ANALYTICAL TECHNIQUES TOWARDS IDENTIFICATION OF HUMAN BRAIN FUNCTIONAL NETWORKS

by

PRIYA AGGARWAL

Under the Supervision of Dr. Anubha Gupta

Indraprastha Institute of Information Technology Delhi

December, 2018

ANALYTICAL TECHNIQUES TOWARDS IDENTIFICATION OF HUMAN BRAIN FUNCTIONAL NETWORKS

by

PRIYA AGGARWAL

Submitted
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

to the
Indraprastha Institute of Information Technology Delhi

December, 2018

Certificate

This is to certify that the thesis titled **Analytical Techniques towards Identification of Human Brain Functional Networks** being submitted by **Priya Aggarwal** to the Indraprastha Institute of Information Technology-Delhi, for the award of the degree of **Doctor of Philosophy**, is an original research work carried out by her under my supervision. In my opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted in part or full to any other university or institute for the award of any degree/diploma.

December, 2018

Dr. Anubha Gupta

Associate Professor,

Deptt. of Electronics & Communication Engineering,

Indraprastha Institute of Information Technology-Delhi,
New Delhi-110020, India.

Acknowledgments

I would like to take this opportunity to express my thanks to those who assisted and supported me throughout my Ph.D. journey. First and foremost, I would like to thank God for giving me strength, ability and perseverance to undertake this research study and complete it satisfactorily. Without his blessings, this work would not have been possible.

I would like to begin by expressing my sincere gratitude to my supervisor Prof. Anubha Gupta for her guidance, mentoring, and encouragement throughout my Ph.D. research. She has truly been one of the most admirable people that I have met through the course of my life. Prof. Gupta's achievements, enthusiasm, and work ethics have been a great source of inspiration throughout all my Ph.D. endeavours. I have learned immensely by working with her since past few years. Our long hours of discussions were extremely helpful to make progress in my research. I would also like to thank her for all her helps when it came to documenting results, writing papers and presenting it.

On a broader note, I want to acknowledge all the professors of IIIT-D who have inspired me directly or indirectly. I would especially like to thank my thesis examiners for their valuable comments: Prof. Saket Anand, Prof. Sumit Darak and Prof. Vivek Bohara (also my M.Tech. co-supervisor). I am extremely grateful to them for their fruitful comments throughout my doctoral research endeavor. My gratitude also extends to the Admin Staff of IIIT-D, especially Priti for being extremely helpful in the fast resolution of all admin related matters.

I am also grateful to Visvesvaraya Ph.D. scheme, Department of Electronics and Information Tech., Ministry of Comm. and IT, Govt. of India for providing financial support for this work and also extremely thankful to Microsoft research for funding my travel to one of the prestigious conference in medical imaging area i.e., MICCAI, 2015. I would also like to sincerely thank the editors and anonymous reviewers for their very helpful and kind comments to assist in improving the presentation of our papers.

Success is not just about knowledge. Without the encouragement and support of my friends, especially Palak, this work would not have been possible. I owe my deepest gratitude to them for all their love and support during the course of my Ph.D. thesis. I would also like to thank all the RAs, undergraduates and

postgraduates students with whom I worked during the course of my Ph.D. They helped me to grasp foothold on my thesis by discussing and guiding them.

Last but not the least, I would like to express my gratitude to my parents whom have always been a source of encouragement throughout my academic career. I can not thank them enough for their tremendous support and inspiration along the way.

Abstract

Human brain is a large scale complex network of different but functionally coupled brain regions. These functionally coupled brain regions together form functional brain networks (fBNs). Identification of these networks help in the diagnosis of various neuropsychiatric disorders. In the field of neuroscience, there are three larger issues to be handled using functional Magnetic Resonance Imaging (fMRI) data. Firstly, reliable functional networks are required to be identified using the spatio-temporal recordings of the brain. Secondly, the dynamic time-varying information processing in the brain results in a faster reorganization of fBNs. Thus, development of advanced methodology to detect dynamic fBNs is one of the biggest challenge in neuroscience. Thirdly, longer scan duration results in poor subjects attentivity to the intended task. This problem relates to poor fMRI data quality. This raises a question as to whether the data can be captured in smaller time such that good data quality fMRI data can be captured.

