derivatives allows us to capture small changes in the latency of the response, on the other hand, the estimate of dispersion derivatives allows us to capture differences in the duration of the peak response.

A stationary Gaussian process with non-spherical covariance matrix is assumed for the noise term. This approach is usually named general linear model and it has been studied in literature at individual level [26]. A common choice for the HRF is the double-gamma function, which can be derived through a deconvolution process between a stimulus function and the BOLD signal. The choice of the stimulus function is straightforward when task based experiments are considered (e.g. a step function), but it is not trivial in case of RfMRI.

The main contribution of this paper consists in the estimate of RfMRI hemodynamic response functions at the ROI level, with two different methodologies: on one hand, we take advantage of the blind deconvolution technique proposed in [29]; on the other hand, we consider a more general method, based on robust filtering, developed in [18]. We then compare the performance of the two methods in terms of spontaneous activation detection and analyse the estimated parameters that characterize the HRF.

Both methods are based on a signal plus noise decomposition,

$$y_t = \mu_t + \varepsilon_t$$

where  $y_t, t = 1, 2, \dots, T$ , is the observed fMRI data (in fMRI, the term functional refers to time series data),  $\mu_t$  is the unobserved BOLD signal and  $\varepsilon_t$  is a noise term.

In the framework of general linear modelling, where the blind deconvolution technique is developed [29], the signal is explained by a set of (estimated, see Sect. 2.2) HRF and the noise is a first order Gaussian autoregressive process, i.e.  $\mu_t = X_t \beta$  and  $\varepsilon_t = \phi \varepsilon_{t-1} + \eta_t$  where  $\eta_t \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$  and  $|\phi| < 1$ .

The alternative model we are considering here assumes that the noise is a Student- $t$  process,  $\varepsilon_t \sim t_\nu(0, \sigma^2)$ , while the signal follows a first order autoregressive model,  $\mu_t = X_t \beta + \phi \mu_{t-1} + \kappa u_{t-1}$ , where  $u_t$  is a martingale difference sequence, i.e.,  $E(u_t | \mathcal{F}_{t-1}) = 0$ ,  $\mathcal{F}_{t-1}$  represents the information set up to time  $t - 1$ , and  $|\phi| \leq 1$ , i.e., the model can account for non stationary components. This model belongs to the class of score driven models, recently introduced in [9] and in [19].

These are nonlinear observation driven models where the dynamic parameters ( $\mu_t$  in our case) are updated by filters that are robust with respect to extreme values. The robustness comes from the properties of the martingale difference sequence  $u_t$ , that

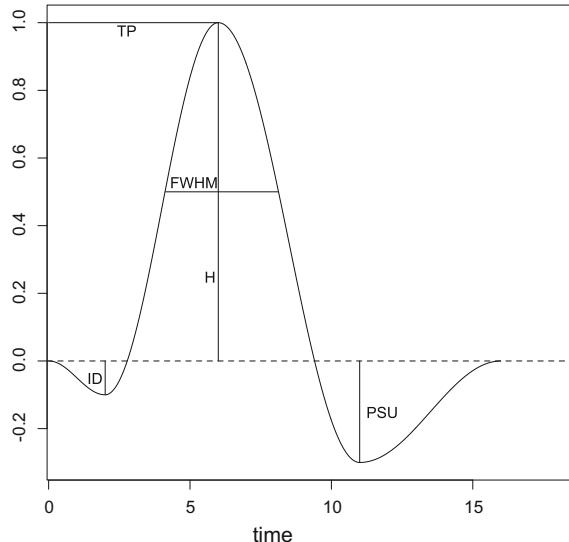
drives the dynamics of the time varying parameter. The main feature of score driven models is that  $u_t$  is proportional to the score of the conditional (with respect to  $\mathcal{F}_{t-1}$ ) likelihood of  $\mu_t$ . When the data come from a heavy tail distribution, the score  $u_t$  is less sensitive to extreme values than the score of a Gaussian distribution or than the innovation error  $v_t = X_t - \mu_t$ . We shall further discuss the robustness properties of the model in Sect. 2.2.

The static parameters,  $\nu$ ,  $\phi$ ,  $\kappa$ ,  $\sigma$  in our case, are consistently estimated by maximum likelihood and asymptotic standard errors can be derived [18, 19]. An important property of the proposed specification is that it encompasses the Gaussian case, in that the score of the Student- $t$  converges to that of the Gaussian distribution when the degrees of freedom tend to infinity. In practice, if a score driven model is estimated when the underlying dataset is in fact Gaussian, a very high value for the degrees of freedom is estimated and a Gaussian model is eventually fitted.

## 2.1 The BOLD Signal

The BOLD signal arises from the interplay of blood flow, blood volume, and blood oxygenation in response to changes in neural activity. Under an active state, the local concentration of oxygenated hemoglobin increases, with a corresponding increase in the homogeneity of magnetic susceptibility, which, in turn, results in an increase of MRI signal. The BOLD signal does not increase instantaneously and does not return to baseline immediately after the stimulus ends. Because these changes in blood flow are relatively slow (evolving over several seconds), the BOLD signal is a blurred and delayed representation of the original neural signal. The HRF can be described as the ideal, noiseless response to an infinitesimally brief stimulus. Five characteristics determine the HRF: time from the stimulus until peak (TP, or time-to-peak), height of response (H), the width of the HRF at half the height (FWHM), poststimulus undershoot (PSU) and initial dip (ID), see Fig. 1 reproduced from [25]. Importantly, there is substantial variability in each of these features of the HRF across brain areas and across individuals. For example, in [21], the time until peak varied between 6 and 11 seconds across voxels in a single subject. In [17], a study of the HRF shape revealed that both the time until peak and width of the HRF varied within subjects across different regions of the brain and across subjects, with intersubject variability higher than intrasubject variability. D’Esposito et al. in [10] reviewed a number of studies that compared the BOLD signal in healthy young and old subjects and found, while the shape of the HRF was similar between the groups, elderly had reduced signal-to-noise ratios in the response magnitudes.

The HRF is not observed, i.e., it is not recorded by medical instruments, but it has to be estimated from the observed fMRI series based on some underlying assumptions on its behaviour in time. To this aim, a widely recognized technique in neuroscience is the convolution between a controlled temporal stimulus and an unobservable hemodynamic response, depending on an unknown static parameter. Additively combined with a zero mean noise, the convolution results in an explana-



**Fig. 1** Example of a hemodynamic response function (HRF). The abscissa reports the time, in seconds, ranging from 0 (occurrence of the stimulus) to 15 s (following the post-stimulus undershoot, when the level returns to the baseline); the ordinates report the value of hemodynamic response function across time on a standardised scale from 0 to 1. The main characteristics of HRF are highlighted: time from the stimulus until peak (TP), height of response (H), the width of the HRF at half the height (FWHM), poststimulus undershoot (PSU) and initial dip (ID)

tory variable for the observed data, which reflects that neural response and the BOLD signal exhibit linear time invariant (LTI) properties. By applying deconvolution models, Friston et al. in [15] and Lange and Zeger in [22] found out that in general HRF could be described by a gamma function. However, this function is not able to capture the poststimulus undershoot which characterizes the HRF, so later on Friston et al. in [14] and Glover in [16] proposed a combination of two gamma functions, known as double-gamma, for describing the HRF. All these studies are related to task based fMRI, when a deconvolution procedure according to the explicit task design is possible. Methods for estimating the HRF and the role of the HRF in case of RfMRI are far less explored in literature.

A first attempt to estimate the HRF in RfMRI through a blind-deconvolution technique is described in [29]. The latter is illustrated in the next section along with our alternative strategy.

## 2.2 HRF Estimation

In task based experiments, HRF estimation is based on the convolution of some basis function,  $x_{BF}$ , with an activation function which accounts for the stimulus,  $s_t$ , that

is usually represented by a step function or by a series of delta functions, according to the experimental design (e.g., finger-tapping, hot-cold exposure).

In resting state analysis, where no external stimuli are recorded, spontaneous activations, commonly referred to as spontaneous low-frequency fluctuations, as in [12] or spontaneous neural state, as in [29], should be estimated, hence the reference to  $\hat{s}_t$ . In the current analysis,  $\hat{s}_t$  will be a binary time series of zeros and ones with one corresponding to activation and zero to rest. The key idea behind the identification of the time series of spontaneous events is that it can be expressed as a sequence of delta functions that are activated in correspondence of extreme values of a Gaussian or Student- $t$  distribution. The idea is motivated by the fact that time series extreme values can be seen as spontaneous neural activations of the brain.

Once the spontaneous events are detected, and this is the central focus of this work, the estimated HRF is obtained by the estimation procedure shown in [29], for the two models illustrated in the introduction of Sect. 2.

The overall HRF estimation strategy can be summarised as follows:

- (i) estimate the spontaneous activations  $\hat{s}_t$ ;
- (ii) find a preliminary estimate of the design matrix  $X_t$  as the convolution of  $\hat{s}_t$  with a suitable basis function (a canonical double gamma function, in this case)  $x_{BF}$ , as follows,

$$X_t = \hat{s}_t \otimes x_{BF} = \int_{-\infty}^{+\infty} \hat{s}_t \cdot x_{BF}(t - \tau) d\tau;$$

- (iii) plug the preliminary estimate of  $X_t$  as the matrix containing the explanatory variables in the Gaussian or Student- $t$  model, in order to obtain an estimate of the coefficients  $\hat{\beta}$ ;
- (iv) obtain the HRF estimate as  $\hat{h}_t = x_{BF}\hat{\beta}$ .

We now discuss the procedure under the Gaussian linear model and the Student- $t$  non linear model described in the introduction of Sect. 2.

Under the Gaussian model in the formulation proposed in [29], spontaneous activations are estimated by standardising the original data and marking as outliers those points that exceed a certain cut-off, fixed at  $1.65\hat{\sigma}$ , i.e., the 10% of extreme values of each series are considered as spontaneous activations. The preliminary estimate of the design matrix  $X_t$  is then obtained by the convolution of  $\hat{s}_t$  with an HRF basis, chosen as a classical double gamma with its first derivative and dispersion derivative. Then, an estimate of  $\beta$  is obtained (along with that of  $\phi$ ) by the Cochrane-Orcutt procedure. Finally,  $\hat{h}_t = x_{BF}\hat{\beta}$ . Figures 10 and 11 will report examples of estimated HRF on the case study considered in the paper. Note that  $\hat{h}_t$ , the estimated HRF is not  $\hat{\mu}_t$ , the estimated signal  $\hat{\mu}_t = \hat{X}_t\hat{\beta}$  though some circularity affects procedure of BOLD estimation in [29]. The aim, here, is mainly to estimate the HRF rather than the BOLD signal.

Under the Student- $t$  model in the formulation proposed in [18], described in the introduction of Sect. 2, spontaneous activations are estimated by considering as extreme values those values of  $v_t = y_t - \hat{\mu}_t$  which exceed the threshold  $\max\{\hat{\kappa}u_t\}$  if  $\hat{\kappa} > 1$  or  $\max\{u_t\}$  if  $\hat{\kappa} < 1$  where, under the assumptions of the model,  $v_t$ , the

one-step-ahead prediction error, has a Student- $t$  distribution while  $u_t$ , the martingale different sequence proportional to the score of the Student- $t$  that drives the dynamics of the BOLD signal, has a thin tailed distribution. The selection of the threshold is motivated as follows. In practice,  $u_t$  itself is a thresholded version of  $v_t$ , where extremes are cut off by construction, as  $u_t = v_t / (1 + v_t^2 / (\nu\sigma^2))$ . The parameter  $\kappa$  further regulates the thresholding of  $v_t$ , and thus the robustness of  $\hat{u}_t$ . A natural choice for the threshold then relates to the case when  $\hat{\kappa} > 1$  versus the case when  $\hat{\kappa} \leq 1$ . We choose to take as extremes those values of  $v_t$  which exceed  $u_t$  by keeping the estimated  $\hat{\kappa}u_t$  when  $\hat{\kappa} > 1$  and  $u_t$  otherwise. The opposite choice would lead to a larger number of outlying observations, i.e., of spontaneous activations. The preliminary estimate of the design matrix  $X_t$  is obtained, as in the Gaussian case, by the convolution of  $\hat{s}_t$  with the HRF basis formed by the double gamma with its first derivative and dispersion derivative. An estimate of  $\beta$  is obtained (along with that of all the static parameters of the model) by maximum likelihood in the case when explanatory variables are included in the model, see [18] Sect. 7. Finally,  $\hat{h}_t = x_{BF}\hat{\beta}$ , see Figs. 10 and 11. The estimated HRF  $\hat{h}_t$  may then be considered as an explanatory variable for the filtered signal,  $\mu_t = \hat{h}_t\gamma + \phi\mu_{t-1} + \kappa u_{t-1}$ . As before, focus here was primarily on HRF estimation, but we aim at further pursuing this analysis in future research.

### 3 Illustrative Examples

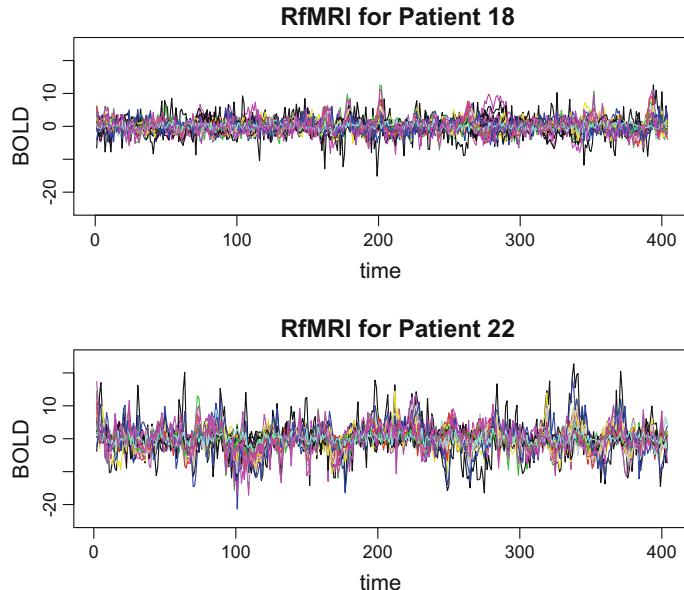
We focus on patients 18 and 22 and on ROI 64, 51, 63, 59, for reasons explained below.

Patient 18 is 46 years old, right-handed and healthy while 22 is 42 years old, right-handed and has a diagnosis of depression with a clinical history of alcohol, cannabis and cocaine abuse. Our aim consists in investigating and comparing brain connections in these two subjects.

According to Fox et al. [13], some brain regions have positive correlations, while other brain regions show negative correlations. In particular, there is a positive correlation among frontal eye field (FEF), intraparietal (IPS) and middle temporal region (MT), while there is a negative correlation among medial prefrontal cortex (MPF), posterior cingulate precuneus (PCC) and lateral parietal cortex (LP). In our dataset, we do not have the same atlas representation of the brain, so we have to find the correspondence between our atlas and the one used in [13]. After this procedure, we analyze the following brain regions:

- **Positive correlation:** ROI 64 (rh-frontal superior, for FEF) and ROI 51 (rh-middle temporal, for MT);
- **Negative correlation:** ROI 63 (rh-rostral middle frontal, for MPF) and ROI 59 (rh-posterior cingulate, for PCC).

In our dataset, the lateral parietal cortex cannot be distinguished from the intraparietal sulcus and this is the reason why we focus only on the correlation between two regions of interest.

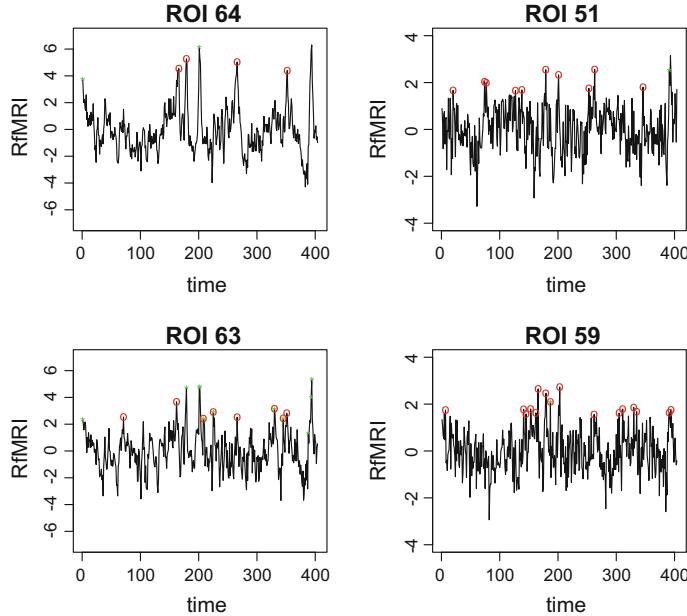


**Fig. 2** Observed RfMRI series at all ROI for subjects 18 and 22

The 70 observed fMRI series for patients 18 and 22, first scan, are represented in Fig. 2. One can immediately notice the difference of range and amplitude of the fluctuations between the two series, being the variability of fMRI of patient 22 larger than that of patient 18.

We proceed to compare the estimated  $\hat{s}_i$  with both methods. Spontaneous events detected in the time series of ROI 64, 51, 63 and 59 for patient 18 are reported in Fig. 3, those detected in the time series of the same ROI for patient 22 are reported in Fig. 4. We remark here that the number of spontaneous activations detected under the hypothesis of a Student- $t$  model for patient 22 drastically drops down if the first extreme observation is removed from the series, compare Figs. 4 and 5. Nevertheless, we illustrate the results of the analysis on the original series including the first observation. The method based on the assumption of a Student- $t$  distribution for the noise is expected to select a smaller number of spontaneous activations than the method based on the Gaussian assumption. However, several series analysed in the dataset turned out to be Gaussian, in which case the two methods give, as expected, similar results.

Despite of the fact that similar peaks are detected (ROI 63 in Fig. 3), we would like to understand whether there are correlations among activations of different brain regions and whether the two methods used for detecting peaks lead to different results. In order to further investigate this idea, we need a concept of distance/similarity for binary vectors. Indeed, standard correlation indices, such as Pearson's correlation index, are not suitable for this case. There are several similarity indices for binary



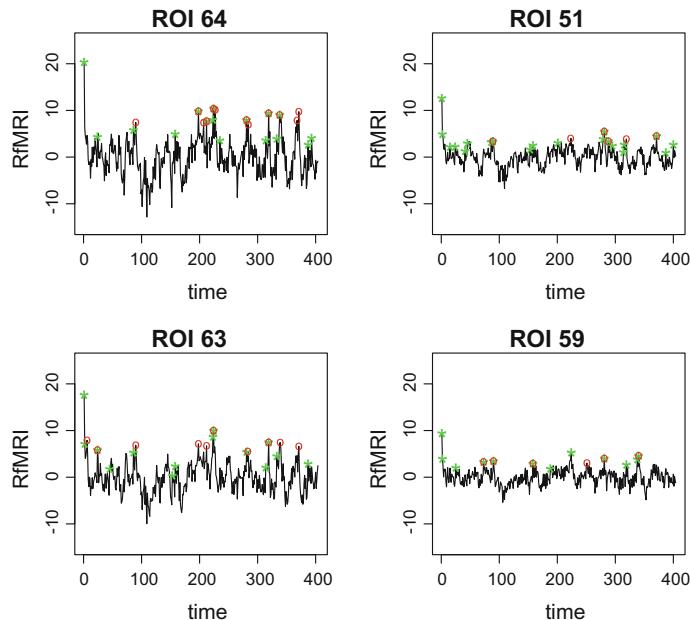
**Fig. 3** Detected spontaneous activations for patient 18 in ROI 64,51,63 and 59. Green stars represent spontaneous activations estimated with robust method, while red dots correspond to the general linear model

vectors in literature, such as Jaccard, Dice, Kulczyńska and Driver and Kroeber (a review is presented in [8]). In this case study, we focus on Dice's similarity index [11] with expression:

$$\frac{2n_{11}}{2n_{11} + n_{10} + n_{01}}.$$

Given  $\hat{s}_t$  related to two different ROIs,  $n_{11}$  is the total number of activations that occurred at the same time in the two ROIs, while  $n_{10}$  and  $n_{01}$  are activations that occurred in one ROI and not in the other. Dice index ranges from 0 to 1. The results for patient 18 are reported in Fig. 6 while those of patient 22 are reported in Fig. 7. In both figures, we notice that computing the Dice index after the application of the robust method leads to a clearer correlation pattern with respect to computing it after the application of the classical method based on linear model.

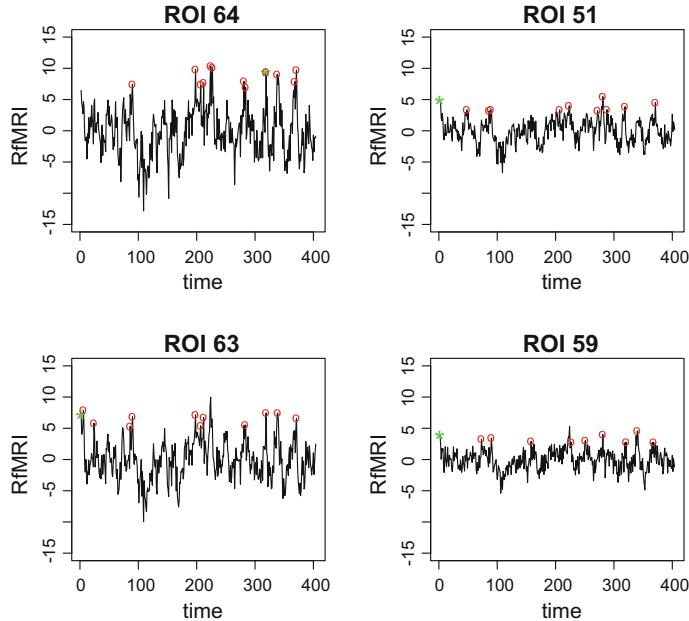
Moreover, the highest value of Dice index recorded for patient 18 according to the robust method allows us to detect the following clusters, that are summarized in Fig. 8:



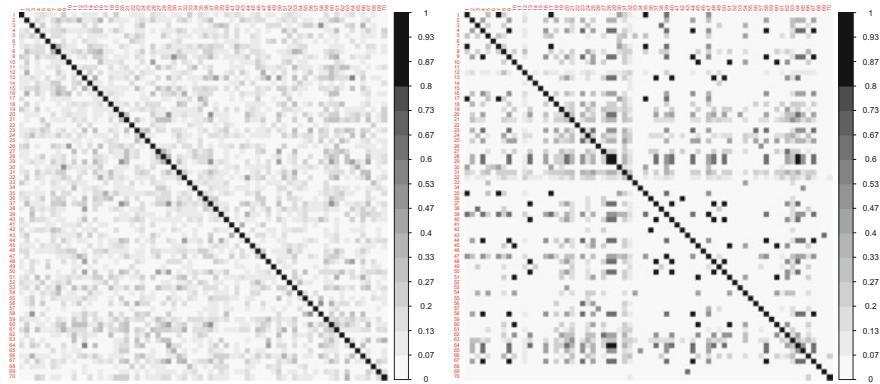
**Fig. 4** Detected spontaneous activations for patient 22 in ROI 64, 51, 63 and 59. Green stars represent spontaneous activations estimated with robust method, while red dots correspond to the general linear model

- lh-unknown, lh-entorhinal, lh-parahippocampal, lh-transversetemporal (ROIs: 1, 7, 17, 35);
- lh-caudalmiddlefrontal, lh-inferiorparietal, rh-inferiorparietal, rh-postcentral, rh-supramarginal (ROIs: 4, 9, 44, 58, 67);
- lh-rostralmiddlefrontal, lh-superiorfrontal, rh-superiorfrontal (ROIs: 28, 29, 64);
- lh-lateralorbitofrontal, rh-bankssts, rh-corpuscallosum, rh-lateralorbitofrontal, rh-medialorbitofrontal (ROIs: 13, 37, 40, 48, 50);
- lh-inferiortemporal, rh-inferiortemporal, rh-middletemporal, rh-precentral (ROIs: 10, 45, 51, 60);
- rh-unknown, rh-entorhinal (ROIs: 36, 42).