The above three issues were the main motivations of this dissertation. In this dissertation, we aim to adapt tools from signal processing, such as multivariate regression with constrained optimization technique, to identify fBNs. Proposed multivariate regression method considers weighted combination of different brain regions and captures sparse and dense fBNs simultaneously. We name the proposed method as Multivariate Vector Regression-based Connectivity (MVRC). Further, we extend the proposed method to identify fBNs at the group-level comprising of

multiple subjects using group-fused constrained optimization.

Second contribution of the proposed work is to identify overlapping fBNs that are often ignored in the literature. By overlapping, we mean that one brain region may be a part of multiple fBNs. This sounds plausible because one stimulus, say auditory, may stimulate memory and other fBNs apart from the auditory network. This indicates a need for identifying overlapping fBNs compared to the commonly identified disjoint fBNs.

Thirdly, in order to identify dynamic fBNs, we utilize state-of-the-art sliding time window approach and consider fBNs to be static within each window. Further, we propose a technique to identify dynamic overlapping fBNs and show the efficacy of proposed approach on openly available Autism dataset.

Recently, some studies have attempted to compute fBNs via estimated intrinsic stimulus at brain regions instead of commonly used activity time-series. These methods are based on the assumption that the functional connection between brain regions is due to the intrinsic stimulus. In this dissertation, we present novel intrinsic stimulus estimation method which overcomes certain limitations of existing methods and can be further utilized to extract fBNs.

In the end of this dissertation, we present work carried out on developing new method for compressed fMRI for faster data acquisition. We aim to build methods that are able to better preserve fBNs compared to the existing methods. This is essential to show that the proposed method is accurate since extraction of fBNs is one of the the crucial motive for studying fMRI data.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AAL	Automated Anatomical Labeling
ADMM	Alternating Direction Multiplier Method
AN	Auditory Network
ASD	Autism Spectrum Disorder
ASR	Adaptive Sparse Representation
BCT	Brain Connectivity Toolbox
BLN	Bilateral Limbic Network
BOLD	Blood Oxygenated Level Dependent
CCN	Cognitive Control Network
CORR	Pearson Correlation
CS	Compressed Sensing
dFC	Dynamic Functional Connectivity
DMN	Default Mode Network
EEG	Electroencephalography
EPI	Echo Planar Imaging
fBNs	Functional Brain Networks

FC	Functional Connectivity
FCP	Functional Connectomes Project
FD	Fourier Domain
fMRI	Functional Magnetic Resonance Imaging
FOV	Field Of View
FWHM	Full Width Half Maximum
GF-MVRC	Group-fused MVRC
HCP	Human Connectome Project
HRF	Heamodynamic Response Function
IFT	Inverse Fourier Transform
JI	Jaccard Index
LASSO	Least Absolute Shrinkage and Selection Operator
LN	Language Network
MC	Matrix Completion
MEG	Magnetoencephalography
MNI	Montreal Neurological Institute
MPRAGE	Magnetization Prepared Rapid Gradient Echo
MRI	Magnetic Resonance Imaging
MVRC	Multivariate Vector Regression-based Connectivity
NMF	Non-negative Matrix Factorization
NMSE	Normalized Mean Square Error

NP	Non-deterministic Polynomial
PCORR	Partial Correlation
RPCA	Robust Principal Component Analysis
rs-fMRI	Resting-State fMRI
SCN	Subcortical Network
sFC	Static Functional Connectivity
SG	Subgroup
SMN	Somato-Motor Network
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
SVT	Singular Value Thresholding
TDC	Typically Developing Controls
TE	Echo Time
TR	Repetition Time
VN	Visual Network

Chapter 1

Introduction

1.1 Motivation

Brain is the most complex part in humans, which comprises of about 10^{11} neurons connected by around 10^{14} synapses [2, 3, 4, 5, 6]. In order to study brain, this complex architecture is generally being viewed from the perspective of networks of the brain. Network, in general, is a set of nodes that are connected together. While analyzing human brain networks, one is interested in analyzing data collected from different brain areas or regions to ascertain whether they are structurally and/or functionally connected [7, 8]. This understanding of human brain networks can aid in a better understanding of various neurological disorders such as autism, schizophrenia, etc., besides facilitating better medical and paramedical care to subjects. Therefore, extracting brain networks has become one of the enduring goals of neuroscience leading to an exponential growth of research in this area in the past one decade with the most comprehensive investigation than ever before. [9, 10, 11, 12, 13, 3, 5, 10, 6].

Complete network map of brain is also termed as Connectome [14]. Here, brain regions and their interconnections are represented as nodes and edges, respectively, also called the elemental building blocks. Brain networks, be it structural or functional, are identified in terms of connected nodes. Clear definition of nodes and edges is of utmost importance for modeling any network's organization [15, 6]. As of today, human brain is believed to be organized at multiple scales comprising of a) individual neurons ($< 1\mu m$), b) to population of neurons ($< 1cm$) and, c) to

large-scale regions such as lobes or hemisphere ($< 10cm$) [16]. There is no single definition of nodes for the case of brain networks due to it's multiscale organization. The definition of nodes at each scale may provide complementary and important information about the brain networks.

In terms of multiscale brain organization, networks can be organized into three categories: microscopic, mesoscopic and macroscopic [6]. Microscopic scale refers to networks constructed at the level of individual neurons. Here, a node can be defined by a single neuron. On the other end, macroscopic scale refers to networks constructed at the level of large-scale population of neurons, such as those characterized with non-invasive neuroimaging modalities including electroencephalography (EEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI). This scale allows us to study brain networks over the entire brain (particularly with fMRI) at the spatial resolution of millimeters or centimeters. Lastly, mesoscopic scale bridges microscopic and macroscopic scales. Unfortunately, so far, there is no single neuroimaging modality that can help in studying brain networks across all scales.

In terms of defining an edge/link of a network, Connectome can be analyzed via structural connectivity (to study structural brain networks) and functional connectivity (FC) (to study functional brain networks (fBNs)). Structural connectivity between nodes stands for anatomical connections between nodes such as synapses at the microscopic scale (measured with electron microscopy) and large-scale fiber bundles at the meso- and macroscopic scales (measured with tract imaging and diffusion weighted imaging (DWI), respectively) [17], whereas FC stands for statistical dependence such as Pearson correlation between distinct brain nodes' physiological recordings. In this dissertation, we focus on FC at the macroscopic scale. This scale is most commonly considered to study fBNs in humans (or in living organisms) due to the ease of capturing data via different noninvasive neuroimaging modalities such as EEG, MEG, and fMRI. The advantages of these modalities include clinical safety, tolerability, and flexibility in allowing studies across wide range of human neurological disorders.

Since its development in the early 1990s, fMRI has played an essential role in visualizing human fBNs [18, 19, 13, 20, 21]. There is a great interest in estimating fBNs from fMRI data due to its flexibility to allow higher spatial coverage of the

brain compared to other popular neuroimaging modalities such as EEG and MEG. Therefore, in this dissertation, we focus on fMRI to study fBN organization. A typical fMRI modality has a spatial resolution of about $1mm^3$, represented as a single volumetric unit, called voxel. A single voxel may contain about 20000-30000 neurons and billions of synapses [22]. Generally, each voxel represents one node, although a group of voxels, forming one functional brain region, is generally considered as one node because it allows easy handling of a large number of voxels in the brain (around 0.5 million). Further details regarding fMRI and nodes definition is provided in Chapter 2.