Note that ‘lh’ and ‘rh’ stand respectively for the left and right hemispheres. It is interesting to notice that symmetrical ROIs belong to the same groups, such as inferiorparietal, second group, superiorfrontal, third group, lateralorbitofrontal, fourth group and inferiortemporal, fifth group. Moreover, this detection is graphically confirmed by the recorded BOLD, indeed BOLD signals from ROIs of the same group show similar trend (Fig. 8). On the other hand, the highest value of Dice similarity index with the standard approach is 0.55, between lh-lateralorbitofrontal and rh-superiorfrontal.

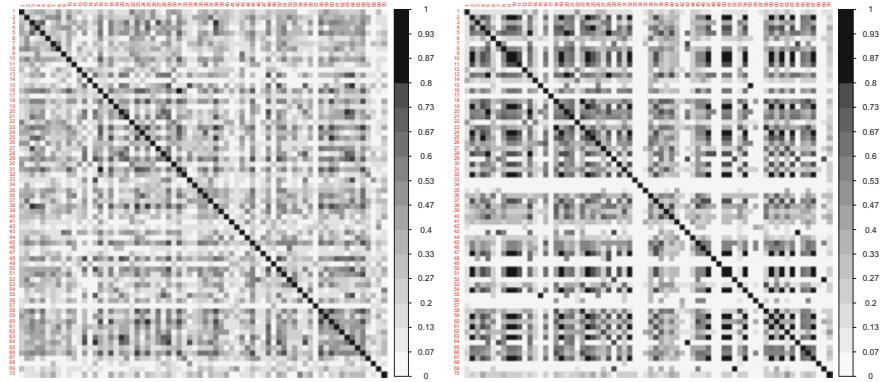


**Fig. 5** Detected spontaneous activations for patient 22 without the first observation in ROI 64,51,63 and 59. Green stars represent spontaneous activations estimated with robust method, while red dots correspond to the general linear model

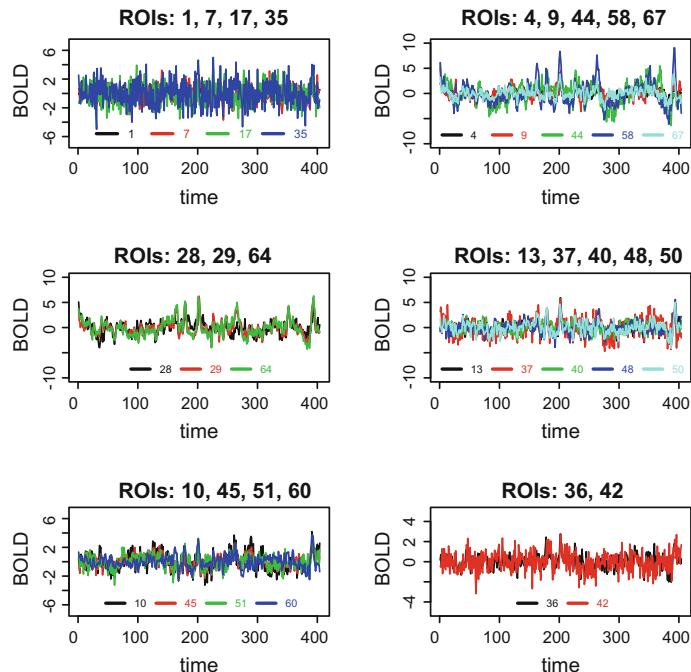


**Fig. 6** Similarity matrix based on Dice distance for patient 18, classical methods (left) robust method (right)

The same analysis can be done on patient 22. For this specific patient, a high peak is recorded in the first time frames as discussed in Sect. 3. If we consider the whole time series, the maximum value of Dice index is recorded for only one group of time series. This group is composed by those RfMRI measures with only one peak at the

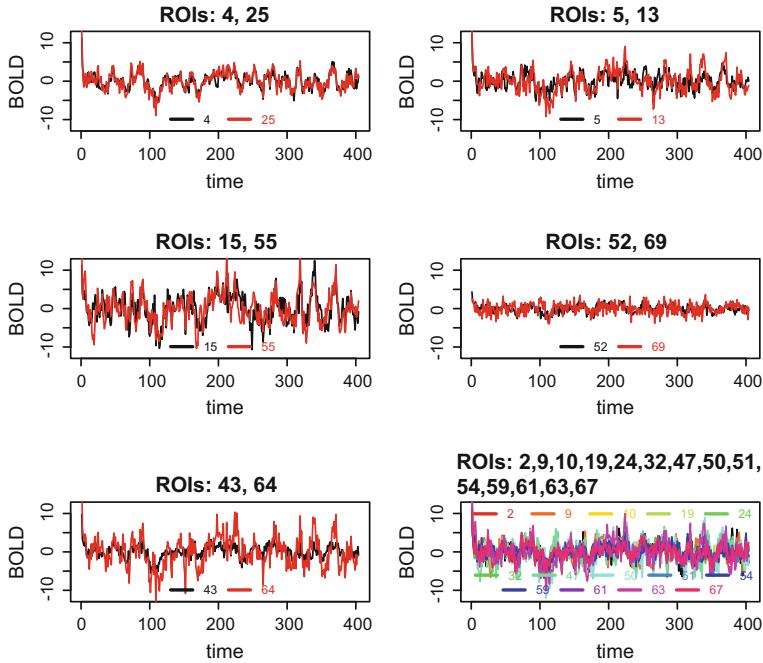


**Fig. 7** Similarity matrix based on Dice distance for patient 22, classical methods (left) robust method (right)



**Fig. 8** BOLD measured in groups with the highest Dice's index for patient 18

beginning. The obtained result is not as informative as the one achieved by detecting peaks on the reduced time series (time series without the first time frames). Focusing on the reduced time series, the highest value of Dice index according to the robust method allows us to detect the following clusters, that are summarized in Fig. 9:

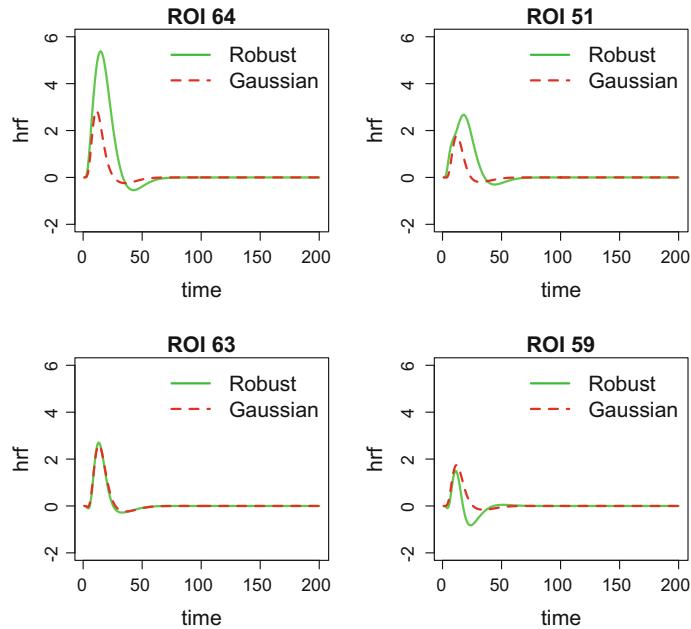


**Fig. 9** BOLD measured in groups with the highest Dice's index for patient 22

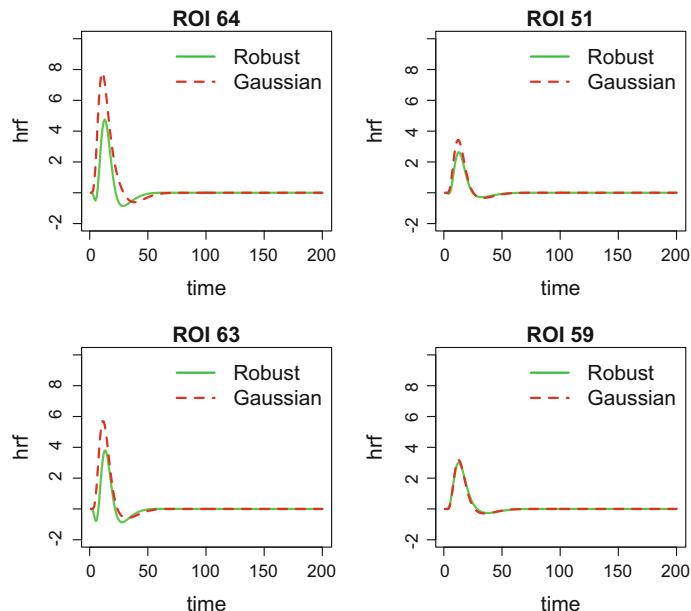
- lh-caudalmiddlefrontal, lh-precentral (ROIs: 4, 25);
- lh-corpuscallosum, lh-lateralorbitofrontal (ROIs: 5, 13);
- lh-medialorbitofrontal, rh-parsorbitalis (ROIs: 15, 55);
- rh-parahippocampal, rh-temporalpole (ROIs: 52, 69);
- rh-fusiform, rh-superiorfrontal (ROIs: 43, 64);
- lh-bankssts, lh-inferiorparietal, lh-inferiortemporal, lh-parsopercularis, lh-posteriorcingulate, lh-supramarginal, rh-lateraloccipital, rh-medialorbitofrontal, rh-middletemporal, rh-parsopercularis, rh-posteriorcingulate, rh-precuneus, rh-rostralmiddlefrontal, rh-supramarginal (ROIs: 2, 9, 10, 19, 24, 32, 47, 50, 51, 54, 59, 61, 63, 67).

The detected pattern is difficult to interpret and requires further investigation.

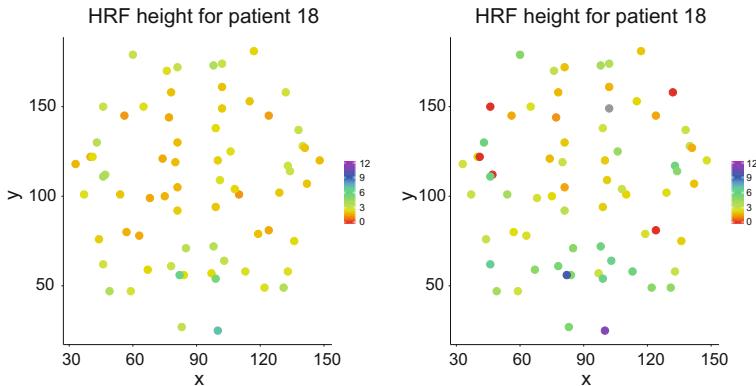
To conclude, we report in Fig. 10 the estimated HRF in ROIs 64, 51, 63 and 59 for patient 18 and in Fig. 11 the corresponding estimates for patient 22. It is immediate to see that the time-to-peak is similar among all the considered regions, while both the height and the FWHM may be quite different.



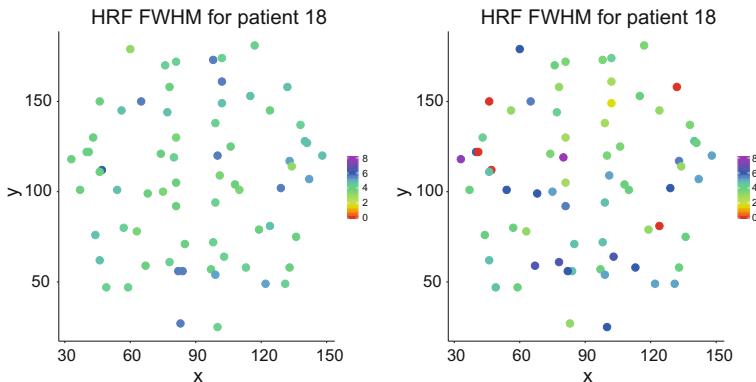
**Fig. 10** Estimated HRF with both methods for patient 18 in ROI 64, 51, 63 and 59



**Fig. 11** Estimated HRF with both methods for patient 22 in ROI 64, 51, 63 and 59

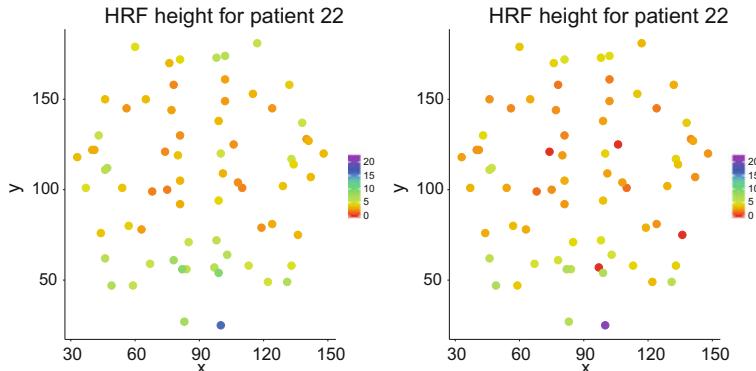


**Fig. 12** Estimated height parameter across ROIs for patient 18 with the classical (left) and robust method (right). This plot shows the estimated height parameter in xy-section of the brain that is obtained by projecting all ROIs centroids in a horizontal plane

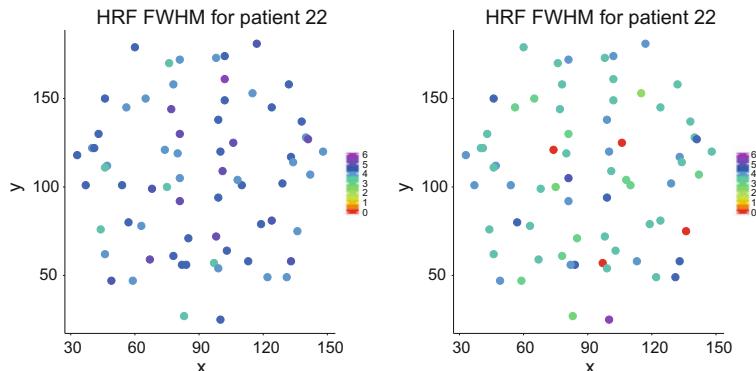


**Fig. 13** Estimated FWHM parameter across ROIs for patient 18 with the classical (left) and robust method (right). This plot shows the estimated FWHM parameter in xy-section of the brain that is obtained by projecting all ROIs centroids in a horizontal plane

In order to have a clearer picture of these differences we plotted the estimated values of these parameters across the regions of interest. The estimated HRF height of patient 18 in a xy-section of brain (horizontal plane of brain) is reported in Fig. 12. The estimated HRF FWHM of patient 18 in a xy-section of brain is reported in Fig. 13. The corresponding pictures for patient 22 are in Figs. 14 and 15. A pattern of height distribution across ROIs seems to be identifiable, but it is still under investigation.



**Fig. 14** Estimated height parameter across ROIs for patient 22 with the classical (left) and robust method (right). This plot shows the estimated height parameter in xy-section of the brain that is obtained by projecting all ROIs centroids in a horizontal plane



**Fig. 15** Estimated FWHM parameter across ROIs for patient 22 with robust method and a more restrictive cut-off. This plot shows the estimated FWHM parameter in xy-section of the brain that is obtained by projecting all ROIs centroids in a horizontal plane

## 4 Concluding Remarks and Further Developments

The paper was concerned with the problem of detecting spontaneous activations in resting state fMRI time series and of estimating the hemodynamic response function. Two methods have been considered, one based on a classical, Gaussian, assumption for the data generating process of fMRI data and another based on the assumption that the data may be generated by an heavy tailed distribution. The assumption of a heavy tailed distribution for RfMRI series is supported by the fact that several of the series in the dataset that we have used for our empirical illustrations have shown evidence of extreme values, confirmed by high kurtosis, indicating a violation of the Gaussian assumption. The cut-off threshold used for identifying an observation as

an extreme one, i.e. for detecting a spontaneous event, was rather restrictive, so that fewer spontaneous activations are detected than in the Gaussian case. Methods for determining an optimal threshold may be the object of future investigation.

Future works aim at taking into account the spatial correlation of the data, e.g., by considering a locally anisotropic stationary spatial model, as in the work of Castruccio et al. [6], or by considering spatio-temporal score-driven models, as in [4, 7].

**Acknowledgements** We thank two anonymous referees for their insightful comments and Federico Crescenzi, Michele Peruzzi and Alexios Polymeropoulos for constructive discussions at the Certosa di Pontignano, Bologna and Milano during the initial stages of the current work. We would like to thank Antonio Canale, Daniele Durante, Lucia Paci and Bruno Scarpa for bringing us together and providing us with the challenging dataset analysed in the paper. These data are provided by Greg Kiar and Eric Bridgeford from NeuroData at Johns Hopkins University, who graciously pre-processed the raw DTI and R-fMRI imaging data available at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html), using the pipelines ndmg and C-PAC. We would also like to thank all the participants of the StartUp Research event held at the Certosa di Pontignano on June 25-27, 2017, for the stimulating and nice discussions.

## References

1. Aston, J., Kirsch, C.: Evaluating stationarity via change-point alternatives with applications to fMRI data. *Ann. Appl. Statist.* **6**(4), 1906–1948 (2012)
2. Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S.: Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* **34**(4), 537–541 (1995)
3. Biswal, B.: Toward discovery science of human brain function. *PNAS* **107**(10), 4734–4739 (2010)
4. Blasques, F., Koopman, S.J., Lucas, A., Schaumburg, J.: Spillover dynamics for systemic risk measurement using spatial financial time series models. *J. Econom.* **195**(2), 211–223 (2016)
5. Bullmore, E., Fadili, J., Breakspear, M., Salvador, R., Suckling, J., Brammer, M.: Wavelets and statistical analysis of functional magnetic resonance images of the human brain. *Statist. Methods Med. Res.* **12**(5), 375–399 (2003)
6. Castruccio, S., Ombao, H., Genton, M. G.: A scalable multi-resolution spatio-temporal model for brain activation and connectivity in fMRI data. *Biometrics* (2018)
7. Catania, L., Billé, A.G.: Dynamic spatial autoregressive models with autoregressive and heteroskedastic disturbances. *J. Appl. Econom.* (2017)
8. Choi, S.S., Cha, S.H., Tappert, C.C.: A survey of binary similarity and distance measures. *J. Syst. Cybern. Inf.* **8**, 43–48 (2010)
9. Creal, D., Koopman, S., Lucas, A.: A dynamic multivariate heavy-tailed model for the time-varying volatility and correlations. *J. Bus. Econom. Statist.* **29**, 552–563 (2011)
10. D’Esposito, M., Deouell, L., Gazzaley, A.: Alterations in the bold fMRI signal with ageing and disease: a challenge for neuro imaging. *Nature Rev. Neurosci.* (4), 863–872 (2003)
11. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302 (1945)
12. Fox, M.D., Raichle, M.E.: Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Rev. Neurosci.* **8**(9), 700 (2007)
13. Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E.: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U. S. A.* **102**(27), 9673–9678 (2005)

14. Friston, K.J., Fletcher, P., Josephs, O., Holmes, A.P., Rugg, M., Turner, R.: Event-related fMRI: characterizing differential responses. *Neuroimage* **7**(1), 30–40 (1998)
15. Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.: Statistical parametric maps in functional imaging: a general linear approach. *Huma. Brain Mapp.* **2**(4), 189–210 (1994)
16. Glover, G.H.: Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* **9**(4), 416–429 (1999)
17. Handwerker, D.A., Ollinger, J.M., D’Esposito, M.: Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* **21**(4), 1639–1651 (2004)
18. Harvey, A., Luati, A.: Filtering with heavy tails. *J. Am. Statist. Assoc.* **109**(507), 1112–1122 (2014)
19. Harvey, A.C.: *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press (2013)
20. Henson, R., Friston, K.: Convolution models for fMRI. Statistical parametric mapping: the analysis of functional brain images, pp. 178–192 (2007)
21. Kruggel, F., von Cramon, D.Y.: Temporal properties of the hemodynamic response in functional MRI. *Huma. Brain Mapp.* **8**(4), 259–271 (1999)
22. Lange, N., Zeger, S.L.: Non-linear fourier time series analysis for human brain mapping by functional magnetic resonance imaging. *J. Royal Statist. Soc. Ser. C (Appl. Statist.)* **46**(1), 1–29 (1997)
23. Lindquist, M.A.: The statistical analysis of fMRI data. *Statist. Sci.* **23**(4), 439–464 (2008)
24. Lund, T.E.: Non-white noise in fMRI: Does modelling have an impact? *Neuroimage* **29**(4), 1639–1651 (2006)
25. Poldrack, R.A., Mumford, J.A., Nichols, T.E.: *Handbook of Functional MRI data Analysis*. Cambridge University Press (2011)
26. Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M.: Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage* **14**(6), 1370–1386 (2001)
27. Worsley, K.J., Liao, C., Aston, J., Petre, V., Duncan, G., Morales, F., Evans, A.: A general statistical analysis for fMRI data. *Neuroimage* **15**(1), 1–15 (2002)
28. Worsley, K.: Detecting activation in fMRI data. *Statist. Methods Med. Res.* **12**(5), 401–418 (2003)
29. Wu, G.R., Liao, W., Stramaglia, S., Ding, J.R., Chen, H., Marinazzo, D.: A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. *Med. Image Anal.* **17**(3), 365–374 (2013)

# Hierarchical Spatio-Temporal Modeling of Resting State fMRI Data



Alessia Caponera, Francesco Denti, Tommaso Rigon,  
Andrea Sottosanti and Alan Gelfand

**Abstract** In recent years, state of the art brain imaging techniques like Functional Magnetic Resonance Imaging (fMRI), have raised new challenges to the statistical community, which is asked to provide new frameworks for modeling and data analysis. Here, motivated by resting state fMRI data, which can be seen as a collection of spatially dependent functional observations among brain regions, we propose a parsimonious but flexible representation of their dependence structure leveraging a Bayesian time-dependent latent factor model. Adopting an assumption of separability of the covariance structure in space and time, we are able to substantially reduce the computational cost and, at the same time, provide interpretable results. Theoretical properties of the model along with identifiability conditions are discussed. For model fitting, we propose a MCMC algorithm to enable posterior inference. We illustrate our work through an application to a dataset coming from the ENKIRS project, discussing the estimated covariance structure and also performing model selection along with network

---

A. Caponera

Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy  
e-mail: [Alessia.caponera@uniroma1.it](mailto:Alessia.caponera@uniroma1.it)

F. Denti

Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy  
e-mail: [f.denti2@campus.unimib.it](mailto:f.denti2@campus.unimib.it)

T. Rigon (✉)

Department of Decision Sciences, Bocconi University, Milan, Italy  
e-mail: [tommaso.rigon@phd.unibocconi.it](mailto:tommaso.rigon@phd.unibocconi.it)

A. Sottosanti

Department of Statistical Sciences, University of Padova, Padua, Italy  
e-mail: [sottosanti@stat.unipd.it](mailto:sottosanti@stat.unipd.it)

A. Gelfand

Department of Statistical Science, Duke University, Durham, North Carolina, USA  
e-mail: [alan@stat.duke.edu](mailto:alan@stat.duke.edu)

analysis. Our modeling is preliminary but offers ideas for developing fully Bayesian fMRI models, incorporating a plausible space and time dependence structure.