In general, functionally connected nodes have stronger edge connection between them. Pearson correlation, signifying similarity between physiological recordings from distinct nodes, is used as a major tool in FC analysis [23]. High value of correlation indicates functional connection between nodes. However, Pearson correlation has certain disadvantages. For example, there may be a case when node A is functionally correlated to node B not because nodes A and B are correlated but because both nodes A and B are correlated to node C. This may provide spurious high correlation between the three nodes that is also known as triangular fBNs in the literature.

To overcome the above stated limitation of correlation, partial correlation is used [1, 24, 25, 26, 27]. It is used to combat spurious correlation values between two nodes by removing influence of all other nodes. Two nodes are said to be connected if partial correlation value between them is non-zero. However, this conclusion is valid only if the data is Gaussian [28]. In case, data does not satisfy this assumption, partial correlation may also lead to erroneous inferences.

In fact, different methods are being used to define edges in the fMRI literature [1]. This suggests that there is no "one method" to identify edges, although many studies have indicated that Pearson correlation based approaches are successful in extracting FC from fMRI data [29, 1]. As a result, the study of fBNs, since its initial days, is relying largely on Pearson correlation, although there are many unanswered questions about the choice of this metric.

Typically, edges are determined by computing edge values for the entire fMRI session duration. Such computation of an edge is known as the static functional connectivity (sFC). However, it has been reported that human brain may not be in

a single network state for a long duration. This implies that fBNs change with time and are, perhaps, non-stationary even during one neuroimaging session. Therefore, it is generally agreed that sFC alone cannot fully account for all aspects of fBNs.

Dynamic functional connectivity (dFC) based analysis is emerging as a promising technique complementary to sFC [30]. It is an extension of sFC stated above, where first the entire session is divided into multiple time windows. FC is computed within every fixed time window, assuming network to be static within this window duration. Collection of sFCs across multiple time windows, called as dFC, is further analyzed to extract dynamic fBNs. Again, recent methods on computing dFC are largely based on computing correlation or partial correlation within each window duration.

Variability in FC over time (i.e., dFC) was earlier explored in EEG recordings [31, 32]. However, lower temporal resolution of fMRI, ranging between 0.5 and 1.5 Hz (depending on the scan acquisition protocol), limits the ability to unravel the dynamic configuration of fBNs [33, 34]. It imposes the constraint of capturing fMRI for larger time duration in one session that makes the subjects uncomfortable inside the scanner or imposes the constraint of reducing the inter-scan time (time-to-repeat, TR) that increases noise in the captured fMRI data [35]. Thus, both these constraints lead to noisy data either due to motion and other physiological artifacts or due to low SNR associated with low TR. This leads to two research questions: 1) to devise methodologies to capture data in shortest possible time, say, using parallel imaging or compressed sensing (CS) based reconstruction in fMRI and 2) to devise robust methods to extract dynamic fBNs using the available fMRI data.

1.2 Research Objectives

Many insights into the human fBNs are made possible by methodological progress. Analysis methods range from simple measures that consider just two nodes at a time (e.g., Pearson correlation) to sophisticated approaches that consider all nodes simultaneously (e.g., Bayes net method), although Bayes net method is not considered that appropriate for fMRI data due to its low quality [1]. Thus, Pearson correlation still remains the most often used method. Newer techniques are needed to

overcome limitations of the existing methods. In this dissertation, we aim to adapt tools from signal processing, such as multivariate regression with constrained optimization technique, to compute FC. The proposed multivariate regression considers weighted combination of all nodes simultaneously, while computing FC between nodes and, captures sparse and dense fBNs simultaneously. We also utilize this proposed method to classify schizophrenic patients from human controls. Further, we extend the proposed method to identify fBNs at the group-level comprising of multiple subjects using group-fused constrained optimization problem.