**Keywords** Bayesian factor analysis · Gaussian processes  
Low-rank factorizations · Separable models

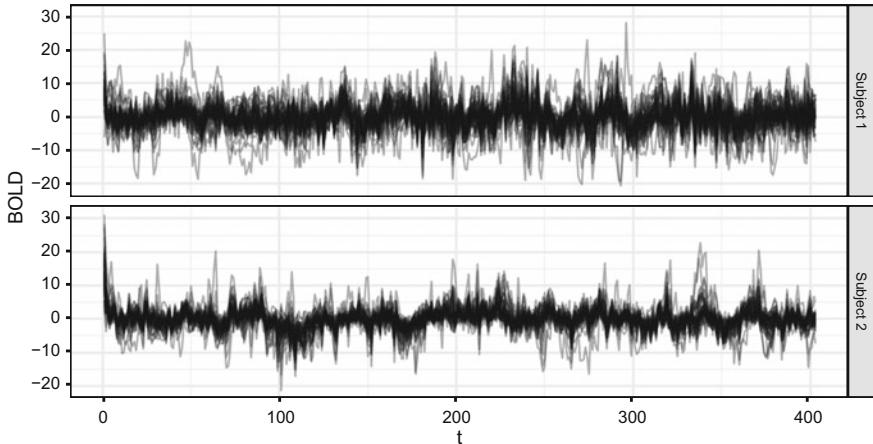
## 1 Introduction

Functional magnetic resonance imaging (fMRI) is an imaging technique which allows the study of human brain activities without being invasive. Such a technique provides a high resolution 3D image reconstruction of a human brain, starting from the blood-oxygen-level dependent (BOLD) signal. The BOLD value is the difference in magnetization between oxygenated and deoxygenated blood, arising from changes in regional cerebral blood flow. In particular, the data at our disposal consist of a collection of BOLD signals obtained from a resting state functional magnetic resonance imaging (rs- fMRI) session. This means that the subjects were not performing any explicit task during the scan. Refer, for instance, to [1–4] for detailed discussions on rs- fMRI data, statistical techniques commonly employed, and medical implications.

From a modeling perspective, what emerges from a rs- fMRI scan is a collection of spatially dependent functional observations. This kind of data collection has encouraged the development of suitable statistical techniques, and indeed, several novel spatio-temporal and dynamic models have been proposed (e.g., [5–10]). Within the Bayesian framework, comprehensive reviews of the main statistical methodologies employed for fMRI data are given in [11, 12].

Since we are proposing a preliminary specification here, we focus on the model of one subject at time, i.e., of a single brain, which is usually referred to as single-subject analysis. Although such an approach does not account for borrowing of information across subjects, it simplifies the modeling process and the related estimation procedures. An early reference to single-subject analysis is given by [13], who propose a general linear model to learn about the blood activity of a single brain. Several contributions appeared afterwards in the context of single-subject modeling and we mention here just a few. For example, the authors of [7] specify a Gaussian random field to capture the spatial correlation, while in [5] the spatial dependence is induced through a hierarchical specification of the parameters.

One of the main goals in the analysis of rs- fMRI data is to study the complex covariance structure between brain regions [3, 14]. In this paper, we propose a Bayesian factor model for fMRI data which is based on the structural assumption of separability. This means that, with regard to the dependence in brain activity across regions, we assume that the covariance structure can be split into two multiplicative components: the spatial and the temporal one. Mainly motivated by the high computational cost that would arise by using a non-separable specification, separability has been employed in several fMRI applications (e.g. [6, 7, 15]). The model presented in this paper benefits from this simplifying assumption, which, in addition,



**Fig. 1** BOLD functional activities for subjects  $i = 1, 2$ , for the  $L$  brain regions of the [16] parcellation and  $t = 1, \dots, 403$ , after the C- PAC pre-processing

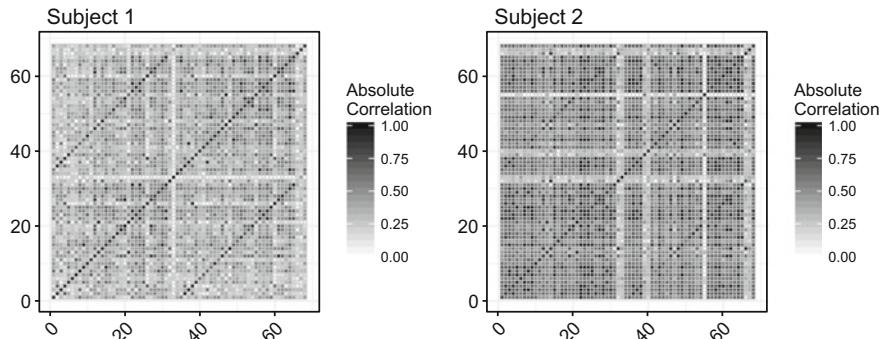
provides interpretable inferential results. This allows the assessment of functional connectivity across brain regions.

The paper is organized as follows. In Sect. 2, we introduce the RS- fMRI dataset at our disposal and we conduct some explorative analysis. In Sect. 3, we specify a single-subject Bayesian factor model for the blood functional activity, which accounts also for temporal dependence. In Sect. 4 we present a Markov Chain Monte Carlo (MCMC) for fitting the proposed model. In Sect. 5, we discuss the performance of our model and we present some empirical results. Concluding remarks are given in Sect. 6.

## 2 The RS- fMRI Dataset

Our dataset comes from the pilot study of the Enhanced Nathan Kline Institute-Rockland Sample project (ENKIRS), which aims at providing a publicly available large sample of multimodal neuroimaging data. Comprehensive information about the project can be found at the link [http://fcon\\_1000.projects.nitrc.org/indi/CorR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CorR/html/nki_1.html). From the original multimodal imaging dataset, we retained the BOLD values of two different subjects, which were randomly chosen among the patients. The BOLD values refer to the  $L = 68$  brain regions of the Desikan atlas parcellation [16], equally divided into the left and right hemispheres, discarding the two regions labeled as unknown.

Each measurement is a functional observation composed of  $T = 403$  equally spaced BOLD values, with a lag of approximately 1400 ms, meaning that our dataset comprises two matrices of size  $L \times T$ , for individuals  $i = 1, 2$ . From the original dataset, two BOLD values were discarded since they were missing. The BOLD functional activities, displayed in Fig. 1, are obtained from the raw RS- fMRI scans through

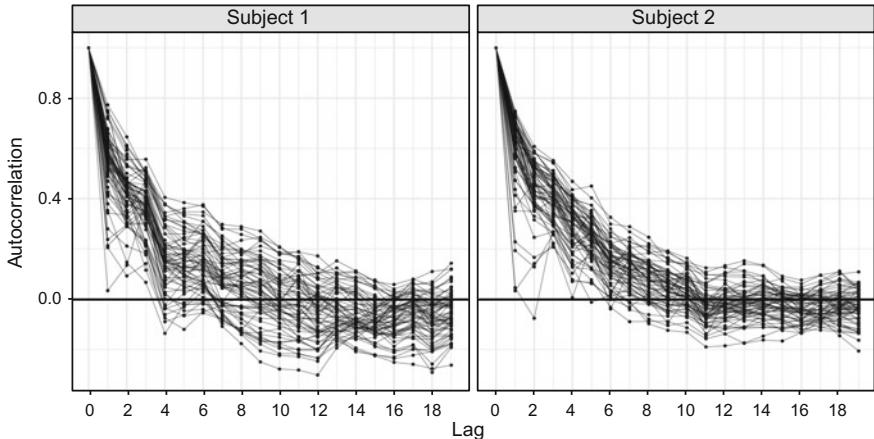


**Fig. 2** Absolute value of the Pearson correlation coefficients among the BOLD functional activities of the 68 brain regions, for subjects  $i = 1, 2$ . Values from 1–34 refer to the left hemisphere, whereas the remaining 35–68 refer to the right hemisphere

an automated pipeline called C-PAC, whose details can be found at <https://fcp-indi.github.io>.

As shown in Fig. 1, the set of regional BOLD functional activities can be regarded as multiple realizations of continuous functions. That is, BOLD, though continuous in time at each region, is evaluated on a finite grid of time  $t = 1, \dots, T$ , while brain regions are specified discretely, being obtained from the Desikan parcellation [16]. There is a considerable statistical literature in spatio-temporal modeling in continuous space and time (see e.g., [17]), particularly in the Bayesian setting [18]. However, data over continuous time and discrete space is rather uncommon in spatio-temporal applications, as pointed out in [19]. In particular, our data should not be modeled, at least in principle, through classical multivariate time series models, since the BOLD activities are continuous in time. Refer for instance to (Chap. 1, [20]) about the use of continuous models for functional observations. Moreover, our data cannot be modeled via standard functional data analysis techniques because some sort of dependence across brain regions is expected. There is need for general modeling methodology for the analysis of this type of RS-fMRI data. We aim to partially fill this gap by introducing a simple spatio-temporal model in the continuous-time and discrete-space framework. Then, we apply it to the RS-fMRI data.

As already mentioned, one of the main goals in the analysis of RS-fMRI data is to study the functional connectivity, e.g. dependence, between brain regions [3, 14]. A simple approach consists in computing the Pearson correlation coefficients between BOLD functional activities, treating the BOLD values as if they were independent over time [2]. The correlation coefficients for subjects  $i = 1, 2$  are shown in Fig. 2. We argue that this strategy, although useful in an explorative phase, could lead to misleading inferential results, for instance revealing fictitious relationships which are due to temporal dependence. In fact, BOLD functional activities are characterized by a non-negligible amount of autocorrelation, as evidenced in Fig. 3, suggesting that Pearson coefficients should be, at the very least, interpreted with care. Nonetheless, correlation matrices like those in Fig. 2 provide an interpretable picture of the BOLD



**Fig. 3** Autocorrelation functions for brain regions  $l = 1, \dots, 68$  and for subjects  $i = 1, 2$

functional connectivity. Additionally, dichotomized versions of these correlations are often at the basis of network analyses for functional connectivity [21, 22]. We aim to preserve this simple structure by seeking a model that naturally leads to an alternative estimate of such a correlation structure, but also takes into account the temporal component.

Some additional difficulties arise when trying to model the spatial component of our RS-fMRI dataset. In particular, areally referenced temporal processes typically rely on some notion of distance, or neighborhood, between different regions, whose definition crucially impacts the results of the analysis. However, given the natural complexity of the brain morphometry, unavoidable questions about the choice of such a distance can be raised [23]. Although there is some evidence that connectivity among brain regions rapidly decays as a function of the Euclidean distance [23, 24], this is often a crude approximation. For instance, as shown in Fig. 2, high levels of connectivity characterize symmetric pairs of brain regions, which are far apart in terms of Euclidean distance. In order to avoid potential misspecification issues, we do not rely on any notion of physical distance between brain regions. Thus, the spatial structure of the BOLD functional activities is reconstructed entirely from the data, without imposing the brain morphometry. This is not to say that there is no potential information in terms of proximity of regions. If a suitable measure reflecting the foregoing caveats were developed, it would provide valuable information and it could improve estimation performance.

### 3 Modeling and Theory

#### 3.1 Low-Rank Multivariate Processes

Consistent with the discussion in Sect. 2, we propose a hierarchical model for rs-fMRI data which (i) accounts for both the spatial and temporal aspects, specifying Gaussian processes for time, and latent factor models for the spatial dimension; (ii) allows a simple interpretation of the functional connectivity among brain regions in terms of a suitable covariance matrix; (iii) avoids misspecification issues by placing few assumptions on the spatial structure. Again, we focus on single-subject models, which means that the two different individuals  $i = 1, 2$  are treated separately and independently, having in common only the model structure. For the sake of notation, we omit the subject index  $i = 1, 2$  and we describe a model for a generic brain. We also note that, with a single subject, we cannot build a regression explanation of response since we cannot include individual level covariate information.

We aim to describe the joint behavior of realizations from a  $L$ -dimensional stochastic process, i.e., the collection of the BOLD functional activities. Formally, we will denote the  $L$ -dimensional stochastic process as  $\mathbf{B}(t)$ , whose entries are the BOLD functional activities  $B_l(t)$ , for  $l = 1, \dots, L$ . We assume a customary additive error structure, that is

$$\mathbf{B}(t) = \mathbf{Z}(t) + \boldsymbol{\epsilon}(t), \quad (1)$$

where  $\boldsymbol{\epsilon}(t)$  is a  $L$ -dimensional pure error, and  $\mathbf{Z}(t)$  is a  $L$ -dimensional process, which we refer to as the mean process. Specifically,  $\boldsymbol{\epsilon}(t)$  is a Gaussian white noise process with variance  $\sigma^2$ , whose entries are independent Gaussian random variables over time and brain regions.<sup>1</sup> Notice that no intercept term is included in specification (1), since, as shown in Fig. 1, the dataset is centered around zero during the C-PAC pipeline.

The overarching goal of our contribution is to infer functional connectivities among brain regions [3]. Therefore, consistent with the available literature and with the descriptive analysis in Sect. 2, the components of the  $L$ -dimensional process  $\mathbf{Z}(t)$  should not be modeled as independent realizations. Moreover, the high dimensionality of our data calls for parsimonious representations, which can be obtained for instance via low-rank approximations. As a notable example, covariance regression models [25, 26] address similar issues and assume that the mean process  $\mathbf{Z}(t)$  can be decomposed as follow:

$$\mathbf{Z}(t) = \mathbf{A}(t)\mathbf{V}(t), \quad (2)$$

where  $\mathbf{A}(t)$  is a  $L \times K$  time varying factor loading matrix, and  $\mathbf{V}(t)$  is a  $K$ -dimensional vector whose entries are independent Gaussian processes—called latent factors in this context. The dimensionality reduction is performed by fixing some

---

<sup>1</sup>Richer modeling might allow heterogeneity in variances, e.g., across regions but we do not consider that here.

$K \ll L$ , e.g.  $K = 3$  or  $K = 5$ . The covariance regression model of [26] is a flexible model for multidimensional stochastic processes, having large support and a familiar interpretation in terms of Bayesian factor models, for any fixed time. Moreover, with discrete time, the covariance regression model could be formally related to the class of dynamic latent factor models (Chap. 10 [27]).

### 3.2 A Time-Dependent Latent Factor Model

We simplify (2) and we set  $\mathbf{A}(t) = \mathbf{A}$ , that is, the matrix  $\mathbf{A}$  is now constant over time and

$$\mathbf{Z}(t) = \mathbf{A}\mathbf{V}(t). \quad (3)$$

Hence,  $\mathbf{A}$  is a  $L \times K$  factor loading matrix. Although such an assumption reduces the global flexibility of the covariance regression model in Eq.(2), it allows the factor loading matrix  $\mathbf{A}$  to be interpreted as a simple measure of dependence, e.g., connectivity, among brain regions, a key feature of our analysis.

In Sect. 5 we provide some empirical support for the factorization (3) as a reasonable assumption for modeling RS- fMRI data while in Sect. 6 we discuss possible extension to the non-stationary case. Additionally, decomposition (3) formally relates our model to the class of latent factor models, which have been used as a dimension reduction tool for instance in genomic applications [28, 29]. Thus, our model can be regarded as a time-dependent extension of a Bayesian factor model, in which the latent factors  $\mathbf{V}(t)$  are independent random functions of time rather than independent draws.

Gaussian processes [30] are a flexible class of stochastic processes to provide random realizations within the space of functions over a specified domain. Therefore, they are a suitable candidate for modeling the time-dependent latent factors  $\mathbf{V}(t)$ . We suppose that the components of  $\mathbf{V}(t)$  are independent and identically distributed Gaussian processes  $\text{GP}(0, \kappa\rho(t, t'))$ , with zero mean and correlation function  $\rho(t, t')$ . As we will discuss in Sect. 3.3, for identifiability purposes we assume  $\kappa = 1$ . Independence among the Gaussian latent factors and the restricted factorization (3) imply a multivariate Gaussian distribution for the mean process  $\mathbf{Z}(t)$  evaluated at a fixed time  $t_0$ , that is

$$\mathbf{Z}(t_0) \sim \mathcal{N}_L(0, \boldsymbol{\Sigma}_A), \quad \text{for any fixed } t_0, \quad (4)$$

with  $\boldsymbol{\Sigma}_A = \mathbf{A}\mathbf{A}^\top$ , which does not depend on time. The role of  $\mathbf{A}$  is now clearer since it can be viewed as the square root of the covariance matrix  $\boldsymbol{\Sigma}_A$ . We remark that  $\boldsymbol{\Sigma}_A$  is singular, being of rank  $K \ll L$ . In turn, this implies that  $\mathbf{Z}(t_0)$  for any fixed  $t_0$  would be a degenerate multivariate Gaussian, lying in a subspace of dimension  $K$ . Factorization (3) effectively induces dependence among the components of the mean process, e.g. among brain regions, but implicitly enforces some form of stationarity, since the spatial dependence structure is constant over time. Such an assumption is

discussed in depth for instance in [31], who suggest that it might be worth looking at non-stationary models to obtain a more complete picture of the phenomenon. However, as argued by [31] themselves, stationarity is also convenient in order to prevent the model from becoming vastly more complex. The stationary temporal dependence of our model can be appreciated by observing the covariance matrix between  $\mathbf{Z}(t)$  and  $\mathbf{Z}(t')$ , that is, the so-called cross-covariance matrix

$$\text{Cov}(\mathbf{Z}(t), \mathbf{Z}(t')) = \rho(t, t') \boldsymbol{\Sigma}_A, \quad t \neq t', \quad (5)$$

whose limit as  $|t - t'| \rightarrow 0$  is  $\boldsymbol{\Sigma}_A = \mathbf{A}\mathbf{A}^\top$ , consistently with Eq. (4). Moreover, for  $t \neq t'$  we have  $\text{Cov}(\mathbf{B}(t), \mathbf{B}(t')) = \text{Cov}(\mathbf{Z}(t), \mathbf{Z}(t'))$ . Thus, the cross-covariance in (5) has an appealing interpretation: dependence between BOLD values is multiplicatively adjusted according to temporal proximity.

In practice, we observe the BOLD functional activities only over a finite grid of times  $t = 1, \dots, T$ ; we denote with  $\mathcal{Z}$  the  $L \times T$  matrix containing the values of  $\mathbf{Z}(t)$  over this time grid. Also, let  $\mathcal{B}$  be a  $L \times T$  observed data matrix having entries  $B_l(t)$ , for  $t = 1, \dots, T$ . We can re-express the model of Eqs. (1), (3) and (4) in terms of matrix Gaussian distributions [32], evaluated over the finite time grid:

$$(\mathcal{B} | \mathcal{Z}, \sigma^2) \sim N_{L,T}(\mathcal{Z}, \sigma^2 I_{L \times L}, I_{T \times T}), \quad (6)$$

$$(\mathcal{Z} | \mathbf{A}) \sim N_{L,T}(0, \boldsymbol{\Sigma}_A, \boldsymbol{\Sigma}_T), \quad (7)$$

where  $\boldsymbol{\Sigma}_T$  denotes the Gram-matrix obtained by evaluating the covariance functions  $\rho(t, t')$  over the finite grid  $t = 1, \dots, T$ . Notice that the stationarity assumption translates into a separability assumption in the finite-dimensional setting, since we have

$$(\text{vec}(\mathcal{Z}) | \mathbf{A}) \sim N(0, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_T \otimes \boldsymbol{\Sigma}_A. \quad (8)$$

This convenient separability result is well described in the spatial literature on multivariate spatial processes: see for instance [18]. Factorization  $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_T \otimes \boldsymbol{\Sigma}_A$  has relevant benefits: it provides a parsimonious representation of the covariance matrix  $\boldsymbol{\Sigma}$  and it facilitates numerical computations. Notice that under the separability assumption the marginal distribution of the rows of  $\mathcal{Z}$  is a multivariate Gaussian with covariance matrix  $\boldsymbol{\Sigma}_T$  and, symmetrically, the columns of  $\mathcal{Z}$  follow a multivariate Gaussian with covariance  $\boldsymbol{\Sigma}_A$ . In other words, the dependence structure over time does not depend on the brain regions, and vice versa.

### 3.3 Identifiability

Without further restrictions, the model described in Eqs. (6) and (7) is not identified. There are two sources of non-identifiability which can be handled by imposing some constraints on the parameters. Notice that  $\mathbf{A}$  appears in Eq. (7) only in terms of

product with its transpose. This means, for instance, that for any orthogonal matrix  $\mathbf{Q}$  such that  $\mathbf{Q}\mathbf{Q}^\top = I_{K \times K}$  we get

$$\mathbf{A}\mathbf{A}^\top = \mathbf{A}\mathbf{Q}\mathbf{Q}^\top\mathbf{A}^\top = \tilde{\mathbf{A}}\tilde{\mathbf{A}}^\top,$$

where  $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{Q}$ . Thus, it is not possible to discriminate between a model with parameter  $\mathbf{A}$  and another with parameter  $\tilde{\mathbf{A}}$ , since they lead to exactly the same likelihood. Thus, we let  $\mathbf{A}$  to be lower triangular with positive diagonal, as commonly done in coregionalization models in spatial statistics [18]. To avoid confusions: since  $\mathbf{A}$  is a  $L \times K$  rectangular matrix, by lower triangular with positive diagonal we mean that the elements  $a_{lk}$  of  $\mathbf{A}$  are such that  $a_{lk} = 0$  for  $k > l$  and  $a_{kk} > 0$  for  $k = 1, \dots, K$ . Thanks to the Cholesky decomposition for positive semi-definite matrices, under these assumptions the matrix  $\mathbf{A}$  is a Cholesky factor, which uniquely identifies  $\Sigma_A = \mathbf{A}\mathbf{A}^\top$ .

The second source of non-identifiability concerns the scale of the covariance matrices in Eq. (7). For any positive constant  $c \in \mathbb{R}^+$  we have that

$$\Sigma_A \otimes \Sigma_T = (c\Sigma_A) \otimes \left( \frac{1}{c}\Sigma_T \right) = \tilde{\Sigma}_A \otimes \tilde{\Sigma}_T,$$

which leads again to non-identifiability. To overcome this difficulty, we set the trace  $\text{Tr}(\Sigma_T)$  equal to some constant, which can be easily obtained by letting the scaling parameter of the covariance function  $\kappa$  to be equal to one. Under these constraints, the model is fully identified. As an alternative, one could impose the first, or the last, diagonal entry to be equal to one.

### 3.4 Prior Specification

We conduct inference within the Bayesian framework and therefore we need to specify prior distributions for both  $\mathbf{A}$  and  $\sigma^2$ . In the latter case, we choose an inverse gamma prior for the residual variance, that is

$$\sigma^{-2} \sim \text{Ga}(a_\sigma, b_\sigma), \quad (9)$$

with  $a_\sigma, b_\sigma > 0$  some fixed hyperparameters. In the former case, we can equivalently deal with the coefficients in  $\mathbf{A}$  or with the covariance matrix  $\Sigma_A$ , since they are in a one-to-one correspondence. We formulate the prior distribution in terms of the coefficients in  $\mathbf{A}$ : we let its elements  $a_{lk}$ , for  $l = 1, \dots, L$  and  $k = 1, \dots, K$ , to be independently distributed as follow

$$\begin{aligned} a_{lk} &\stackrel{\text{iid}}{\sim} N(0, \gamma^2), & l = 1, \dots, L, & k = 1, \dots, K, & k < l, \\ a_{kk}^2 &\stackrel{\text{ind}}{\sim} \gamma^2 \chi_{K-k+1}^2, & k = 1, \dots, K, \\ a_{lk} &= 0, & \text{otherwise,} \end{aligned} \tag{10}$$

for some variance hyperparameter  $\gamma^2 > 0$ . By employing specification (10), we automatically deal with the identifiability constraints of Sect. 3.3.

## 4 Posterior Inference

Posterior inference cannot be conducted in closed form. We need to turn to simulation-based fitting techniques to obtain samples from the posterior distribution. Generally, we would prefer to work with a marginal specification in order to reduce the dimensionality of the problem as much as possible before doing any computation. The *normal-normal* conjugacy enables marginalization of Eq. (6) over  $\mathcal{Z}$ , leading to the following Gaussian model, no longer having a factorized specification

$$(\text{vec}(\mathbf{B}) \mid \mathbf{A}, \sigma^2) \sim N(0, \mathbf{C}), \tag{11}$$

where  $n = L \times T$ , and  $\mathbf{C} = \boldsymbol{\Sigma}_T \otimes \boldsymbol{\Sigma}_A + \sigma^2 I_{n \times n}$ . The covariance matrix in Eq. (11) is diagonally dominant and thus invertible, allowing to ignore singularity issues that would arise when considering  $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_T \otimes \boldsymbol{\Sigma}_A$  alone. In Appendix A we describe a simple Metropolis-Hastings (M-H) model fitting algorithm with multivariate Gaussian random walk which is based on the marginal specification (11) and is sufficient to guarantee a satisfying mixing. For this purpose, it is convenient to parametrize the residual variance  $\sigma^2$  on the logarithmic scale, i.e.,  $\tau = \log \sigma^2$ . Computational details concerning the M-H sampler are provided in Sect. 4.1, where we describe how to exploit the separability assumption for fast computations.

Suppose we are able to draw posterior samples for  $\mathbf{A}$  and  $\tau$ , for instance by using the M-H in Algorithm 1 (in Appendix). Then, predicting the new BOLD values at a new time and brain region, conditionally on the data, is relatively simple and can be obtained by means of the so-called kriging equations (e.g., [18]), Chap. 2. We remark that kriging the BOLD values is not of direct interest in the analysis of RS-fMRI data. However, as we will discuss in Sect. 5, this procedure is useful to conduct model assessment in terms of out-of-sample prediction performance. Let  $\mathbf{B}_0$  be the  $L_0 \times T_0$  matrix of unobserved BOLD values over a new grid of time values with length  $T_0$  for some subset of  $L_0$  brain regions from the original  $L$  regions. We are interested in finding the predictive distribution

$$p(\mathbf{B}_0 \mid \mathbf{B}) = \int p(\mathbf{B}_0 \mid \mathbf{B}, \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \mathbf{B}) d\boldsymbol{\theta}, \tag{12}$$

where we have defined  $\boldsymbol{\theta} = (\text{vec}(\mathbf{A}), \tau)$ . The conditional distribution  $p(\text{vec}(\mathcal{B}_0) | \mathcal{B}, \boldsymbol{\theta})$  is available in closed form, being a multivariate Gaussian distribution

$$\begin{aligned} (\text{vec}(\mathcal{B}_0) | \mathcal{B}, \boldsymbol{\theta}) &\sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0), \\ \boldsymbol{\mu}_0 &= \tilde{\mathbf{C}}_0^\top \mathbf{C}^{-1} \text{vec}(\mathcal{B}), \\ \boldsymbol{\Sigma}_0 &= \mathbf{C}_0 - \tilde{\mathbf{C}}_0^\top \mathbf{C}^{-1} \tilde{\mathbf{C}}_0, \end{aligned} \quad (13)$$

where  $\mathbf{C}$  is the covariance matrix of  $\text{vec}(\mathcal{B})$  given the parameters,  $\mathbf{C}_0$  is the covariance matrix of  $\text{vec}(\mathcal{B}_0)$ , and finally  $\tilde{\mathbf{C}}_0$  represent the cross-covariance matrix between  $\text{vec}(\mathcal{B})$  and  $\text{vec}(\mathcal{B}_0)$ . Thus, draws from the predictive distribution in (12) can be obtained by composition sampling, by first drawing posterior values for  $\mathbf{A}$  and  $\tau$  and then by sampling from the multivariate Gaussian distribution in (13).

## 4.1 Computational Difficulties

Some useful matrix identities can be exploited to reduce the computational burden both for the M-H algorithm and for Eq. (13). We start by inspecting the log-posterior distribution of the marginal model (11) which is equal, up to an additive constant, to the following quantity

$$\mathcal{L}(\mathbf{A}, \tau; \mathcal{B}) = -\frac{1}{2} \log |\mathbf{C}| - \frac{1}{2} \text{vec}(\mathcal{B})^\top \mathbf{C}^{-1} \text{vec}(\mathcal{B}) + \log p(\mathbf{A}) + \log p(\tau), \quad (14)$$

where  $p(\mathbf{A})$  and  $p(\tau)$  denote the probability density functions of the priors for  $\mathbf{A}$  and  $\tau = \log \sigma^2$ , respectively. During the MCMC chain the log-posterior is evaluated several times and therefore it is crucial to maintain computations as fast as possible. A potential computational bottleneck is represented by the inverse of the matrix  $\mathbf{C} = \boldsymbol{\Sigma}_T \otimes \boldsymbol{\Sigma}_A + \sigma^2 I_{n \times n}$ , which in our case is a  $n \times n$  matrix, with  $n = T \times L$ . In the separable case, thanks to the properties of the Kronecker product, this issue can be attenuated by exploiting the following decomposition of the inverse of  $\mathbf{C}$ , being equal to

$$\mathbf{C}^{-1} = (\mathbf{U}_T \otimes \mathbf{U}_A)(\boldsymbol{\Lambda}_T \otimes \boldsymbol{\Lambda}_A + \sigma^2 I_{n \times n})^{-1}(\mathbf{U}_T \otimes \mathbf{U}_A)^\top, \quad (15)$$

where  $\boldsymbol{\Sigma}_T = \mathbf{U}_T \boldsymbol{\Lambda}_T \mathbf{U}_T^\top$  and  $\boldsymbol{\Sigma}_A = \mathbf{U}_A \boldsymbol{\Lambda}_A \mathbf{U}_A^\top$  are the spectral decompositions of the matrices. Detailed calculations leading to (15) are given in Appendix A. The above spectral decompositions are relatively cheap in our context. Notice also that the decomposition of  $\boldsymbol{\Sigma}_T$  has to be computed only once, since it does not depend on unknown parameters in our formulation. More importantly, the matrix  $(\boldsymbol{\Lambda}_T \otimes \boldsymbol{\Lambda}_A + \sigma^2 I_{n \times n})^{-1}$  is diagonal, and can be, therefore, inverted directly.

Decomposition (15) allows easy evaluation of the log-determinant of  $\mathbf{C}$ , which is given as a simple function of the previously obtained eigenmatrices  $\boldsymbol{\Lambda}_T$  and  $\boldsymbol{\Lambda}_A$ ,

$$\log |\mathbf{C}| = \log |\boldsymbol{\Lambda}_T \otimes \boldsymbol{\Lambda}_A + \sigma^2 I_{n \times n}| = \sum_{i=1}^n \log (\lambda_i + \sigma^2),$$

where  $\lambda_i$  is the  $i$ -th entry of the diagonal matrix  $\boldsymbol{\Lambda}_T \otimes \boldsymbol{\Lambda}_A$ , for  $i = 1, \dots, n$ . Notice that, as long as  $K \ll L$ , the covariance matrix  $\boldsymbol{\Sigma}_A$  is not full rank, meaning that some of the eigenvalues  $\lambda_i$  are exactly equal to zero.

Leveraging the decomposition (15) and using some simple properties of the Kronecker product, we can express the quadratic form  $\text{vec}(\mathbf{B})^\top \mathbf{C}^{-1} \text{vec}(\mathbf{B})$  in (14) as follows:

$$\text{vec}(\mathbf{B})^\top \mathbf{C}^{-1} \text{vec}(\mathbf{B}) = \text{vec} (\mathbf{U}_A^\top \mathbf{B} \mathbf{U}_T)^\top (\boldsymbol{\Lambda}_T \otimes \boldsymbol{\Lambda}_A + \sigma^2 I_{n \times n})^{-1} \text{vec} (\mathbf{U}_A^\top \mathbf{B} \mathbf{U}_T).$$

This drastically reduces the computational burden, since it avoids storage in memory of very large  $n \times n$  matrices. With similar reasoning, the kriging Eq. (13) can also be obtained quite cheaply, adopting fast algorithms for products between matrix involving Kronecker products, implemented for instance in the `klin` R package [33]. The code used in the paper is made available at the link <https://github.com/tomasorigon/StartUpResearch>.

## 5 Data Analysis

### 5.1 Model Checking

We now apply the spatio-temporal model presented in Sect. 3 to the RS- fMRI dataset. However, before proceeding with the interpretation of the results, it is crucial to check the adequacy of the fit to the data to assess the plausibility of the proposed model (Chap. 6, [34]). We measure the goodness of fit of our model by means of out-of-sample predictions, DIC indices, and by direct graphical inspection.

In performing Bayesian inference, we employ the priors described in Sect. 3.4, which require the specification of some tuning parameters. The hyperparameter  $\gamma^2$  controls the prior variability of the coefficients in  $\mathbf{A}$ . By choosing  $\gamma^2 = 100$  we incorporate vague prior information into the model. Following a similar rationale, we let the parameters of the residual variance  $\sigma^2$  to be equal to  $a_\sigma = b_\sigma = 1$ , which induces a fairly noninformative prior for the residual variance.

As discussed in Sect. 3, the temporal component is controlled by the Gaussian processes in  $\mathbf{V}(t)$ , which in turn are characterized by their correlation function  $\rho(t, t')$ . Depending on the choice of such a function, the latent processes  $\mathbf{V}(t)$  could behave quite differently. An extreme example consists in setting  $\rho(t, t') = \mathbb{1}(t = t')$ , with  $\mathbb{1}(\cdot)$  denoting the indicator function, which would imply that the processes  $\mathbf{V}(t)$  are independent over time, and the model in Eqs. (6) and (7) reduces to a simple Bayesian factor model. Instead, by letting  $\rho(t, t') = \exp\{-\psi|t - t'|\}$ , with  $\psi = 3 \times 10^{-2}$ , we introduce temporal dependence favoring stationary and quite regular paths for

**Table 1** For different values of  $K = 1, 3, 5$ , and for subjects  $i = 1, 2$ , the DIC index, the total number of parameters and the out-of-sample root mean squared error (RMSE) are reported. For each individual, it is also shown the out-of-sample RMSE of a random forest model. For both the subjects, the bold values represent the best model according to each index; in both cases, the lower the better

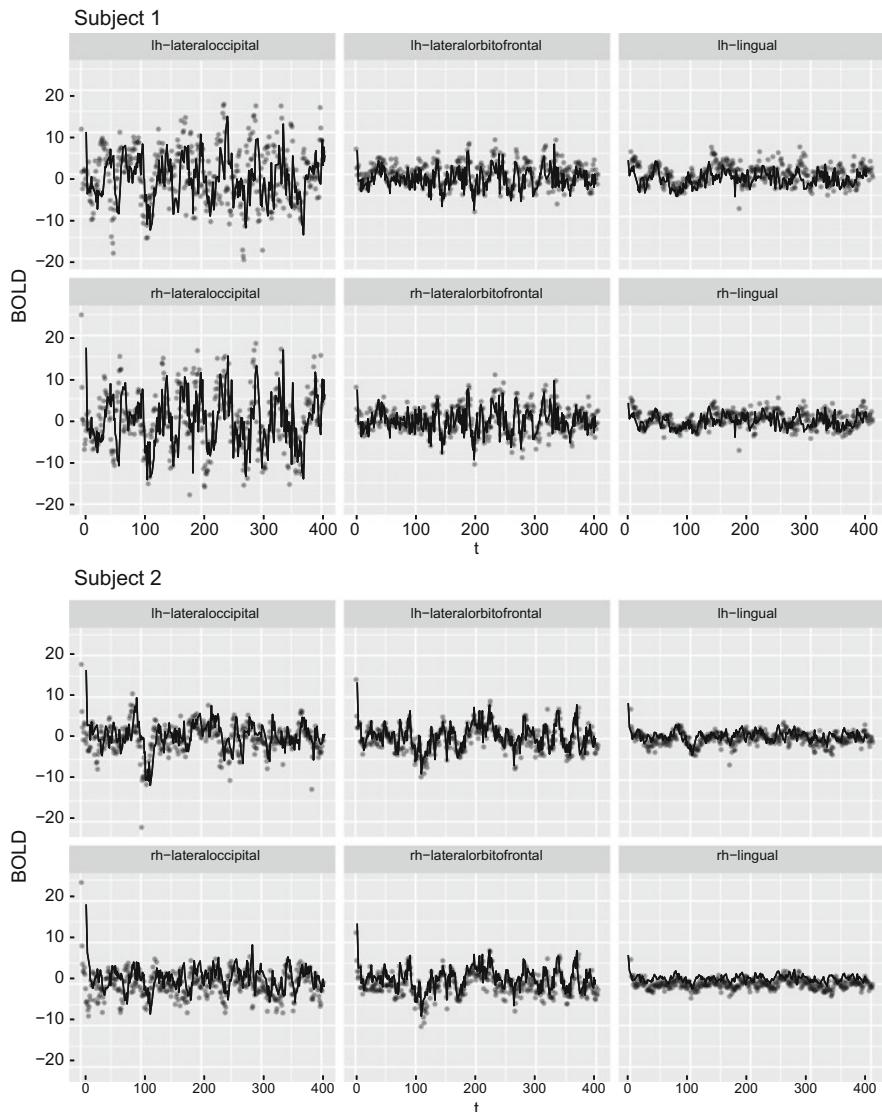
	$K$	DIC	Total # of parameters	RMSE
Subject 1	1	65988.81	69	3.63
	3	60204.04	202	3.34
	5	<b>54800.08</b>	331	<b>3.17</b>
	Random forest	—	—	3.21
Subject 2	1	48108.30	69	2.26
	3	35941.24	202	2.06
	5	<b>31555.90</b>	331	<b>2.00</b>
	Random forest	—	—	2.24

the latent processes  $\mathbf{V}(t)$ . In fact, such a correlation function implicitly induces a continuous-time first order autoregressive process for each element of  $\mathbf{V}(t)$ , with autocorrelation coefficient equal to  $\exp\{-3 \times 10^{-2}\} \approx 0.97$ , which favors fairly regular latent trajectories.

Finally, the number of latent processes  $K$  has to be carefully selected, since its choice critically impacts the computational performance. Indeed, the number of parameters grows linearly as a function of  $K$  and therefore overly complex models become harder to fit using the M-H algorithm. More sophisticated and efficient approaches might mitigate this issue, and a brief discussion is given in Sect. 6. We set  $K = 5$  mainly because of these practical considerations, but we provide below some empirical evidence which offers some reassurance that a model based on such a choice is sufficiently flexible to capture the brain connectivity structure of our data.

To assess whether our model leads to reasonable inferential conclusions and to discriminate between competing models, we conduct some posterior checks, obtaining measures of out-of-sample accuracy as well as the DIC indices [35]. The original dataset is split in two parts: the first one is used for estimation and it comprises the 75% of randomly selected columns of  $\mathcal{B}$ , i.e., different time instants, selected at random, while the remaining 25% is used as a test set to compute for instance the out-of-sample root mean squared error (RMSE).

In Table 1 we compare our model with alternatives involving a smaller number of latent processes, e.g., with  $K = 1$ , or  $K = 3$ , showing that for both the subjects we obtain improved accuracy and lower DIC indices with  $K = 5$ . The out-of-sample predictions are obtained by means of the kriging Eq. (13), after plugging in the MAP estimate. Formally, this is incorrect and may lead to the underestimation of the predictive uncertainty. More correctly, we should take the average of the kriged estimates over posterior realizations of the parameters. However, the latter procedure turns out to be computationally too expensive, so we adopt the aforementioned plug-in alternative. Posterior samples for all the competing models and both the subjects are



**Fig. 4** Scatter plots of the BOLD values for subjects  $i = 1, 2$ , over the time grid  $t = 1, \dots, 403$ , for 6 selected brain regions. Three of these regions are located in the left hemisphere (lh-lateraloccipital, lh-lateralorbitofrontal, lh-lingual), while the others are their symmetric correspondent of the right hemisphere (rh-lateraloccipital, rh-lateralorbitofrontal, rh-lingual). The solid lines represent the predicted values obtained by means of the kriging Eq. (13), after plugging-in the MAP estimate

obtained using the M-H algorithm, which rely on a Gaussian random walk proposal, each with its own covariance matrix. These matrices, one for each competing model in Table 1, have been carefully tuned essentially by trial and error, to ensure good mixing and quick convergence of each MCMC chain. For each MCMC chain we retain 250,000 thinned samples from a chain of 5,000,000 iterations, after a burn-in period of about 100,000 draws. The trace plots show no evidence against convergence and a decent mixing.

As shown in Table 1, we further compare our model with a benchmark method for regression, random forests [36], in which the BOLD response values are fitted as a function of time and brain regions. Although the latter method is specifically designed to provide accurate predictions of response values, our proposal seems to have better out-of-sample performance.

Finally, in Fig. 4 we graphically explore the predictive performance of our model by comparing the original BOLD values with their predictions. For illustrative reasons we displayed only few brain regions, but we remark that the other cases present similar patterns. The graphs of Fig. 4 further corroborate the reasonableness of our proposal, which is able to capture the main trends and the differences in variability of the BOLD values among brain regions.

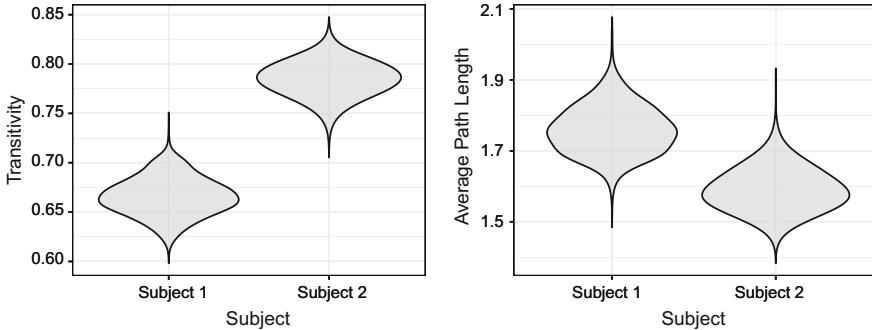
We remark that, in order to reduce the computational burden, the DIC indices of Table 1 and the results in Sect. 5.2 are also based on this 75% partition of the observations, which we believe well-represents the whole dataset.

## 5.2 Network Analysis

In neurological applications it is common practice to explore functional connectivity networks exploiting graph theoretical approaches. As summarized in [21], the typical pipeline of the analysis of structural and functional brain networks consists of the following steps: the identification of the brain regions of interests, the estimation of a continuous measure of association between regions, the application of a threshold to generate a binary adjacency matrix, and the computation of network indices on the obtained undirected graph. In our case, the regions of interests are those obtained from the Desikan parcellation [16], whereas a continuous measure of association can be obtained from the covariance matrix  $\Sigma_A$ , appropriately standardized. Following [22], we define a  $L \times L$  binary adjacency matrix  $\mathbf{G}$  as the truncation of a correlation matrix, that is

$$[\mathbf{G}]_{ll'} = \mathbb{1} ([\text{Cor}(\Sigma_A)]_{ll'} > \text{threshold}), \quad \text{for } l \neq l', \quad (16)$$

and  $[\mathbf{G}]_{ll} = 0$  for  $l = 1, \dots, L$  and  $l' = 1, \dots, L$ , where threshold is a constant between 0 and 1, and  $\text{Cor}(\Sigma_A)$  denotes the correlation matrix obtained by standardizing  $\Sigma_A$ . The covariance matrix has a direct interpretation, but the precision matrix, i.e., its inverse, might be considered as well. The choice of the threshold is crucial in determining  $\mathbf{G}$  but, unfortunately, there are no general guidelines. Indeed,



**Fig. 5** Posterior distributions (violin plots) of the transitivity indices and the average path lengths, for subjects  $i = 1, 2$ , evaluated on a graph  $G$  with threshold = 0.8

for any value of the threshold, we could obtain a graph having different sparsity and network properties. To mitigate this issue, we explored a range of plausible thresholds [21] and we noticed that in our setting the inferential conclusions are insensitive to moderate variations of the threshold.