Second contribution of the proposed work is identifying overlapping fBNs that are often not studied in the FC literature [36]. By overlapping, we mean that one node may be a part of multiple fBNs. This sounds plausible because one stimulus, say auditory, may stimulate memory and other fBNs apart from the auditory network. This indicates a need of identifying overlapping fBNs compared to the commonly identified disjoint fBNs [36, 10].

In order to identify dFC by considering state-of-the-art time window approach, we utilize the proposed multivariate regression method instead of CORR to compute dFC within each window. Further, we propose a technique to identify dynamic overlapping fBNs and show the efficacy of proposed approach on an openly available Autism dataset.

Recently, some studies have attempted to compute FC via estimated intrinsic stimulus at nodes instead of commonly used activity time-series of brain nodes. The idea behind these methods is that functional connection between nodes is mainly due to intrinsic stimulus at different nodes. In this dissertation, we present novel intrinsic stimulus estimation method that overcomes certain limitations of existing methods that can be further utilized to build fBNs.

In the end, we present work carried out on developing new method for CS based reconstruction of fMRI data for faster data acquisition. We aim to build methods that are able to better preserve fBNs compared to the existing methods. This is essential to show that the proposed method is accurate and is able to preserve fBNs, which is one of the most important objectives for studying fMRI data.

1.3 Publications

The work presented in this dissertation resulted in some journals and conferences publications that are listed below.

Journal Papers

- Priya Aggarwal and Anubha Gupta, "Low rank and sparsity constrained method for identifying overlapping functional brain networks," *PloS one*, vol. 13, no. 11, pp. e0208068, 2018.

- Priya Aggarwal, Anubha Gupta, and Ajay Garg, "Multivariate brain network graph identification in functional MRI," *Medical Image Analysis*, vol. 42, pp. 228–240, 2017.

- Priya Aggarwal and Anubha Gupta, "Double temporal sparsity based accelerated reconstruction of compressively sensed resting-state fMRI," *Computers in Biology and Medicine*, vol. 91, pp. 255–266, 2017.

- Priya Aggarwal, Parth Shrivastava, Tanay Kabra, and Anubha Gupta, "Opt-shrink LR+S: accelerated fMRI reconstruction using non-convex optimal singular value shrinkage," *Brain Informatics*, vol. 4, no. 1, pp. 65–83, 2017.

- Priya Aggarwal and Anubha Gupta, "Accelerated fMRI reconstruction using matrix completion with sparse recovery via split bregman," *Neurocomputing*, vol. 216, pp. 319–330, 2016.

- Priya Aggarwal and Anubha Gupta, "Group-fused multivariate regression modeling for group-level brain networks," Submitted, *Neurocomputing*, Feb. 2018.

- Priya Aggarwal and Anubha Gupta, "Multivariate graph learning for detecting aberrant connectivity of dynamic brain networks in autism," Submitted, *Medical Image Analysis*, Aug. 2018.

International Conferences

- Priya Aggarwal, Anubha Gupta, and Ajay Garg, "Joint estimation of activity signal and hrf in fMRI using fused lasso." *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 829–833, 2015.

- Priya Aggarwal, Anubha Gupta, and Ajay Garg, "Joint estimation of hemodynamic response function and voxel activation in functional MRI data," *Medical Image Computing and Computer Assisted Intervention Society (MICCAI)*, pp. 142–149, 2015.

1.4 Outline of Thesis

The thesis is structured as follows:

In an introductory Chapter-2, we present an overview of how fBNs can be studied using the fMRI data. This chapter serves as a background for the published or submitted articles, which are reproduced in further chapters. This chapter also presents overview of preliminaries required for this dissertation.

In Chapter-3, we propose a multivariate regression-based connectivity method for FC analysis. The proposed method considers all nodes simultaneously while computing FC between nodes. This method also incorporates optimization constraints to identify sparse and dense fBNs simultaneously. We show that FC computed using the proposed method comprises collection of fBNs. Nodes within each fBN are highly connected and share same FC property. Results have been tabulated on healthy controls using fMRI dataset.