Given the threshold, the adjacency matrix  $G$  is a random quantity whose posterior distribution can be easily approximated using the output of the MCMC. In particular, it is possible to quantify the uncertainty of any network characteristic one could be possibly interested in. Among several alternatives, a relevant network index is the so called clustering coefficient, also known as transitivity in the statistical literature, or fraction of transitive triples, which is a measure of global cohesion of the graph  $G$ . Another index which provides a measure of global connectivity of a given graph is the average path length, defined as the average minimal distance between two brain regions.

We expect these indicators to be negatively correlated in our application: broadly speaking, a high number of transitive triples suggests that two brain regions require a low number of steps to be connected. We refer to [37] for the formal definition of these indices and their theoretical properties. In Fig. 5 we reported the posterior distributions of these measures for both of the subjects. We see substantial differences. In particular, subject 2 presents a much higher functional activity compared to subject 1, in terms of both indices. Understanding the qualitative reasons for such a marked distinction between the two subjects is beyond the aim of this paper. Nonetheless, we remark that our proposal was able to capture the differential traits of the two brains, thus providing a tool for detecting differences in functional connectivity and for quantifying the related uncertainty.

## 6 Discussion

In this paper we proposed a spatio-temporal Bayesian factor model for the analysis of RS- fMRI data. Both for interpretational and computational reasons, we employed a separable structure. We discussed how to obtain posterior inference using a M-H algorithm, providing also some technical details that could speed up computations. Finally, we applied our model to a real RS- fMRI dataset and we provided an example.

Although the model we described is designed for a single-subject analysis, it could be extended to the multi-subject case adding a further layer in the hierarchical specification of Sect. 3. One possibility is to borrow information across individuals assuming exchangeable prior distributions for the subject-specific covariance matrices  $\Sigma_A^{(i)}$ . In particular, if we let  $(\Sigma_A^{(i)} | V) \sim \text{InverseWishart}(K, V)$  independently and identically distributed, we could then induce dependence across subjects by placing a hyperprior distribution for  $V$ , which in turn could be interpreted as the baseline covariance structure, common to all the individuals. Additionally, in the multi-subject setting it might be possible to explore the effect of individual covariates on functional connectivity, which we did not attempt, having considered only two subjects.

Another possible extension, already mentioned in Sect. 3.2, could be the implementation of a dynamic model. This would take us to a non separable model by specifying a factor loading matrix  $A(t)$  that also evolves in time. This issue is examined in depth, in a different applied context, by [25, 26]. To capture the evolution of  $A(t)$ , avoiding, at the same time, naive approaches with poor performances, they use independent Gaussian processes with unit variance as a set of basis functions. Thus, the factor loading matrix  $A(t)$  would itself be a time-varying random function, implying that  $\Sigma_A(t) = A(t)A^\top(t)$ , for any fixed  $t$ . As a consequence, the evaluation of the adjacency matrix in Eq. (16) for each correlation matrix would generate a dynamic network.

Generalizing our model beyond separability can be done in several other ways. For instance, one could assume that the latent Gaussian processes in  $V(t)$  are independent but not identically distributed, and characterized by different correlation functions  $\rho_k(t, t')$ . This would imply a more sophisticated and non stationary covariance structure for the mean process  $Z(t)$ . Both the above settings are arguably more realistic [31], but unfortunately they do not lead to the simple interpretation which follows from our separable model.

Besides the difficulties in interpretation that could arise from the above generalizations, the main challenge is on the computational side. The algorithm for posterior inference we described in Sect. 4 can be improved in several different directions. For instance, the default prior setting and the parameter expansion strategy of [38] could be adapted to our framework to provide better mixing. In multi-subject scenarios, or whenever the number of brain regions is massive, and therefore MCMC computations are prohibitive, one could attempt deterministic model fitting approximations like variational Bayes. In the context of fMRI data this approach was developed by [9], and it could possibly be adapted to our model.

**Acknowledgements** We are grateful to Greg Kiar and Eric Bridgeford from NeuroData at Johns Hopkins University, who graciously pre-processed the raw DTI and R-fMRI imaging data available at [http://fcon\\_1000.projects.nitrc.org/indi/CoRR/html/nki\\_1.html](http://fcon_1000.projects.nitrc.org/indi/CoRR/html/nki_1.html), using the pipelines NDMG and C-PAC. The authors are also thankful to the organizers of StartUp Research for coordinating such a stimulating event.

## 7 Computational Details

In this appendix we describe a simple Metropolis-Hastings for posterior inference. The algorithm is summarized in Algorithm 1. Additionally, we derive the identity in Eq. (15). Because of orthogonality, we have that  $\mathbf{U}_T \mathbf{U}_T^\top = I_{T \times T}$  and  $\mathbf{U}_A \mathbf{U}_A^\top = I_{L \times L}$  and recall also that the matrices  $\boldsymbol{\Lambda}_T$  and  $\boldsymbol{\Lambda}_A$  are diagonal, containing the eigenvalues of  $\boldsymbol{\Sigma}_T$  and  $\boldsymbol{\Sigma}_A$ , respectively. Exploiting the spectral decompositions of  $\boldsymbol{\Sigma}_T$  and  $\boldsymbol{\Sigma}_A$  and the basic properties of the Kronecker product, we get

$$\begin{aligned} \mathbf{C} &= (\mathbf{U}_T \boldsymbol{\Lambda}_T \mathbf{U}_T^\top) \otimes (\mathbf{U}_A \boldsymbol{\Lambda}_A \mathbf{U}_A^\top) + \sigma^2 I_{n \times n} \\ &= (\mathbf{U}_T \otimes \mathbf{U}_A)(\boldsymbol{\Lambda}_T \otimes \boldsymbol{\Lambda}_A)(\mathbf{U}_T \otimes \mathbf{U}_A)^\top + \sigma^2 I_{n \times n}. \end{aligned}$$

Then, we can write the identity matrix  $I_{n \times n} = (\mathbf{U}_T \mathbf{U}_T^\top) \otimes (\mathbf{U}_A \mathbf{U}_A^\top) = (\mathbf{U}_T \otimes \mathbf{U}_A)(\mathbf{U}_T \otimes \mathbf{U}_A)^\top$ . Rearranging the above equation, we get

---

### Algorithm 1: Metropolis-Hastings algorithm for posterior inference

---

```

begin
    Let the matrix  $\boldsymbol{\Sigma}_{\text{Metropolis}}$  be a tuning parameter and let the superscript  $(r)$  denote the
    value of the corresponding parameter at the  $r$ -th step of the MCMC chain.
    for  $r$  from 1 to  $R$  do
        Step [1]. Sample a proposed value  $\theta^* = (\text{vec}(A^*), \tau^*)$  from a multivariate Gaussian
        
$$\theta^* \sim N(\theta^{(r)}, \boldsymbol{\Sigma}_{\text{Metropolis}}),$$

        having care that the upper triangular values of  $A^*$  should be equal to zero. This can
        be accomplished either by ignoring them in the MCMC step, or forcing the
        corresponding elements of the  $\boldsymbol{\Sigma}_{\text{Metropolis}}$  matrix to have zero variance. Step [2]. Set
        the acceptance probability  $\alpha$  equal to
        
$$\alpha = \min \left\{ 1, \exp \left\{ \mathcal{L}(A^*, \tau^*; \mathcal{B}) - \mathcal{L}(A^{(r)}, \tau^{(r)}; \mathcal{B}) \right\} \right\},$$

        where  $\mathcal{L}(A, \tau; \mathcal{B})$  denotes the log-posterior distribution (14), up to an additive
        constant.
        Step [3]. With probability  $\alpha$ , accept the proposed value  $\theta^*$  and set
        
$$A^{(r+1)} \leftarrow A^* \quad \text{and} \quad \tau^{(r+1)} \leftarrow \tau^*.$$


```

---

$$\mathbf{C} = (\mathbf{U}_T \otimes \mathbf{U}_A)(\Lambda_T \otimes \Lambda_A + \sigma^2 I_{n \times n})(\mathbf{U}_T \otimes \mathbf{U}_A)^\top,$$

from which decomposition of  $\mathbf{C}^{-1}$  in Eq. (15) follows directly.

## References

1. Biswal, B., Yetkin, F.Z., V.M., H., Hyde, J.: Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* **34**(4), 537–541 (1995). <https://doi.org/10.1002/mrm.1910340409>
2. Lee, M.H., Smyser, C.D., Shimony, J.S.: Resting state fMRI: a review of methods and clinical applications. *Am. J. Neuroradiol.* **34**(10), 1866–1872 (2013). <https://doi.org/10.3174/ajnr.A3263>
3. Poldrack, R.A., Mumford, J.A., Nichols, T.E.: *Handbook of Functional MRI Data Analysis*. Cambridge University Press, Cambridge (2011)
4. Smitha, K., Akhil Raja, K., Arun, K., Rajesh, P., Thomas, B., Kapilamoorthy, T., Kesavadas, C.: Resting state fMRI: A review on methods in resting state connectivity analysis and resting state networks. *Neuroradiol. J.* **30**(4), 305–317 (2017). <https://doi.org/10.1177/1971400917697342>
5. Bowman, F.D., Caffo, B., Bassett, S.S., Kilts, C.: A Bayesian hierarchical framework for spatial modeling of fMRI data. *NeuroImage* **39** (2008). <https://doi.org/10.1016/j.neuroimage.2007.08.012>
6. Hartvig, N.G.: A stochastic geometry model for functional magnetic resonance images. *Scand. J. Stat.* **29**(3), 333–353 (2002). <https://doi.org/10.1111/1467-9469.00294>
7. Quirós, A., Diez, R.M., Gamerman, D.: Bayesian spatiotemporal model of fMRI data. *NeuroImage* **49**(1), 442–456 (2010). <https://doi.org/10.1016/j.neuroimage.2009.07.047>
8. Stephan, K.E., Kasper, L., Harrison, L.M., Daunizeau, J., den Ouden, H.E., Breakspear, M., Friston, K.J.: Nonlinear dynamic causal models for fMRI. *NeuroImage* **42**(2), 649–662 (2008). <https://doi.org/10.1016/j.neuroimage.2008.04.262>
9. Zhang, L., Guindani, M., Versace, F., Engelmann, J.M., Vannucci, M.: A spatiotemporal nonparametric Bayesian model of multi-subject fMRI data. *Ann. Appl. Stat.* **10**(2), 638–666 (2016). <https://doi.org/10.1214/16-AOAS926>
10. Zhang, L., Guindani, M., Versace, F., Vannucci, M.: A spatio-temporal nonparametric Bayesian variable selection model of fMRI data for clustering correlated time courses. *NeuroImage* **95**, 162–175 (2014). <https://doi.org/10.1016/j.neuroimage.2014.03.024>
11. Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M.: Bayesian analysis of neuroimaging data in FSL. *NeuroImage* **45** (2009). <https://doi.org/10.1016/j.neuroimage.2008.10.055>
12. Zhang, L., Guindani, M., Vannucci, M.: Bayesian models for functional magnetic resonance imaging data analysis. *Comput. Stat.* **7**(1), 21–41 (2015). <https://doi.org/10.1002/wics.1339>. Wiley Interdisciplinary Reviews
13. Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.P., Frith, C.D., Frackowiak, R.S.J.: Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* **2**(4), 189–210 (1994). <https://doi.org/10.1002/hbm.460020402>
14. Craddock, R.C., Jbabdi, S., Yan, C., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C., Heberlein, K., Colcombe, S., Milham, M.P.: Imaging human connectomes at the macroscale. *Nat. Methods* **6**, 524–539 (2013). <https://doi.org/10.1038/nmeth.2482>
15. Erhardt, E.B., Allen, E.A., Wei, Y., Eichele, T., Calhoun, V.D.: Simtb, a simulation toolbox for fmri data under a model of spatiotemporal separability. *NeuroImage* **59**(4), 4160–4167 (2012). <https://doi.org/10.1016/j.neuroimage.2011.11.088>

16. Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006). <https://doi.org/10.1016/j.neuroimage.2006.01.021>
17. Cressie, N., Wikle, C.K.: Statistics for Spatio-Temporal Data. Wiley, New York (2011)
18. Banerjee, S., Carlin, B.P., Gelfand, A.E.: Hierarchical Modeling and Analysis for Spatial Data, 2nd edn. CRC Press, Boca Raton (2014)
19. Quick, H., Banerjee, S., Carlin, B.P.: Modeling temporal gradients in regionally aggregated California asthma hospitalization data. *Ann. Appl. Stat.* **7**(1), 154–176 (2013). <https://doi.org/10.1214/12-AOAS600>
20. Ramsay, J., Silverman, B.W.: Functional Data Analysis. Springer, New York (2005)
21. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009). <https://doi.org/10.1038/nrn2575>
22. van den Heuvel, M., Stam, C.J., Boersma, M., Hulshoff Pol, H.E.: Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *NeuroImage* **43**, 528–539 (2008). <https://doi.org/10.1016/j.neuroimage.2008.08.010>
23. Alexander-Bloch, A.F., Vértes, P., Stidd, R., Lalonde, F., Clasen, L., Rapoport, J., Giedd, J., Bullmore, E.T., Gogtay, N.: The anatomical distance of functional connections predicts brain network topology in health and schizophrenia. *Cereb. Cortex* **23**(1), 127–138 (2013). <https://doi.org/10.1093/cercor/bhr388>
24. Salvador, R., Suckling, J., Coleman, M., Pickard, J., Menon, D., Bullmore, E.: Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb. Cortex* **15**, 1332–1342 (2005). <https://doi.org/10.1093/cercor/bhi016>
25. Durante, D., Scarpa, B., Dunson, D.B.: Locally adaptive factor processes for multivariate time series. *J. Mach. Learn. Res.* **15**, 1493–1522 (2014). <http://jmlr.org/papers/v15/durante14a.html>
26. Fox, E., Dunson, D.: Bayesian nonparametric covariance regression. *J. Mach. Learn. Res.* **16**, 2501–2542 (2015). <http://jmlr.org/papers/v16/fox15a.html>
27. Prado, R., West, M.: Time Series. Modeling, Computation and Inference. CRC Press (2010)
28. Carvalho, C.M., Chang, J., Lucas, J.E., Nevins, J.R., Wang, Q., West, M.: High-dimensional sparse factor modeling: applications in gene expression genomics. *J. Am. Stat. Assoc.* **103**(484), 1438–1456 (2008). <https://doi.org/10.1198/016214508000000869>
29. West, M.: Bayesian factor regression models in the large p, small n paradigm. In: J. Bernardo, M.J. Bayarri, J.O. Berger, A.P. Dawid, D. Heckerman, A.F.M. Smith, M. West (eds.) Bayesian Statistics, vol. 7, pp. 733–742. Oxford University Press (2003)
30. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. The MIT Press (2006)
31. Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., Calhoun, V.D.: Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* **24**, 663–676 (2014). <https://doi.org/10.1093/cercor/bhs352>
32. Dawid, A.P.: Some matrix-variate distribution theory: notational considerations and a Bayesian application. *Biometrika* **68**(1), 265–274 (1981). <https://doi.org/10.2307/2335827>
33. Papp, T.K.: klin: Linear Equations with Kronecker Structure R package version 2007-02-05 (2012). <https://CRAN.R-project.org/package=klin>.
34. Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B.: Bayesian Data Analysis, 3rd edn. CRC Press (2014)
35. Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A.: Bayesian measures of model complexity and fit. *J. R. Stat. Society. Ser. B: Stat. Methodol.* **64**(4), 583–616 (2002). [10.1111/1467-9868.00353](https://doi.org/10.1111/1467-9868.00353)
36. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001). <https://doi.org/10.1023/A:1010933404324>
37. Kolaczyk, E.D.: Statistical Analysis of Network Data. Springer, New York (2009). <https://doi.org/10.1007/978-0-387-88146-1>
38. Ghosh, J., Dunson, D.B.: Default prior distributions and efficient posterior computation in Bayesian factor analysis. *J. Comput. Graph. Stat.* **18**(2), 306–320 (2009). <https://doi.org/10.1198/jcgs.2009.07145>

# Challenges in the Analysis of Neuroscience Data



Michele Guindani and Marina Vannucci

**Abstract** In the last two decades, our understanding of the mechanisms underlying the functioning and disruption of the human brain has advanced considerably. The previous chapters of the book have provided a compelling argument for demonstrating the advantages of thoughtful, non-naive, statistical approaches for analyzing brain imaging data. Here, we provide a review of the main themes highlighted in those chapters, and we further discuss some of the challenges that statistical imaging is currently confronted with. In particular, we emphasize the importance of developing analytical frameworks that allow to characterize the heterogeneity typically observed in brain imaging both within- and between- subjects, by capturing the main sources of variability in the data. More specifically, we focus on clustering methods that identify groups of subjects characterized by similar patterns of brain responses to a task; on dynamic temporal models that characterize the heterogeneity in individual functional connectivity networks; and on multimodal imaging analysis and imaging genetics that combine information from multiple data sources in order to achieve a better understanding of brain processes.

**Keywords** Brain imaging data · fMRI data · Clustering  
Dynamic functional connectivity · Multimodal analysis · Imaging genetics

## 1 Introduction

In the last two decades, our understanding of the mechanisms underlying the functioning and disruption of the human brain has advanced considerably. The development of a number of innovative technologies has spurred unparalleled enthusiasm

---

M. Guindani (✉)

Department of Statistics, University of California, Irvine, USA  
e-mail: [michele.guindani@uci.edu](mailto:michele.guindani@uci.edu)

M. Vannucci

Department of Statistics, Rice University, Houston, USA  
e-mail: [marina@rice.edu](mailto:marina@rice.edu)

for research in the Neurosciences. Major breakthroughs have escaped the bounds of academic labs, and have been often widely publicized by the media. For example, in February 2014, a special issue of the *National Geographic* magazine enthusiastically hailed the new technologies that are “shedding light on biology’s greatest unsolved mystery: how the brain really works”. In the future, these technologies are expected to have a profound impact on the type of clinical treatments administered by physicians. On September 21st 2013, the British journal *The Guardian* dedicated an article to the new landscape of psychiatry, where the use of widely employed anti-depressant drugs has been called into question in favor of alternative treatments directly targeting the functioning of specific neural circuits. By studying how brain areas interact differently in healthy and depressed patients, the hope is to decode the determinants of complex human emotion and behavior.

Statistical methods play a crucial role in the quest for a better understanding of brain mechanisms, and their disruption in the face of disease. As an illustration, in the analysis of many types of brain imaging data, it is customary to employ *statistical parametric maps*, e.g., localized maps of  $p$ -values or posterior probabilities, to inform on the significance and spatiotemporal organization of the observed signal across distinct brain regions [1]. Those images provide a synthetic representation of significant areas of the brain, which may be targeted for further research and, also, to improve clinical diagnosis or intervention. However, early approaches based on naive  $t$ -tests or ANOVA statistics have shown limitations, especially due to their inability to take into account the complexity and specific characteristics of the data. Thus, the need for fairly sophisticated statistical techniques has emerged, e.g. to address the typically weak signal, high dimensionality and complex spatio-temporal correlation structure of the data.

The previous chapters of the book have provided a compelling argument for demonstrating the advantages of thoughtful, non-naive, statistical approaches for analyzing brain imaging data. Here, we provide a review of the main themes highlighted in those chapters, and we further discuss some of the challenges that statistical imaging is currently confronted with. More specifically, in Sect. 2 we provide a summary review of the main inferential objectives associated with structural and functional brain imaging modalities, and discuss general modeling strategies that have been developed to achieve such inferences. In Sect. 3, we discuss the importance of developing analytical frameworks that allow to characterize the heterogeneity typically observed in brain imaging both within- and between- subjects. In Sect. 4, we examine clustering approaches, that allow to identify groups of subjects characterized by similar patterns of brain responses to a task. In Sect. 5, we discuss dynamic temporal models to capture the heterogeneity of functional connectivity network states experienced by subjects in the course of an experiment. In Sect. 6, we present recent modeling trends, which aim at combining information from multiple data sources in order to achieve a better understanding of brain processes: multimodal imaging analysis and imaging genetics are examples of those developments. In Sect. 7, we provide some concluding remarks.

## 2 Statistical Analysis of Brain Imaging Data

We start our discussion by noting that the statistical methods employed in the analysis of brain imaging data necessarily depend on the specific type of technology employed and need to be necessarily informed by the expert knowledge of neuroscientists. Brain imaging technologies can be roughly separated into three categories: *structural*, *functional* and *molecular imaging* technologies. Each technology aims at capturing different characteristics of brain mechanisms, and therefore requires specifically tailored methods.

### 2.1 Structural Imaging

Structural brain imaging aims at providing a description of the anatomical structure of the brain. As an illustration, computed axial tomography (CT) uses X-rays to quickly identify different levels of density and tissues inside a solid organ, and can be used to obtain clinical evidence of trauma, e.g., a stroke. MRI scans use powerful magnetic fields and radio frequency pulses to create high-resolution images, and thus they are able to depict the brain anatomy in greater detail. Signal change and cerebral atrophy visible on structural MRI can be used to identify diagnostically relevant imaging features to help the clinical diagnosis of neurodegenerative dementias.

Diffusion tensor imaging (DTI) is a popular MRI-based technique which allows to identify fiber tracts connecting brain regions by estimating the diffusion of the water molecules along their main direction. More specifically, the three-dimensional diffusion of water is mapped and characterized as a function of spatial location. The diffusion tensor describes the magnitude, the degree of anisotropy, and the orientation of diffusion anisotropy, that is how the water molecules differently move in the directions parallel to the fiber tracts rather than in the two orthogonal dimensions. Many different measures of diffusion anisotropy have been proposed to visualize and quantify the properties of the diffusion tensor [2]. The most commonly used parameters are fractional anisotropy (FA), a measure of the orientation of diffusion, and (rotationally indifferent) mean diffusivity (MD). DTI has been suggested as an indirect marker for white-matter integrity. For example, in epilepsy, the epileptogenic hippocampus demonstrates increased MD and decreased FA [3].

The two chapters by Crispino et al. and Cabassi et al. in this volume (pp. 1 and 37) provide interesting modeling approaches for the analysis of DTI data. Cabassi et al. argue that the quality of diffusion-weighted images could be affected by several types of artifacts, due to the low signal-to-noise ratio and the relatively long scan time required by the DTI tractography [4]. In particular, those artifacts may cause underestimation of diffusion coefficients and bias anisotropy measures. To address such issues, Cabassi et al. propose a hierarchical Bayesian model to estimate the effective unknown number of white matter fibers connecting each pair of brain regions. More precisely, they assume a discrete measurement error model, where each

observed white matter fiber count is assumed to be Binomially sampled from the true unknown population of white matter fibers, which is assigned a latent Poisson prior. The model leverages available information both at the subject and the brain region scale. These results provide some evidence that the fiber-counts may be indeed severely underestimated.

The chapter by Crispino et al. provides an exploratory analysis of how structural connectivity may inform patterns of activation captured by functional imaging techniques among regions of interest (ROIs). This is an issue which we will discuss again in Sect. 6.2 later on in this chapter. In their latent space model for the DTI data, Crispino et al. consider the structural imaging data as an observed network of connections between ROIs and model the probability of observing an edge in the network (i.e., the probability that at least one white matter fiber connects two ROIs) as a function of how close/far the regions are. They conclude that the inferred latent space of the DTI is highly correlated with the physical one represented by ROIs locations, although the two may not completely overlap.

Also Durante and Dunson [5] have recently developed a statistical model to infer expected network structures from DTI data, which takes into account that fiber tracking pipelines are subject to measurements error. More specifically, they consider a latent variable framework, where the probability mass function of the network is characterized using a mixture of low-rank factorizations. Within each mixture component, connections among pairs of nodes are characterized as conditionally independent Bernoulli random variables given component-specific edge probabilities, which are further obtained as a function of node-specific latent variables. The model allows for group dependence in the mixing probabilities, which can be used to conduct global and local testing for differences in brain connectivity networks between two groups of subjects. The study of undirected connections estimated from structural imaging data will certainly be the objective of further investigations in the future.

## 2.2 Functional Imaging

Functional brain imaging involves the study of brain functioning, both in terms of its *specialization* (i.e., which parts of the brain respond to a given task) as well as its *integration* (i.e., how different brain regions interact with each other). Perhaps the two most popular functional brain imaging techniques are electroencephalography (EEG) and functional MRI (fMRI). EEG data record the electrical activity of the brain from the scalp. They are characterized by high temporal resolution. However, they present low spatial resolution, due to the configuration of the electrodes on the scalp. Due to the early influence of signal processing, statistical methods for EEG data often involve spectral time series representations of the temporal signal. With respect to EEG data, fMRI data are characterized by higher spatial resolution but lower temporal resolution. fMRI data provide an indirect measure of brain activity, since they record the metabolic activity in the brain, as represented by differences

in local blood flow (blood-oxygen level-dependent, BOLD, signal). It is beyond the scope of this chapter to provide further details about the physiology of fMRI signals, for which we refer to Poldrack et al. [6]. We only mention that, due to their high spatial resolution, fMRI data have been typically employed to identify changes in brain activity across different brain regions, and also over time, although their ability to identify brain events over very short time periods may be somewhat limited.

**The analysis of fMRI data** Statistical methods for fMRI data vary widely according to the experiment design (e.g., task-based or resting-state experiment) and the objectives of the study. In a task-based experiment, for example, the whole brain is scanned at multiple times while a subject performs a series of tasks. Therefore, a typical objective is to detect which brain regions get activated by the external stimuli (*activation detection*). Statistical methods for this analysis typically include linear and nonlinear models, as well as mixture models, for both single- and multiple-subject studies.

The chapter by Gasperoni and Luati in this volume (p. 91) highlights the importance of taking into proper account the physiology of the different neuroimaging experiments in the statistical analysis of fMRI data. The hemodynamic response function (HRF) models the vascular response to neuronal activity, which contributes to the observed fMRI signal. Since the estimation of neural activity is a major interest of fMRI studies, the interpretation of fMRI findings may be severely impaired if the hemodynamic response were not accurately taken into account in the analysis [7, 8]. The HRF varies considerably over different brain regions and across subjects. Most fMRI studies have primarily focused on estimating the amplitude of evoked HRFs in task-based experiments. However, the influence of the hemodynamic response has been shown also in resting state experiments, to characterize the BOLD signal in response to spontaneous neuronal activity. For example, Rangaprakash et al. in [9] have shown that the variability of the HRF across the brain may alter functional connectivity estimates obtained from resting-state fMRI. In their chapter, Gasperoni and Luati extend a multi-step blind-deconvolution approach, first presented in [10], to estimate the HRF from spontaneous brain activity. In particular, they robustify the procedure by assuming a Student-*t* distribution for the noise affecting the BOLD signal and then they identify spontaneous activations as extreme values of the residuals obtained from a robust procedure for signal extraction. They discuss how the method based on the assumption of a Student-*t* distribution for the noise should select a smaller number of spontaneous activations than the method based on the Gaussian assumption. This is certainly an area of continuous interest in the fMRI literature, as it affects the validity of any subsequent inferences.

**Brain connectivity** Another important task in fMRI studies, which has received increased attention in recent years, is to infer *brain connectivity*. In general terms, connectivity looks at how brain regions interact with each other and how information is transmitted between them, with the aim of uncovering the actual mechanisms of how our brain functions. In particular, it is customary to distinguish between functional (undirected) and effective (directed) connectivity, as first defined by [11]. In the study of *functional connectivity*, the goal is to identify multiple brain areas that

exhibit similar temporal profiles, either task-related or at rest. On the other hand, effective connectivity seeks to estimate the directed influence of one brain region on another. In the classical literature, simple approaches to capture functional connectivity are based on temporal correlations between regions of interest, or between a “seed” region and other voxels throughout the brain. Alternative approaches include clustering methods, to partition the brain into regions that exhibit similar temporal characteristics, and multivariate methods for dimension reduction, such as Principal Components Analysis (PCA) [12] and Independent Components Analysis (ICA) [13], which determine spatial patterns that account for most of the variability in the time series data. Approaches that allow to estimate partial correlations between pre-defined regions of interest (ROIs) have also been proposed, for example by using the graphical Lasso (GLasso), which estimates a sparse precision matrix [14].

In the Bayesian literature, Bowman et al. (2009) in [15] employed a two-stage modeling approach to capture short-range task-related (or between-group) connectivity between voxels within a given anatomical region. The model assumes that voxels within anatomically defined regions exhibit task-related activity that deviates around an overall mean for that region. By appropriate modeling of a flexible unstructured covariance matrix for regional mean parameters, the model allows to estimate spatial correlations which are interpretable as task-related functional connections. This modeling framework also allows to develop a measure of inter-regional (or long-range) connectivity between two regions. Long-range connectivity is observed whenever relatively distant pair of voxels exhibit high positive correlations, even when compared to a more proximal pair of voxels. For example, Broca’s area and Wernicke’s area are two noncontiguous anatomical regions that may exhibit long range correlations, given their joint involvement in speech generation, processing and understanding. More recently, Zhang et al. in [16] allow clustering of spatially remote voxels that exhibit fMRI time series with similar characteristics, by imposing a Dirichlet Process (DP, [17]) prior on the parameters of a long memory error term. The induced clustering can be viewed as an aspect of functional connectivity, as it naturally captures statistical dependencies among remote neurophysiological events.

Many of the chapters in this volume have focused on estimating functional connectivity. For example, the chapter by Caponera et al. in this volume (p. 111) proposes an elegant Bayesian time-dependent latent factor model, where the factor loading matrix can be interpreted as a simple measure of connectivity. Their method can be seen as a further contribution to the collection of multivariate methods for dimension reduction discussed above. A key assumption of their approach, which we will discuss further on in this chapter, is stationarity, i.e., the spatial dependence structure is assumed constant over time. Another interesting aspect of their work is the discussion of the graph theoretical approach to explore functional connectivity networks, according to the paradigm of analysis discussed in Bullmore and Sporns [18]. In neurological applications it is common practice to report the brain network structure by thresholding the estimated association measures (e.g., correlation matrices). The thresholding generates binary adjacency matrices which can be used to compute network indices to summarize the topological properties of the network. A vast number of graph theory measures of network topology have been recently studied in various neuro-

logical diseases. The majority of those features relate to various aspects of global network integration or local segregation. A relevant subset of features identifies the nodes that have a strong influence on the communication of the network, which are known as centrality or hub measures. The simplest of those centrality measures is degree centrality, which counts the number of edges connected to each node. Other centrality measures capture more nuanced quantities, such as eigenvector centrality, which identifies nodes that are connected to other highly central nodes, or betweenness centrality, which captures the number of shortest paths that pass through a node [19]. In addition, deviations from a small-world configuration have been consistently found to characterize various types of brain diseases, including Alzheimer's disease, epilepsy, brain tumors, and traumatic brain injury [20]. Therefore, by investigating the inference on the graph theory measures of network topology induced by a particular modeling approach, it is possible to achieve additional understanding about the clinical implications of the estimated functional connectivity networks.

An alternative approach, which has been explored in a few chapters of this volume, regards the fMRI time-series as instances of functional data, to be considered in an object-oriented data analysis in non-Euclidean spaces. Instead of comparing networks based on a set of connectivity measures summarizing the topological properties of functional brain networks, the chapter by Cabassi et al. in this volume develops a procedure for testing group differences in the network structure based on several types of non-Euclidean metrics. Also Ginestet et al. (2017) in [21] have recently proposed to employ statistical inference on manifolds to develop one- and two-sample tests for network data objects. Similarly, the chapter by Cappozzo et al. in this volume (p. 57) considers a functional data analysis approach to define a rescaled covariance operator for functional random processes, in the Riemannian manifold defined by positive semi-definite symmetric matrices. All contributions show that global tests may result in more statistical power than when using a mass-univariate approach, which is the standard approach in the field. On the other hand, global tests may be limited as in practice the interest of many investigators is often focused on local discrepancies in the network structure. Methodologies, like the one in [5], which allow for both global and local testing of differences in brain connectivity networks, may perhaps be adapted to this object-oriented data analysis framework.

Differently than functional connectivity, which relates to undirected associations between time series, *effective connectivity* refers to the influence that “one neural system exerts over another” [22]. Effective connectivity refers to causal dependence, as opposed to simple association. Therefore, commonly used approaches for capturing effective connectivity include many of the methods typically employed to represent causal relationships: structural equation modeling (SEM, [23, 24]), dynamic causal modeling (DCM, [25]) vector autoregressive (VAR) models [26], Granger causality [27] and Bayesian networks [28]. It should be pointed out, however, that even though such methods allow inference on directed connections between brain regions, they do not necessarily imply *physiological* causality. Due to the nature of fMRI experiments, the models can only be used to assess causality at the hemodynamic level rather than the neuronal level. Brain scientists are typically more interested to make inference on neural activity. However, the connectivity estimated at

the hemodynamic level can still yield interesting results. More appropriately, physiological causality should be assessed through a carefully crafted experimental design [29]. In particular, the often-used notion of Granger causality is based on the idea that causes always precede effects. Therefore, past signal values from one brain region can be used to predict current values in another region. Gorrostieta et al. in [30] have developed a Bayesian hierarchical VAR model for investigating Granger causality and effective connectivity in multiple subjects, accounting for the variability in the connectivity structure within and between subjects. Yu et al. in [31] have further extended this framework, for simultaneously estimating brain activation and effective connectivity in a study of how brain motor function is altered in patients who have suffered a stroke, with respect to healthy subjects. With the hierarchical structure, subject-specific estimates for activation and connectivity are obtained by pooling information from other subjects. The approach allows to study local activation and connectivity between brain regions, and to compare the inferred patterns for stroke patients and healthy controls in order to explore the effects of stroke on brain motor function.

In this section, we have provided a limited overview of the main goals typically associated with functional imaging studies. We refer to [32] for a review of modeling approaches to study functional and effective connectivity, causal modeling, connectomics, and multivariate analyses of distributed patterns of brain responses. Bowman in [33] provides a more extensive background on various types of neuroimaging data and analysis objectives that are commonly targeted in brain imaging studies. Stephan and Friston in [34] provide an extensive review of the conceptual and methodological basis of linear and nonlinear DCMs for characterizing effective connectivity using fMRI data.

In the following sections, we discuss a few of the most recent interests and arising challenges in the analysis of neuroimaging data.

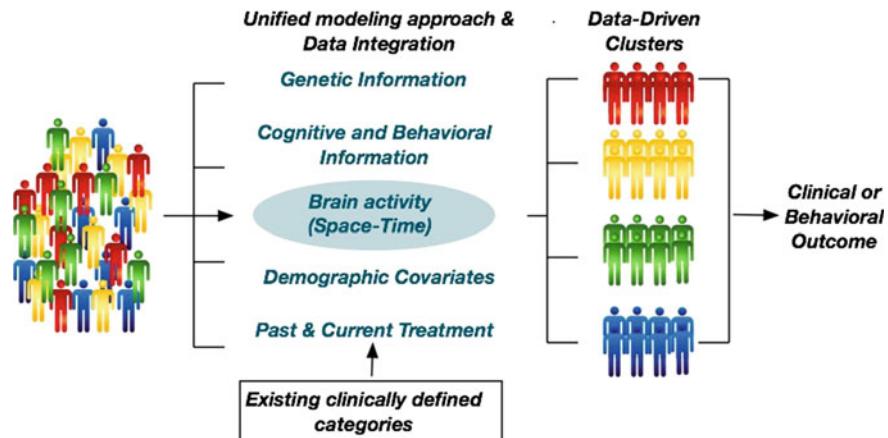
### 3 Describing the Heterogeneity of Brain Mechanisms

One of the main objectives in the analysis of brain imaging data is to characterize the heterogeneity typically observed both within- and between- subjects, especially in subjects affected by behavioral and psychiatric disorders. An improved understanding of the heterogeneity of brain mechanisms is considered key for enabling clinicians to deliver targeted, precision, medicine to individuals affected by such disorders. Current medical practice often relies on symptom-based diagnostic criteria. Despite the progress enabled by neuroimaging technologies in the understanding of the pathophysiology of the major psychiatric disorders, the diagnosis or treatment of individual patients have not been yet significantly impacted by such revolution [35]. On the other hand, traditional diagnostic criteria are increasingly recognized as inappropriate to describe the variety of the disorders actually observed in individuals, which are progressively seen as the result of the interplay of different characteristics [36, 37]. In 2010, the United States National Institutes of Mental

Health (NIMH) started the Research Domain Criteria (RDoC) project to develop new ways for classifying mental disorders, on the basis of experimental research criteria rather than traditional diagnostic categories. The RDoC assumes that further insights and progress in the understanding and diagnosis of psychiatric disorders will be achieved by integrating many levels of information (from genomics to neuroimaging and self-reports). This holistic approach will allow to investigate both the normal and the disrupted dimensions of brain functioning and human behavior at a deeper level than it has been currently achieved. The ultimate long-term goal of the NIMH RDoC initiative is precision medicine. Data from genetics and clinical neuroscience will eventually allow the identification of prognostic and predictive biomarkers. That is, the goal is to develop an analytical framework that allows to incorporate the specific genomic and neuroimaging characteristics of a subject into a predictive decision-making paradigm, so that clinicians may optimize the choice of individual treatments based on their expected predicted outcome [38, 39].

Statistics can provide innovative tools for a data-driven classification of subjects, by combining the neuro-imaging data with the available genomic, behavioral and clinical information on the subjects. Figure 1 illustrates the general scheme underlying the unified approach sought for better understanding the heterogeneity of the brain disorders. Here, we will focus on a few approaches that can be used to capture the main sources of variability in fMRI data, with respect to

- identifying clusters of subjects, characterized by similar patterns of brain responses to a task;
- characterizing the heterogeneity in the individual dynamics of functional connectivity networks;



**Fig. 1** Understanding the heterogeneity of the brain disorders based on neuro-imaging data and other information on the subjects in a unified framework is key for attaining the goal of precision medicine

- (c) relating the observed imaging patterns to additional available information on the same subjects, including genetic covariates and other observable clinical or behavioral outcomes.

## 4 Clustering Subject-Specific Imaging Patterns

In single-subject analysis, the clustering of fMRI time-series has emerged as a way to classify the regions of the brain according to the temporal pattern of the BOLD response. For example, the chapter by Bertarelli et al. in this volume (p. 75) proposes  $k$ -means and functional clustering approaches to cluster fMRI time-series beyond the traditional statistical methods which are typically used to evaluate the level of activation of individual voxels. In the analysis of fMRI data, unsupervised clustering methods have been used also in the context of Gaussian mixture models applied to processed data (either “contrast” maps or simple  $z$ -statistic images), to capture distinct clusters of activations, e.g., for pre-surgical assessment of peritumoral brain activation [40, 41]. Alternatively, Zhang et al. (2014) in [16] provide a joint analytical framework to detect regions of the brain which exhibit neuronal activity in response to a stimulus and, simultaneously, infer the association, or clustering, of spatially remote voxels that exhibit fMRI time series with similar characteristics.

In multi-subject analyses, clustering methods have been used to identify groups of subjects that are characterized by similar patterns of brain activity. The chapter by Cappozzo et al. in this volume proposes functional clustering of networks based on the definition of a suitable distance between covariance operators, or alternatively on a low dimensional representation of the correlation matrices. Woolrich et al. in [42] and Xu et al. in [43] model the inter-subject variability in brain activity via (possibly infinite) Gaussian mixture models that estimate the probability that an individual has an activation at a particular location. Zhang et al. in [44] leverage on more advanced multi-level Bayesian nonparametric approaches to allow for the separate inferential objectives within and between subjects. More precisely, they employ a hierarchical Dirichlet Process prior construction to induce clustering among voxels within a subject at one level of the hierarchy and across subjects at the second level. This formulation allows, in particular, to capture spatial correlation among potential activations of distant voxels, within a subject (an aspect of functional connectivity), while simultaneously borrowing strength in the estimation of the parameters from subjects with similar activation patterns. Let  $Y_{iv} = (Y_{iv1}, \dots, Y_{ivT})^\top$  be the  $T \times 1$  vector of the BOLD response data at the  $v$ th voxel in the  $i$ th subject, with  $i = 1, \dots, N$ ,  $v = 1, \dots, V$ , and with the symbol  $(\cdot)^\top$  indicating the transpose operation. The BOLD time-series response is then modeled with a general linear model

$$Y_{iv} = X_{iv}\beta_{iv} + \varepsilon_{iv}, \quad \varepsilon_{iv} \sim N_T(0, \Sigma_{iv}), \quad (1)$$

where  $X_{iv}$  is a  $T \times p$  covariate matrix,  $\beta_{iv} = (\beta_{iv1}, \dots, \beta_{ivp})^\top$  is a  $p \times 1$  vector of regression coefficients and  $\varepsilon_{iv} = (\varepsilon_{iv1}, \dots, \varepsilon_{ivT})^\top$  is a  $T \times 1$  vector of errors. Typically, the matrix  $X_{iv}$  contains the design matrix, i.e., the convolved hemodynamic response function, which captures the change in the metabolism of the BOLD contrast due to an outside stimulus. Thus, each column of  $\mathbf{X}_{iv}$  is modeled through the convolution

$$\int_0^t x(s) h_v(t-s) ds, \quad t = 1, \dots, T$$

of the external time-dependent stimulus function for a given task,  $x(s)$ , which is known and corresponds to the experimental paradigm (for example, a vector defined with elements set to 1 when the stimulus is “on” and 0 when it is “off”), and a parametrically specified hemodynamic response function  $h_v(\cdot)$ . In addition, the matrix  $\mathbf{X}_{iv}$  can also include precision covariates that incorporate motion correction estimates obtained from the preprocessing steps. Of course, additional individual specific covariates may also be included (e.g., demographic and clinical information), depending on the specific study objectives.

In model (1) the detection of brain voxels that activate in response to the stimulus reduces to a problem of variable selection, i.e., the identification of the nonzero  $\beta_{iv}$ 's and is achieved, in the Bayesian framework, by imposing a mixture prior, often called *spike-and-slab* prior, on the regression coefficients. Zhang et al. [44] embed the selection into a clustering framework and effectively define a multi-subject nonparametric variable selection prior with spatially informed selection within each subject. More specifically, they employ a hierarchical Dirichlet Process (HDP) prior [45], which implies that the non-zero  $\beta_{iv}$ 's within subject  $i$  are drawn from a mixture model and possibly shared between subjects. Let  $\gamma_{iv}$  be the binary indicator of whether voxel  $v$  in subject  $i$  is active or not, i.e.,  $\gamma_{iv} = 0$  if  $\beta_{iv} = 0$  and  $\gamma_{iv} = 1$  otherwise. Zhang et al. [44] impose a spiked HDP prior on  $\beta_{iv}$ , i.e., a spike-and-slab prior where the slab distribution is modeled by a HDP prior,

$$\begin{aligned} \beta_{iv} | \gamma_{iv}, G_i &\sim \gamma_{iv} G_i + (1 - \gamma_{iv})\delta_0 \\ G_i | \eta_1, G_0 &\sim DP(\eta_1, G_0) \\ G_0 | \eta_2, P_0 &\sim DP(\eta_2, P_0) \\ P_0 &= N(0, \tau), \end{aligned} \tag{2}$$

with  $\delta_0$  a point mass at zero, with  $\tau$  fixed,  $\eta_1, \eta_2$  the mass parameters and  $P_0$  the base measure. The spike-and-slab formulation enforces sparsity in the pattern of activations within each subject. The HDP prior allows for non-zero coefficients to be shared within and across subjects, potentially highlighting regions characterized by similar intensity of brain activity across subjects. Since the number of mixture components is unknown and inferred from the data, this prior formulation provides an unsupervised clustering framework to account for between-subjects heterogeneity in neuronal activity. In order to take into account information on the anatomical structure

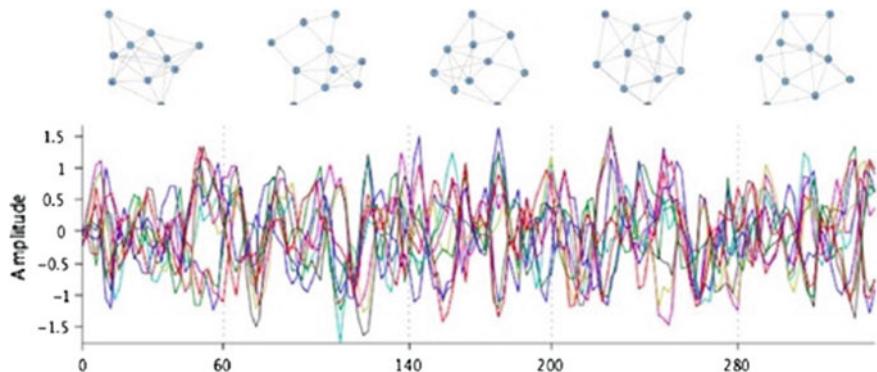
of the brain, in particular the correlation between neighboring voxels, they further place a Markov Random Field (MRF) prior on the selection parameter  $\gamma_{iv}$ .