In Chapter-4, novel method of learning FC is proposed by constraining its rank and the sum of non-zero coefficients. In addition, large-scale overlapping functional brain networks are identified on resting-state fMRI data by employing non-negative matrix factorization.

In Chapter-5, we extend the proposed multivariate regression method for identifying fBNs at the group-level of multiple subjects. Proposed work first computes group-fused FC across each subject and then identify group-level fBNs. Next, we introduce weighted-subject framework via iterative spectral clustering that includes weights of each subject for extracting group-level fBNs. Thus, this approach

not only identifies group-level fBNs, but also utilizes each subject's data appropriately for explaining these networks. Results on fMRI dataset are used to validate the proposed framework.

In Chapter-6, we extend the proposed multivariate regression method for dFC. The proposed work can be used to infer altered fBNs in brain disorders at the group-level. As an illustration, we apply the proposed method to the publicly available fMRI dataset of autism disorder. We observe that the dynamic fBNs of the autism group show alterations compared to those obtained with healthy controls. Thus, the proposed method may prove to be a useful tool for studying alterations in fBNs.

In Chapter-7, we present proposed method to identify intrinsic stimulus across nodes using dual optimization iterative framework. The proposed method has been evaluated on fMRI dataset.

In Chapter-8, we describe our proposed CS fMRI methods for faster data acquisition. We apply both matrix completion and sparse recovery based optimization constraints to improve the reconstruction accuracy. We test the efficacy of proposed methods in terms of fBN reproducibility compared to other existing methods.

In Chapter-9, we present conclusions of all the work done and present some directions for future work.

Chapter 2

Probing the Human Brain Connectome with fMRI

There has been a great interest to study human fBNs using fMRI data. This chapter provides an overview of fMRI (a neuroimaging modality for measuring hemodynamic changes after enhanced neural activity) and its uses for probing the human fBNs [18, 37, 38, 39].

2.1 fMRI technique and preprocessing

fMRI is one of the most widely used noninvasive neuroimaging modality to study fBNs, owing to its high spatial resolution compared to other modalities [20, 29]. Since its introduction in the early 1990s, it has made a real impact on neuroscience research. It has drawn considerable attention of researchers worldwide in the past few years.

fMRI imaging consists of three-dimensional (3D) brain image captured over time [40]. In other words, a set of 3D images (or slices) consisting of multiple voxels (3D volumetric pixel) is acquired. It is based on the magnetic properties of blood and used to measure dynamic changes in blood flow, known as hemodynamics, after enhanced neuronal activity. When neuronal activity increases, active neurons demand more oxygen to metabolize glucose (primary energetic substrate for neuronal signaling). To meet this end, oxygenated blood flows to the activated brain area resulting in decreased deoxygenated blood in that area due to more

oxygen supply than in need. Oxygenated and deoxygenated blood behave differently when placed in large magnetic field, such 1.5 - 7 Tesla fields generated by scanners. Oxygenated blood is diamagnetic (weakly repelled by the field) whereas deoxygenated blood is paramagnetic (weakly attracted to the field). Thus, magnetic properties of that area changes due to influx of oxygenated blood caused by neuronal activity. This change is measured by the scanner and used as an indirect hemodynamic marker of local neuronal activity. Measured signal is known as blood oxygenated level dependent (BOLD) signal that is an indirect measure of neuronal activity [40]. Often, the collected BOLD signal is said to be the outcome of a system that is excited by an intrinsic stimulus to that voxel. The system is characterized via its impulse response called as the hemodynamic response function (HRF) and BOLD signal is known to be the convolution of an intrinsic stimulus with HRF.

fMRI signal can be captured either in task paradigm (aka task fMRI) or in resting-state (aka resting-state fMRI (rs-fMRI)). Task fMRI is captured by triggering subjects with external task paradigm and is used to measure relative changes from baseline in the BOLD signal during the performance of a task. In recent years, there has been an increase in interest in the rs-fMRI. Here, individuals lie quietly in the scanner with their eyes open or closed, without performing any specific task. rs-fMRI is being widely used to characterize fBNs by investigating synchronous activations between spatially distinct brain nodes and therefore, sometimes fBNs is also called as resting-state networks (RSNs).