A single fMRI experiment can yield hundreds of thousands of high frequency time series for each subject, arising from spatially distinct locations. Therefore, computational efficiency is essential for the practical relevance of any statistical method. This is particularly true for multi-subject studies. In particular, Bayesian methods face a significant challenge, since typically Markov chain Monte Carlo sampling algorithms are too slow and inefficient for this type of problems. Thus, there is a need for computational methods which approximate the posterior distribution for faster inference. Variational Bayes methods have been employed successfully in Bayesian models for single-subject fMRI data [46–50]. Typically, these approaches provide good estimates of means, although they tend to underestimate posterior variances and also to poorly estimate the correlation structure of the data. In a comparative study on simulated data, Zhang et al. [44] show that a variational Bayes algorithm approximating the posterior distribution of model (1)–(2) achieves robust estimation results at a much reduced computational costs, therefore allowing scalability of their method. Additionally, they demonstrate on synthetic data how their unified, single-stage, multiple-subject modeling approach, with variational Bayes inference, achieves improved estimation performance with respect to two-stage approaches which may be employed to ease the computational burden of multi-subject analyses.

The availability of user-friendly software implementations is also a required condition for the general adoption of novel statistical methods by the neuroscientist. For example, the model by Zhang et al. [44] has been implemented in a MATLAB GUI (*NPBayesfMRI*, [51]), comprising two components, one for model fitting and another one for visualization of the results. Within the model fitting interface, the user can define the type of analysis (voxel-based or whole-brain parcellation into regions of interest, i.e., ROIs) and the model parameters. Users have the option of a pre-defined default setting for all parameters. Alternatively, they can set the parameters according to customized choices, depending on the available prior information. We should also mention *Neuroconductor* (<https://neuroconductor.org/>), an open-source R-platform for medical imaging analysis [52]. The platform provides data, methods, and software packages designed to support the analysis of populations of images using the publicly available statistical software R.

## 5 Dynamic Functional Connectivity

Behavioral and psychiatric disorders have been associated to differences in the brain functional connectivity networks, i.e., the set of interactions that take place between spatially segregated but temporally related regions of the brain [53]. Traditionally, brain network studies have assumed functional connectivity as spatially and temporally stationary, i.e., connectivity patterns are assumed not to change throughout the scan period [54]. However, in practice, the interactions among brain regions may vary during an experiment. For example, different tasks, or fatigue, may trigger vary-



**Fig. 2** Dynamic functional connectivity assumes that the functional connectivity networks may change over time

ing patterns of interactions among different brain regions. Therefore, more recent work has pointed out that it is more appropriate to regard functional connectivity as *dynamic* over time [55]. Figure 2 provides a pictorial representation of the new paradigm. Current approaches for studying dynamic connectivity typically rely on multi-step approaches for inference, where the analysis may comprise the following steps. First, the fMRI time courses are segmented by selecting a sequence of sliding windows. Then, a covariance (or precision) matrix is estimated separately within each window, e.g., by using graphical Lasso. Finally,  $k$ -means clustering methods are used to identify re-occurring patterns of functional connectivity state [56]. Differences between states are assessed by computing and comparing descriptive graph metrics that capture structural properties of the networks, such as their clustering coefficient and efficiency. Arguably, those approaches are straightforward but present some major limitations. For example, the length of the window is arbitrarily selected before the analysis, through a trial-and-error process. This trial-and-error process can potentially lead to an increased number of false positive and false negative detections in the estimation of the networks, and ultimately affects the reproducibility of the findings. Indeed, Lindquist et al. in [57] show that the choice of the window length can affect inference in unpredictable ways. To partially obviate the issue, Cribben et al. in [58] and Xu and Lindquist in [59] have recently investigated greedy algorithms, which automatically detect change points in the dynamics of the functional networks. Their approach recursively estimates precision matrices using GLasso on finer partitions of the time course of the experiment, and selects the best resulting model based on the Bayesian Information Criterion (BIC). The algorithm estimates independent brain networks over noncontiguous time blocks. Of course, this is not so desirable, as it may be preferable to borrow strength across similar connectivity states in order to increase the accuracy of the estimation. Another issue is related to greedy searches, which often fail to achieve global optima.

Chiang et al. in [60] investigate the stationarity of the brain network topology, as measured by the graph theory measures of functional connectivity networks. The aim

of their study is to identify which aspects of network topology exhibit less within-scan temporal variability in resting-state networks, with the objective of evaluating which graph theory metrics may be robustly estimated using static functional connectivity analyses. In particular, they argue that some aspects of brain topology, such as the level of small-worldness, may exhibit greater temporal stationarity, whereas others, such as local measures, may be more susceptible to local dynamics and more likely to traverse multiple configurations. They use a Bayesian hidden Markov model to estimate the transition probabilities of various graph theoretical network measures using resting-state fMRI (rs-fMRI) data and to investigate the stationarity of different graph theory measures. They further propose two estimators of temporal stationarity, which can be used to assess different aspects of the temporal stationarity of functional networks: a deterministically-based estimator of the number of change-points, and a probabilistically-based estimator that takes into account stochastic variation in the estimated states. They show that small-world index, global integration measures, and betweenness centrality exhibit greater temporal stationarity than network measures of local segregation. This may reflect the organization of the resting-state brain, in which the small-world architecture of the brain is thought to have evolved in order to create systems that support efficiency in both local and global processing. Since long-range connections are generally thought to ensure the interaction between distant neuronal clusters, a large component of fluctuations between neuronal clusters (e.g., long-range connections) may therefore occur downstream to fluctuations within neuronal clusters (e.g., local connections), resulting in slightly greater temporal stationarity among global relative to local connections. On the other hand, connectivity within local subgraphs may be more susceptible to local cell dynamics and likely to fluctuate over time.

The chapter by Crispino et al. in this volume discusses a penalized likelihood approach to estimate time-varying Bayesian networks, based on a first-order Markovian assumption to model the connectivity dynamics. The strength of the interaction between two brain regions is a function of how often two regions are connected by an edge at different time points.

Warnick et al. in [61] propose a principled, fully Bayesian approach for studying dynamic functional network connectivity, that avoids arbitrary partitions of the data in sliding windows. More specifically, they cast the problem of inferring time-varying functional networks as a problem of dynamic model selection in the Bayesian setting. As we have previously discussed, brain networks can be mathematically described as graphs. A graph  $G = (\mathcal{V}, \mathcal{E})$  specifies a set of nodes (or vertices)  $\mathcal{V} = \{1, 2, \dots, V\}$  and a set of edges  $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ . Here, the nodes represent the neuronal units, whereas the edges represent their interconnections. For example, nodes could be intended as either single voxels or macro-areas of the brain which comprise multiple voxels at once. Let  $\mathbf{Y}_t = (Y_{t1}, \dots, Y_{tV})^\top$  be the vector of fMRI BOLD responses of a subject measured on the  $V$  nodes at time  $t$ , for  $t = 1, \dots, T$ . Then, the general linear model (1) can be re-expressed as follows,

$$\mathbf{Y}_t = \boldsymbol{\mu} + \sum_{k=1}^K \mathbf{X}_t^k \circ \boldsymbol{\beta}_k + \boldsymbol{\epsilon}_t, \quad (3)$$

where  $\circ$  denotes the element-by-element (Hadamard) product,  $\mathbf{X}_t^k$  is the  $V \times 1$  design vector for the  $k$ -th stimulus,  $\boldsymbol{\mu}$  the  $V$ -dimensional global mean and  $\boldsymbol{\beta}_k = (\beta_{1k}, \dots, \beta_{Vk})^\top$  the stimulus-specific  $V$ -dimensional vector of regression coefficients. A spike-and-slab prior is imposed on the coefficients  $\beta_{vk}$  to identify brain activations and allow decoupling of the task-related activations from the functional connectivity states. To characterize possibly distinct connectivity states, i.e., network structures, within different time blocks, Warnick et al. (2018) assume that functional connectivity may fluctuate among one of  $S > 1$  different states during the course of the experiment. Let  $s = (s_1, \dots, s_T)^\top$ , with  $s_t = s$ , for  $s \in \{1, \dots, S\}$ , denoting the connectivity state at time  $t$ . Then, conditionally upon  $s_t$ , they assume

$$(\boldsymbol{\epsilon}_t | s_t = s) \sim N_V(0, \boldsymbol{\Omega}_s), \quad (4)$$

where  $\boldsymbol{\Omega}_s \in \mathbb{R}^V \times \mathbb{R}^V$  is a symmetric positive definite precision matrix, i.e.,  $\boldsymbol{\Omega}_s = \boldsymbol{\Sigma}_s^{-1}$ , with  $\boldsymbol{\Sigma}_s$  the covariance matrix. The zero elements in  $\boldsymbol{\Omega}_s$  encode the conditional independence relationships that characterise state  $s$ , that is graph  $G_s = (\mathcal{V}, \mathcal{E}_s)$ . Specifically,  $\omega_{ij}^{(s)} = 0$  if and only if edge  $(i, j) \notin \mathcal{E}_s$ . Many of the estimation techniques for Gaussian graphical models rely on the assumption of sparsity in the precision matrix, which is generally considered realistic for the small-world properties of brain connectivity in fMRI data. Thus, a  $G$ -Wishart distribution is considered as a conjugate prior on the space of the precision matrices  $\boldsymbol{\Omega}$  with zeros specified by the underlying graph  $G$  [62, 63]. The estimation of the unknown connectivity states at each of the time points is treated as a problem of change points detection, by modeling the temporal persistence of the states through a Hidden Markov Model (HMM). The approach is in line with recent evidence in the neuroimaging literature which suggests a state-related dynamic behavior of brain connectivity with recurring temporal blocks driven by distinct brain states [64, 65]. In the model proposed by Warnick et al. (2018), however, the change points of the individual connectivity states are automatically identified on the basis of the observed data, thus avoiding the use of a sliding window. Furthermore, they adapt a recent proposal put forward by Peterson et al. in [66] to conduct inference on the multiple related connectivity networks. The model formulation assumes that the connectivity states active at the individual time points may be related within a *super-graph* and imposes a sparsity inducing Markov Random field (MRF) prior on the presence of the edges in the super-graph. Thus, the estimation of the active networks between two change points is obtained by borrowing strength across related networks over the entire time course of the experiment, also avoiding the use of post-hoc clustering algorithms for estimating shared covariance structures.

## 6 Combining Information from Multiple Data Sources

The term “big data” is often employed to indicate the high-dimensionality and the complexity of data captured by modern technologies. With this meaning, brain imaging data can be regarded as inherently “big”. However, in Sect. 2, we have described how each neuroimaging technology is able to capture only specific characteristics of brain processes. Therefore, each single technology is also inherently limited in its ability to shed light on relevant brain mechanisms. Multi-modal analysis combines different neuroimaging modalities, and possibly information from different data platforms, to achieve a more comprehensive understanding of brain functioning. In this section, we review some recent interesting trends and contributions in this area.

### 6.1 Covariate-Dependent Analysis and Predictive Modeling

It is often of interest to study how imaging-based inferences vary depending on known covariates or risk factors, and to make predictions on a clinical or behavioral response based on the estimated individual’s brain activity.

For example, the chapter by Aliverti et al. in this volume (p. 23) proposes a sequential hierarchical approach, which starts by using a penalized GLasso approach to estimate functional connectivity. Then the connection probabilities are modeled through a latent logit regression involving both phenotypical and brain-region information. The covariates include the age of the subject, an indicator of mental health diagnosis, and another indicator of shared lobe membership for each pair of edges.

As an example of a modeling approach aimed at improving clinical prediction, we refer to Chiang et al. in [67]. They consider positron emission tomography (PET) imaging data from a study on Temporal lobe epilepsy (TLE), the most common form of adult epilepsy and the most common epilepsy refractory to anti-epileptic drugs. PET imaging is a well-developed technique in which the subject is injected with a positron-emitting isotope, such as  $^{18}\text{F}$ -FDG, and a PET image reconstructed of the isotope concentration based on the incidence of gamma rays from the positron-electron annihilation. In PET studies, the quantity that is clinically assessed is a scalar rate of regional glucose uptake. This quantity is then normalized relative to an internal reference standard, such as the whole-brain activity and compared to the expected level for a normal subject. The assessed quantity therefore provides a measure of the level of metabolic activity in each region, relative to that expected in healthy controls. Uptake levels may be quantified on the single-pixel level or based on the mean uptake within fixed regions of interest. Chiang et al. (2017) develop a Bayesian predictive modeling framework to identify whole-brain biomarkers from PET imaging which are associated to the prediction of post-surgical seizure recurrence following anterior temporal lobe resection. Post-surgical seizure recurrence is often due to the incomplete resection of the epileptogenic zone, which is defined as the area of cortex necessary and sufficient for initiating seizures, and whose removal

is necessary for seizure abolition. Indeed, patients with different epileptogenic zone configurations may be expected to exhibit different risks of post-surgical seizure recurrence. The epileptogenic zone, however, cannot be identified pre-operatively. In their model formulation, Chiang et al. (2017) take this into account by looking at the observed PET brain measurements as the phenotypic manifestation of latent individual pathological states that are assumed to vary across the population. More precisely, the joint distribution of the data is factored into the product of two conditionally independent submodels, an outcome model that relates the post-surgical outcome to the latent states, and a measurement model that relates those latent states to the observed brain measurements. For the latter, they employ mixture models for clustering and variable selection priors that capture spatial correlation among neighboring brain regions. Thus, subjects are clustered into subgroups with different latent states, i.e., different epileptogenic zone configurations, while simultaneously identifying discriminatory brain regions that characterize the subgroups. A logistic regression model relates the latent states to the binary clinical outcome. Alternative predictive modeling approaches for neuroimaging include the use of pattern recognition techniques, such as Linear Discriminant Analysis [68], Support Vector Machines [69, 70] and Bayesian classifiers [71, 72]. We refer to the review in [73] for a discussion of Bayesian methods for classification and prediction.

## 6.2 Multi-modal Imaging Analysis

Multi-modal imaging refers to imaging performed using different instrumentation platforms, although a given modality may also provide multiple types of imaging outcomes. The objective is to obtain a more accurate understanding of brain processes by combining two or more datasets obtained with different instruments. For example, in the study of epilepsy, simultaneous acquisition of EEG and fMRI has been employed to improve the spatio-temporal resolution of either data with the aim of localizing epileptic foci [74]. Statistical models for multi-modal analysis are necessarily integrative. In particular, Bayesian methods are well suited for the analysis of multi-modal data, due to their ability to integrate the data into a hierarchical model. We refer to the reviews in [75, 76] for a discussion of general strategies for multi-modal analysis and to [73] for a review of Bayesian methods. Jorge et al. in [77] present a review of the most relevant EEG-fMRI integration approaches for the study of human brain function.

For example, Kalus et al. in [78] use EEG-informed spatial priors in their Bayesian variable selection approach to detect brain activation from fMRI data. Specifically, they relate the prior activation probabilities to a latent predictor stage  $\zeta = (\zeta_1, \dots, \zeta_v)^\top$  via a probit link  $p(\gamma_v = 1) = \Phi(\zeta_v)$ , with  $\Phi$  the standard normal cdf and  $\zeta_v$  consisting of an intercept term and an EEG effect, that is

$$\zeta_v = \zeta_{0,v} + \zeta_{EEG,v} = \begin{cases} \zeta_{0,v}, & \text{if predictor 0} \\ \zeta_{0,v} + \zeta_G J_v, & \text{if predictor } glob, \\ \zeta_{0,v} + \zeta_v J_v, & \text{if predictor } flex \end{cases} \quad (5)$$

where  $J_v, v = 1, \dots, V$  is the continuous spatial EEG information and where 0, *glob* and *flex* indicate three types of predictors: predictor 0 contains a spatially-varying intercept  $\zeta_0 = (\zeta_{0,1}, \dots, \zeta_{0,V})^\top$ , and corresponds to an fMRI activation detection scheme without incorporating EEG information; predictor *glob* contains a global EEG effect  $\zeta_G$  in addition to the intercept; predictor *flex* contains a spatially-varying EEG effect  $\zeta = (\zeta_1, \dots, \zeta_V)^\top$ .

An interesting avenue of research is the development of methods for the integration of fMRI and structural imaging data. Here, we mention a recent proposal by Chiang et al. in [79], where the authors develop a multi-subject multi-modal vector autoregressive (VAR) modeling approach for inference on effective connectivity based on resting-state functional MRI data. More in detail, their method uses Bayesian variable selection techniques to allow for simultaneous inference on effective connectivity at both the subject- and group-level. Furthermore, it accounts for multi-modal data by integrating structural imaging information into the prior model, encouraging effective connectivity between structurally connected regions.

### 6.3 Imaging Genetics

Recent developments in molecular genetics have lowered the cost of individual genetic profiling, creating the opportunity to collect massive amounts of genetic information and neuroimaging data on the same subjects. Thus, the field of imaging genetics has emerged as a promising approach for investigating the genetic determinants of brain processes and related behaviors or psychiatric conditions. Ultimately, the objective is to identify specific brain activity features and genetic variants that can be used as biomarkers to assist medical decision making. However, the high-dimensionality and complexity of the data add challenges to statistical analysis. On one hand, there is a problem of variable selection and multiple decision testing, due to the large number of variables' calls and the necessity to identify a sparse set of relevant fMRI features or genetic covariates. On the other hand, naive multi-step multivariate approaches may lead to results that are difficult to interpret, especially if existing biological information is not incorporated at some stage of the analysis.

Nathoo et al. in [80] provide a comprehensive review of recent statistical approaches for the joint analysis of high-dimensional imaging and genetic data, with particular consideration for approaches proposed within the frequentist paradigm. In particular, they distinguish massive univariate and voxel-wise approaches, where the spatial association among separate brain regions is not explicitly modeled, from more sophisticated multivariate approaches, either through regression techniques or low rank regression, mixture models, and group sparse multi-modal regression.

In the Bayesian literature, Stingo et al. in [81] have proposed a hierarchical mixture model based on ROI summary measures of BOLD signal intensities measured on schizophrenic patients and healthy subjects. The model incorporates prior knowledge via network models that capture known dependencies among the ROIs. More specifically, let  $\{x_{ij}, i = 1, \dots, n, j = 1, \dots, p\}$  indicate the ROI-based summaries of BOLD signal intensity on a set of  $p$  features (the anatomical ROIs) in  $n$  subjects. The authors envision that some of the features could discriminate the  $n$  subjects into  $K$  separate known groups (e.g., schizophrenia cases and healthy controls). Let  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_p)^\top$  be a latent binary vector such that  $\gamma_j = 1$  if the  $j$ -th feature is discriminatory and  $\gamma_j = 0$  otherwise. By employing a discriminant analysis framework, they model the data as a mixture model of the general type

$$f_k(x_{ij}|\gamma_j) = (1 - \gamma_j) f_0(x_{ij}; \theta_{0j}) + \gamma_j f(x_{ij}; \theta_{kj}), \quad k = 1, \dots, K, \quad (6)$$

where  $f_0(x_{ij}; \theta_{0j})$  describes the distribution of the “null” model for the non-discriminatory features, while  $f(x_{ij}; \theta_{kj})$  is the distribution of the measurements on the discriminatory features for subjects in group  $k$ . Gaussian distributions are assumed for the mixture components, that is  $f_0(x_{ij}; \theta_{0j}) = N(0, \sigma_{0j}^2)$ , and  $f(x_{ij}; \theta_{kj}) = N(\mu_{kj}, \sigma_{kj}^2)$ . A spatial MRF prior that captures available knowledge on connectivity among regions of the brain is employed to select ROIs that discriminate schizophrenic from healthy controls:

$$P(\gamma_j|\gamma_i, i \in N_j) = \frac{\exp(\gamma_j F(\gamma_j))}{1 + \exp(F(\gamma_j))}, \quad (7)$$

where  $F(\gamma_j) = e + f \sum_{i \in N_j} (2\gamma_i - 1)$  and  $N_j$  is the set of direct neighbors of feature  $j$  in the network. The parameter  $e$  controls the sparsity of the model, while higher values of  $f$  encourage neighboring features to take on the same  $\gamma_j$  value. Note that if a feature does not have any neighbor, then its prior distribution reduces to an independent Bernoulli, with parameter  $\exp(e)/[1 + \exp(e)]$ , a prior often adopted in the Bayesian variable selection literature. The model also allows the group-specific components to depend on selected covariates (e.g., single nucleotide polymorphisms—SNPs) measured on the individual subjects. Let  $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{iR})^\top$  denote the set of available covariates for the  $i$ -th individual. The vectors of the means of the discriminating components are modeled as subject-specific parameters

$$\boldsymbol{\mu}_{ik(\gamma)} = \boldsymbol{\mu}_{0k(\gamma)} + \boldsymbol{\beta}_{k(\gamma)}^\top \mathbf{Z}_i, \quad k = 1, \dots, K, \quad (8)$$

where  $\boldsymbol{\mu}_{0k(\gamma)}$  is a baseline process which captures long-range brain connectivity and  $\boldsymbol{\beta}_{k(\gamma)}$  is a  $R \times p_\gamma$  matrix of coefficients describing the effect of the covariates on the observed measurements. This model formulation uses component-specific parameters that determine how covariates, and other relevant spatial characteristics, affect the observed measurements  $\mathbf{x}_{i(\gamma)}$ , on the  $n$  subjects, given the selected features. In this respect, the classification of the  $n$  subjects in  $K$  groups is driven by the

subjects' covariates. Different covariates are allowed to affect the individual mixture components, by modeling the  $\beta_{k(\gamma)}$  through spike-and-slab priors. Posterior inference will result in the simultaneous selection of a set of discriminatory ROIs and the relevant SNPs, together with the reconstruction of the correlation structure of the selected regions.

More recently, Greenlaw et al. in [82] have developed a hierarchical Bayesian model with regularizing shrinkage priors, such that the posterior mode corresponds to the estimator proposed by Wang et al. in [83], in order to obtain uncertainty estimates on the regression parameters. Chekouo et al. in [84] have extended the proposal in [81] by developing an integrative Bayesian risk prediction model, which directly links genetic and imaging data with the clinical outcome (e.g., a clinical diagnosis of schizophrenia). The model allows for the identification of a regulatory network between SNPs and ROI intensities, thus exploiting the imaging features as an intermediate phenotype, and further assumes that: (i) genetic factors may affect non-discriminatory brain regions (as endophenotypes); and that (ii) genetic factors may be independently associated with disease status without the mediation of a discriminatory imaging endophenotype. With respect to other approaches, the risk predictive framework allows a direct assessment of the individual probability of being affected by schizophrenia as a function of the observed fMRI and SNP biomarkers, and can also be seen as an extension of recently proposed scalar-on-image regression models to the challenging setting of imaging genetics.

## 7 Conclusions

The chapters in this volume provide a stimulating outlook over many current trends in the analysis of brain imaging data. Well-thought statistical models contribute to a deeper understanding of brain functioning, and its disruption as a consequence of disease. The approaches need to take appropriately into account the physiology of the different neuroimaging experiments. However, the involvement of a large community of statisticians in the analysis of this type of data is relatively recent. The section on *Statistics in Imaging* of the American Statistical Association was only founded in 2012, with the goal to increase the influence of statistics and statisticians on imaging science.

All the contributions in this volume show how the use of novel advanced statistical methods could contribute greatly to future developments in neuroimaging. For example, the chapters by Cabassi et al. and by Cappozzo et al. call attention to the possibilities offered by recent developments in object-oriented data analysis in non-Euclidean spaces. The chapter by Bertarelli et al. also proposes functional data approaches for clustering fMRI time-series. The chapter by Gasperoni and Luati uses a modern robust filtering method for detecting spontaneous activations in resting state fMRI time series and thus improving the estimation of the hemodynamic response function. The chapter by Caponera et al. emphasizes the use of established spatio-temporal modeling techniques to take appropriately into account the depen-

dence structure of the data, achieve dimension reduction, and provide an interpretable assessment of functional connectivity across brain regions. The chapter by Aliverti et al. uses a sequential hierarchical approach that leverages multiple available methods in literature, in order to remove noise from the fMRI signal, estimate the functional brain connectivity networks and investigate the association between phenotypes and functional connectivity patterns. Finally, the chapter by Crispino et al. employs latent space models from network analysis to estimate the structural connectivity information provided by DTI data and examine how structural connectivity may inform patterns of activation captured by functional imaging techniques among regions of interest.

The fast developments in the Neurosciences will keep proposing new challenges to the applied statistician. Multimodal analysis, imaging genetics, and predictive modeling techniques are still at their infancy, in the attempt to identify satisfactory biomarkers for targeted intervention. Novel efficient algorithms may fully exploit the information of existing technologies. For example, fMRI time courses are originally complex-valued signals giving rise to both magnitude and phase data. However, most studies—including all those discussed in this volume—typically use only the magnitude signals and thus irreversibly discard half of the data that could potentially contain important information. Multiple studies show that detectability in low signal-to-noise regions of magnetic resonance images is improved by using the full complex-valued fMRI data. Yu et al. in [85] have recently proposed a Bayesian variable selection approach for detecting brain activation at the voxel level from complex valued fMRI data, where inference is conducted via a complex-valued extension of the Expectation-Maximization (EM) algorithm for Bayesian variable selection of [86] that allows for fast detection of active voxels in large-dimensional complex-valued fMRI. By considering both the real and imaginary information, their approach is able to detect more true positives and less false positives than magnitude-only models, especially when the signal-to-noise ratio is small.

New high-resolution imaging technologies promise to deliver more accurate representations of brain processes. In the last few years, the US NIH Brain Initiative has sponsored multiple grants for developing several next generation human imaging techniques. For example, investigators at University of California, Berkeley are now working on MR Corticography (MRCoG), a new tool for studying neuronal circuitry that improves resolution by an order of magnitude, making it possible to visualize cortical layers and microcircuit columns throughout the whole brain. Researchers at Stanford University are developing a novel PET photon detector concept that promises to enhance substantially PET image reconstruction and should permit joint PET-MR (magnetic resonance) imaging. Joint PET-MR collection would allow multi-modal, simultaneous image acquisition of neuron receptor function, functional MR, and high-resolution neuroanatomy. Other technological developments promise to enhance the spectrum of experimental designs available to investigators. Boto et al. in [87] have recently introduced a magnetoencephalography system that can be worn like a helmet, allowing free and natural movement during scanning. The system would make it easier to conduct experiments with subjects who are traditionally difficult to study under a fixed scanner, such as young children with epilepsy or patients

affected by Parkinson's disease. One of the experiments conducted by the investigators to test the new technology also included a simple "ping-pong" ball-game in which subjects were asked to bounce a table tennis ball on a bat!

In addition to new technologies, new directions of research will surface. For example, the so-called gut-brain axis has been recently implicated in multiple conditions. The enteric nervous system in our abdomen has been shown to communicate directly with the brain through the vagus nerve, which connects the brain with many of our major organs. For this reason, the enteric nervous system is often referred to as our "second brain". Feelings of appetite and satiety are mediated through complex pathways where gut hormones play crucial roles. Understanding the brain-gut mechanisms of appetite and weight control may help the identification of novel therapeutic interventions. The gut microbiome has been implicated also in the development of irritable bowel syndrome as a consequence of anxiety and stress, as well as of neurological/behavioral disorders like autism, ADHD, and various mood disorders. Due to the complexity of the data employed for those investigations, the contribution of advanced statistical models will be necessary to ensure interpretable and reproducible findings for clinical diagnosis and future therapeutic research.

## References

1. Friston, K.J., Ashburner, J.T., Kiebe, Nichols, T.E., Penny, W.D.: Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press (2007)
2. Prados, F., Boada, I., Prats-Galino, A., Martin-Fernandez, J.A., Feixas, M., Blasco, G., Puig, J., Pedraza, S.: Analysis of new diffusion tensor imaging anisotropy measures in the three-phase plot. *J. Magn. Reson. Imaging* **31**(6), 1435–1444 (2010)
3. Weber, B., Fliessbach, K., Elger, C.: Magnetic resonance imaging in epilepsy research: recent and upcoming developments. In: Schwartzkroin, P.A. (ed.) Encyclopedia of Basic Epilepsy Research, pp. 1549–1554. Academic Press, Oxford (2009)
4. Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., Johnson, H., Styner, M.: Dtprep: quality control of diffusion-weighted images. *Front. Neuroinformatics* **8**, 4 (2014)
5. Durante, D., Dunson, D.B.: Bayesian inference and testing of group differences in brain networks. *Bayesian Anal.* **13**(1), 29–58 (2018)
6. Poldrack, R., Mumford, J., Nichols, T.: Handbook of fMRI Data Analysis. Cambridge University Press (2011)
7. Handwerker, D.A., Gonzalez-Castillo, J., D'Esposito, M., Bandettini, P.A.: The continuing challenge of understanding and modeling hemodynamic variation in fMRI. *NeuroImage* **62**(2), 1017–1023 (2012)
8. Handwerker, D.A., Ollinger, J.M., D'Esposito, M.: Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage* **21**(4), 1639–1651 (2004)
9. Rangaprakash, D., Wu, G.R., Marinazzo, D., Hu, X., Deshpande, G.: Hemodynamic response function (HRF) variability confounds resting-state fMRI functional connectivity. *Magn. Reson. Med.* (2018)
10. Wu, G.R., Liao, W., Stramaglia, S., Ding, J.R., Chen, H., Marinazzo, D.: A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. *Med. Image Anal.* **17**(3), 365–374 (2013)
11. Friston, K.: Functional and effective connectivity in neuroimaging: a synthesis. *Hum. Brain Mapp.* **2**, 56–78 (1994)

12. Andersen, A., Gash, D., Avison, M.: Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework. *Magn. Reson. Imaging* **17**(6), 795–815 (1999)
13. Calhoun, V., Adali, T., Pearlson, G., Pekar, J.: A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* **14**(3), 140–151 (2001)
14. Varoquaux, G., Gramfort, A., Poline, J., Thirion, B., Zemel, R., Shawe-Taylor, J.: Brain covariance selection: better individual functional connectivity models using population prior. *Adv. Neural Inf. Process. Syst.* (2010)
15. Bowman, F., Caffo, B., Bassett, S., Kilts, C.: A Bayesian hierarchical framework for spatial modeling of fMRI data. *NeuroImage* **39**(1), 146–156 (2008)
16. Zhang, L., Guindani, M., Versace, F., Vannucci, M.: A spatio-temporal nonparametric Bayesian variable selection model of fMRI data for clustering correlated time courses. *NeuroImage* **95**, 162–175 (2014)
17. Ferguson, T.S.: A Bayesian analysis of some nonparametric problems. *Ann. Stat.*, 209–230 (1973)
18. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**(3), 186–198 (2009)
19. van den Heuvel, M.P., Sporns, O.: Network hubs in the human brain. *Trends Cogn. Sci.* **17**(12), 683–696 (2013)
20. Stam, C.J., Reijneveld, J.C.: Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomed. Phys.* **1**, 3–3 (2007)
21. Ginestet, C.E., Li, J., Balachandran, P., Rosenberg, S., Kolaczyk, E.D.: Hypothesis testing for network data in functional neuroimaging. *Ann. Appl. Stat.* **11**(2), 725–750 (2017)
22. Friston, K.J., Frith, C.D., Frackowiak, R.S.J.: Time-dependent changes in effective connectivity measured with pet. *Hum. Brain Mapp.* **1**(1), 69–79 (1993). <https://doi.org/10.1002/hbm.460010108>
23. Büchel, C., Friston, K.: Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb. Cortex* **7**(8), 768–778 (1997)
24. McIntosh, A., Gonzalez-Lima, F.: Structural equation modeling and its application to network analysis in functional brain imaging. *Hum. Brain Mapp.* **2**(1), 2–22 (1994)
25. Friston, K., Harrison, L., Penny, W.: Dynamic causal modelling. *NeuroImage* **19**(4), 1273–1302 (2003)
26. Harrison, L., Penny, W., Friston, K.: Multivariate autoregressive modeling of fMRI time series. *NeuroImage* **19**(4), 1477–1491 (2003)
27. Goebel, R., Roebroeck, A., Kim, D., Formisano, E.: Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn. Reson. Imaging* **21**(10), 1251–1261 (2003)
28. Zheng, X., Rajapakse, J.: Learning functional structure from fMR images. *NeuroImage* **31**(4), 1601–1613 (2006)
29. Yu, Z., Pluta, D., Shen, T., Chen, C., Xue, G., Ombao, H.: Statistical challenges in modeling big brain signals. *ArXiv e-prints* (2018)
30. Gorrostieta, C., Fiecas, M., Ombao, H., Burke, E., Cramer, S.: Hierarchical vector autoregressive models and their applications to multi-subject effective connectivity. *Front. Comput. Neurosci.* **7** (2013)
31. Yu, Z., Prado, R., Quinlan, E.B., Cramer, S.C., Ombao, H.: Understanding the impact of stroke on brain motor function: a hierarchical Bayesian approach. *J. Am. Stat. Assoc.* **111**(514), 549–563 (2016)
32. Friston, K.: Functional and effective connectivity: a review. *Brain Connect.* **1**(1), 13–36 (2011)
33. Bowman, F.: Brain imaging analysis. *Annu. Rev. Stat. Its Appl.* **1**, 61–85 (2014)
34. Enno, S.K., J., F.K.: Analyzing effective connectivity with functional magnetic resonance imaging. *Wiley Interdiscip. Rev. Cogn. Sci.* **1**(3), 446–459 (2010)

35. Savitz, J.B., Rauch, S.L., Drevets, W.C.: Clinical application of brain imaging for the diagnosis of mood disorders: the current state of play. *Mol. Psychiatry* **18**, 528 EP (2013)
36. Insel, T., Cuthbert, B.: Brain disorders? Precisely. *Science* **348**(6234) (2015)
37. Paulus, M.P., Stein, M.B.: Role of functional magnetic resonance imaging in drug discovery. *Neuropsychol. Rev.* **17**(2), 179–188 (2007)
38. Kaufman, J., Gelernter, J., Hudziak, J.J., Tyrka, A.R., Coplan, J.D.: The Research Domain Criteria (RDoC) project and studies of risk and resilience in maltreated children. *J. Am. Acad. Child Adolesc. Psychiatry* **54**(8), 617–625 (2015)
39. Kose, S., M., C.: The research domain criteria framework: transitioning from dimensional systems to integrating neuroscience and psychopathology. *Psychiatry Clin. Psychopharmacol.* **27**(1), 1–5 (2017)
40. Johnson, T., Liu, Z., Bartsch, A., Nichols, T.: A Bayesian non-parametric Potts model with application to pre-surgical fMRI data. *Stat. Methods Med. Res.* **22**(4), 364–381 (2013)
41. Kim, S., Smyth, P., Stern, H.: A nonparametric Bayesian approach to detecting spatial activation patterns in fMRI data. *Med. Image Comput. Comput. Assist. Interv.*, 217–224 (2006)
42. Jbabdi, S., Woolrich, M., Behrens, T.: Multiple-subjects connectivity-based parcellation using hierarchical Dirichlet process mixture models. *NeuroImage* **44**(2), 373–384 (2009)
43. Xu, L., Johnson, T., Nichols, T., Nee, D.: Modeling inter-subject variability in fMRI activation location: a Bayesian hierarchical spatial model. *Biometrics* **65**(4), 1041–1051 (2009)
44. Zhang, L., Guindani, M., Versace, F., Engelmann, J.M., Vannucci, M.: A spatiotemporal non-parametric Bayesian model of multi-subject fMRI data. *Ann. Appl. Stat.* **10**(2), 638–666 (2016)
45. Teh, Y.W., Jordan, M.I., Beal, M.J., Blei, D.M.: Hierarchical dirichlet processes. *J. Am. Stat. Assoc.* **101**(476), 1566–1581 (2006)
46. Flandin, G., Penny, W.: Bayesian fMRI data analysis with sparse spatial basis function priors. *NeuroImage* **34**(3), 1108–1125 (2007)
47. Harrison, L., Green, G.: A Bayesian spatiotemporal model for very large data sets. *NeuroImage* **50**(3), 1126–1141 (2010)
48. Penny, W., Kiebel, S., Friston, K.: Variational Bayesian inference for fMRI time series. *NeuroImage* **19**(3), 727–741 (2003)
49. Penny, W., Trujillo-Barreto, N., Friston, K.: Bayesian fMRI time series analysis with spatial priors. *NeuroImage* **24**(2), 350–362 (2005)
50. Woolrich, M., Behrens, T., Smith, S.: Constrained linear basis sets for HRF modelling using variational Bayes. *NeuroImage* **21**(4), 1748–1761 (2004b)
51. Kook, J.H., Guindani, M., Zhang, L., Vannucci, M.: NPBayes-fMRI: Non-parametric Bayesian General Linear Models for Single- and Multi-Subject fMRI Data (2017, in press)
52. Muschelli, J., Gherman, A., Fortin, J.P., Avants, B., Whitcher, B., Clayden, J.D., Caffo, B.S., Crainiceanu, C.M.: Neuroconductor: an R platform for medical imaging analysis (2018, in press)
53. Fornito, A., Zalesky, A., Pantelis, C., Bullmore, E.T.: Schizophrenia, neuroimaging and connectomics. *NeuroImage* **62**(4), 2296–2314 (2012)
54. Li, J., Wang, Z., Palmer, S., McKeown, M.: Dynamic Bayesian network modeling of fMRI: a comparison of group-analysis methods. *NeuroImage* **41**(2), 398–407 (2008)
55. Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Penna, S.D., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., Handwerker, D.A., Keilholz, S., Kiviniemi, V., Leopold, D.A., de Pasquale, F., Sporns, O., Walter, M., Chang, C.: Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage* **80**(0), 360–378 (2013). Mapping the Connectome
56. Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., Calhoun, V.D.: Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* (2012)
57. Lindquist, M.A., Xu, Y., Nebel, M.B., Caffo, B.S.: Evaluating dynamic bivariate correlations in resting-state fMRI: a comparison study and a new approach. *NeuroImage* (2014)
58. Cribben, I., Haraldsdottir, R., Atlas, L., Wager, T., Lindquist, M.: Dynamic connectivity regression: determining state-related changes in brain connectivity. *NeuroImage* **61**, 907–920 (2012)

59. Xu, Y., Lindquist, M.A.: Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fMRI data. *Front. Neurosci.* **9**(285) (2015)
60. Chiang, S., Casse, A., Guindani, M., Vannucci, M., Yeh, H.J., Haneef, Z., Stern, J.M.: Time-dependence of graph theory metrics in functional connectivity analysis. *NeuroImage* **125**, 601–615 (2015)
61. Warnick, R., Guindani, M., Erhardt, E., Allen, E., Calhoun, V., Vannucci, M.: A Bayesian approach for estimating dynamic functional network connectivity in fMRI data. *J. Am. Stat. Assoc.* **113**(521), 134–151 (2018)
62. Dobra, A., Lenkoski, A., Rodriguez, A.: Bayesian inference for general gaussian graphical models with application to multivariate lattice data. *J. Am. Stat. Assoc.* **106**(496), 1418–1433 (2011)
63. Roverato, A.: Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models. *Scand. J. Stat.* **29**(3), 391–411 (2002)
64. Baker, A., Brookes, M., Rezek, A., Smith, S., Behrens, T., Penny, J., Smith, R., Woolrich, M.: Fast transient networks in spontaneous human brain activity. *eLife* **3**(3), 1–18 (2014)
65. Balqis-Samdin, S., Ting, C.M., Ombao, H., Salleh, S.H.: A unified estimation framework for state-related changes in effective brain connectivity. *IEEE Trans. Biomed. Eng.* **64**(4), 844–858 (2017)
66. Peterson, C., Stingo, F.C., Vannucci, M.: Bayesian inference of multiple gaussian graphical models. *J. Am. Stat. Assoc.* **110**(509), 159–174 (2015)
67. Chiang, S., Guindani, M., Yeh, H.J., Dewar, S., Haneef, Z., Stern, J.M., Vannucci, M.: A hierarchical Bayesian model for the identification of pet markers associated to the prediction of surgical outcome after anterior temporal lobe resection. *Front. Neurosci.* **11**, 669 (2017)
68. Haynes, J.D., Rees, G.: Predicting the stream of consciousness from activity in human visual cortex. *Curr. Biol.* **15**(14), 1301–1307 (2005)
69. LaConte, S., Strother, S., Cherkassky, V., Anderson, J., Hu, X.: Support vector machines for temporal classification of block design fMRI data. *NeuroImage* **26**(2), 317–329 (2005)
70. Mitchell, T., Hutchinson, R., Niculescu, R., Pereira, F., Wang, X., Just, M., Newman, S.: Learning to decode cognitive states from brain images. *Mach. Learn.* **57**(1–2), 145–175 (2004)
71. Arribas, J., Calhoun, V.D., Adali, T.: Automatic Bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from fMRI data. *IEEE Trans. Biomed. Eng.* **57**(12) (2010)
72. Burge, J., Lane, T., Link, H., Qiu, S., Clark, V.P.: Discrete dynamic Bayesian network analysis of fMRI data. *Hum. Brain Mapp.* **30**, 122–137 (2009)
73. Zhang, L., Guindani, M., Vannucci, M.: Bayesian models for functional magnetic resonance imaging data analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **7**(1), 21–41 (2015)
74. Uludag, K., Roebroeck, A.: General overview on the merits of multimodal neuroimaging data fusion. *NeuroImage* **102**, 3–10 (2014)
75. Valdes-Sosa, P.A., Kotter, R., Friston, K.J.: Introduction: multimodal neuroimaging of brain connectivity. *Philos. Trans. R. Soc. B Biol. Sci.* **360**(1457), 865–867 (2005)
76. Biessmann, F., Plis, S., Meinecke, F.C., Eichele, T., Muller, K.R.: Analysis of multimodal neuroimaging data. *IEEE Rev. Biomed. Eng.* **4**, 26–58 (2011)
77. Jorge, J., van der Zwaag, W., Figueiredo, P.: EEG-fMRI integration for the study of human brain function. *NeuroImage* **102**, 24–34 (2014)
78. Kalus, S., Sämann, P., Czisch, M., Fahrmeir, L.: fMRI activation detection with EEG priors. Technical report, University of Munich (2013)
79. Chiang, S., Guindani, M., Yeh, H.J., Haneef, Z., Stern, J.M., Vannucci, M.: Bayesian vector autoregressive model for multi-subject effective connectivity inference using multi-modal neuroimaging data. *Hum. Brain Mapp.* **38**(3), 1311–1332 (2016)
80. Nathoo, F., Kong, L., Zhu, H.: A review of statistical methods in imaging genetics. Technical report, ArXiv (2018)
81. Stingo, F.C., Guindani, M., Vannucci, M., Calhoun, V.D.: An integrative Bayesian modeling approach to imaging genetics. *J. Am. Stat. Assoc.* **108**(503), 876–891 (2013)

82. Greenlaw, K., Szefer, E., Graham, J., Lesperance, M., Nathoo, F.S.: The Alzheimer's Disease Neuroimaging Initiative: A Bayesian group sparse multi-task regression model for imaging genetics. *Bioinformatics* **33**(16), 2513–2522 (2017)
83. Wang, H., Nie, F., Huang, H., Risacher, S.L., Saykin, A.J., Shen, L.: The Alzheimer's Disease Neuroimaging Initiative: Identifying disease sensitive and quantitative trait-relevant biomarkers from multidimensional heterogeneous imaging genetics data via sparse multimodal multitask learning. *Bioinformatics* **28**(12), i127–i136 (2012)
84. Chekouo, T., Stingo, F.C., Guindani, M., Do, K.A.: A Bayesian predictive model for imaging genetics with application to schizophrenia. *Ann. Appl. Stat.* **10**(3), 1547–1571 (2016)
85. Yu, C.H., Prado, R., Ombao, H., Rowe, D.: A Bayesian variable selection approach yields improved detection of brain activation from complex-valued fMRI. *J. Am. Stat. Assoc.*, 1–61 (2018)
86. Rockova, V., George, E.I.: EMVS: the EM approach to Bayesian variable selection. *J. Am. Stat. Assoc.* **109**(506), 828–846 (2014)
87. Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S.S., Muñoz, L.D., Mullinger, K.J., Tierney, T.M., Bestmann, S., Barnes, G.R., Bowtell, R., Brookes, M.J.: Moving magnetoencephalography towards real-world applications with a wearable system. *Nature* **555**, 657 EP (2018)