Preprocessing

It is common for fMRI data to first undergo preprocessing to remove contributions from non-neuronal physiological sources of noise, such as head motion [41, 42, 43, 44, 45]. Poor quality of BOLD leads to lower signal- to-noise ratio (SNR). Hence, fMRI data is first pre-processed before FC analysis.

Pre-processing starts with the rejection of first few brain volumes so as to allow the magnetization to reach the steady state. Next step is slice time correction followed by motion correction. Slice time correction ensures voxel-to-voxel correspondence among brain slices across time. Motion correction step helps to

remove head motion above 2 mm or 2°. Next, images are spatially normalized onto Montreal Neurological Institute (MNI) space so that results across multiple subjects can be compared and spatial concordance with coordinate systems can be achieved [27]. Further, data is smoothed with a Gaussian kernel to improve SNR. Smoothing is specified with full width half maximum (FWHM) parameter. Finally, regression is done to remove nuisance variables (such as head motion parameters, average signal from ventricles, and white matter), which include high proportions of noise related to cardiac and respiratory signals into BOLD time series [46, 47]. Followed by band pass filtering generally in the range of 0.01-0.1 Hz to remove signal from non-neuronal causes and to improve the SNR [47]. The signal recorded at these low frequencies is generally being observed to infer organization of fBNs. We preprocessed fMRI data using SPM12 (Statistical Parametric Mapping) [1].

2.2 Mapping large-scale networks with resting-state fMRI

In this section, we present an overview of how fBNs is studied with fMRI data. The purpose of this section is to provide review of pipeline used for FC analysis.

2.2.1 Defining nodes

First, we begin with the definition of nodes. A simplest way to define nodes is to treat each measurement point as an individual node. For example, individual voxels in fMRI correspond to distinct nodes [48, 49]. This approach offers flexibility of analyzing at the native resolution of the imaging technique. However, the boundary of functionally synchronous neurons population may extend beyond the boundary of voxel. This problem is more daunting in fMRI due to partial volume effects, causing lack of independence between measurements from spatially adjacent neighborhood voxels. In this scenario, the connectivity between adjacent neighboring voxels may be spuriously inflated due to measurement process. Another way of defining nodes is by utilizing predefined MNI registered parcellation atlas rather than relying on measurement points. These atlas parcellate human

[1] `http://www.fil.ion.ucl.ac.uk/spm/software/spm12/`

brain into number of brain regions, comprising of number of similar voxels in terms of functionality. Defining nodes by atlas rather than voxels also helps to get rid of large number of voxels in the brain.

In this thesis, we utilize popular Automated anatomical labeling (AAL) atlas in fMRI to parcellate human brain into 90 cerebrum regions [50, 51]. In other words, it has a total of 90 nodes as shown in Table 2.1. 45 nodes lie in the left brain hemisphere and other 45 nodes lie in the right brain hemisphere. Table 2.1 lists these nodes where odd numbers represent left hemisphere nodes and even numbers represent right hemisphere nodes.

Once nodes are obtained, nodes' representative time-series are computed for finding FC across them. In general, nodes' representative time-series are obtained by averaging BOLD signals of all the voxels within a node or region [52, 53]. Next, these nodes' representative time-series are used to estimate edges between nodes as described in the next section of this chapter.

2.2.2 Defining edges

In general, Pearson correlation between nodes' time-series is the most commonly used measure in FC analysis. It is usually computed using the amplitude fluctuations of the time-series themselves. High value of correlation indicates functional connection between nodes, whereas zero value of correlation indicates no connection. This measure is both simple and versatile, and has been applied to plenty of studies. First study which attributed the FC using rs-fMRI was made in [18], where they correlated the time-series of a seed node (left somatosensory cortex) to all other nodes and observed that the other homologous nodes showed high correlations with seed node. Since then, the number of studies using correlation in rs-fMRI to identify fBNs has rapidly increased [54, 55, 56, 57, 58].

In spite of huge popularity of Pearson correlation, this measure has certain disadvantages. First, correlation coefficient is sensitive to a third node's indirect effect. Hence, may provide spurious correlation coefficients between the three nodes, also known as triangular fBNs. Second, BOLD signal contamination from other shared inputs, such as pyshiological noisy fluctuations might also cause signals to be correlated though there is no change in signal of interest [59, 60, 61].

Table 2.1: **List of indices and regions in the AAL atlas**

Index	**Region**	**Index**	**Region**
(1,2)	Precental gyrus	(47,48)	Lingual gyrus
(3,4)	Superior frontal gyrus, dorsolateral	(49,50)	Superior occipital gyrus
(5,6)	Superior frontal gyrus, orbital part	(51,52)	Middle occipital gyrus
(7,8)	Middle frontal gyrus	(53,54)	Inferior occipital gyrus
(9,10)	Middle frontal gyrus, orbital part	(55,56)	Fusiform gyrus
(11,12)	Inferior frontal gyrus, opercular part	(57,58)	Postcentral gyrus
(13,14)	Inferior frontal gyrus, triangular part	(59,60)	Superior parietal gyrus
(15,16)	Inferior frontal gyrus, orbital part	(61,62)	Inferior parietal gyrus
(17,18)	Rolandic operculum	(63,64)	Supramarginal gyrus
(19,20)	Supplementary motor area	(65,66)	Angular gyrus
(21,22)	Olfactory cortex	(67,68)	Precuneus
(23,24)	Superior frontal gyrus, medial	(69,70)	Paracentral lobule
(25,26)	Superior frontal gyrus, medial orbital	(71,72)	Caudate nucleus
(27,28)	Gyrus rectus	(73,74)	Putamen
(29,30)	Insula	(75,76)	Pallidum
(31,32)	Anterior cingulate gyrus	(77,78)	Thalamus
(33,34)	Median cingulate gyrus	(79,80)	Heschl's gyrus
(35,36)	Posterior cingulate gyrus	(81,82)	Superior temporal gyrus
(37,38)	Hippocampus	(83,84)	Superior temporal pole
(39,40)	Parahippocampal gyrus	(85,86)	Middle temporal gyrus
(41,42)	Amygdala	(87,88)	Middle temporal pole
(43,44)	Calcarine cortex	(89,90)	Inferior temporal gyrus
(45,46)	Cuneus		

The odd and even indices represent left- and right- brain hemisphere regions, respectively.
Central = (1-2, 57-58); Frontal = (3-28, 69-70); Limbic = (31-42); Occipital = (43-56); Parietal = (59-68); Subcortical = (71-76); Temporal = (77-88).
Visual = (43-56); Auditory = (13-16, 79-82); Bilateral Limbic = (21, 22, 27, 28, 37-42, 83, 84, 87, 88); Default Mode Network (DMN) = (5, 6, 9, 10, 23-26, 35, 36, 65-68, 77, 78); Motor = (1, 2, 19, 20, 69, 70); Subcortical = (73-76); Somatosensory = (57-64); Memory = (85, 86, 89, 90).

Hence, Pearson correlation may provide spurious connection values [61]. In this thesis, we propose new method of FC computation as presented in next chapter.

2.3 Connectivity Matrices

A central goal of Connectome is to comprehensively map connections between anatomically distributed nodes' time-series. Number of possible connections for a network of N nodes would be of the order N^2. Owing to this, it is common to represent FC between every pair of nodes as a two-dimensional square matrix (known as connectivity matrix) for succinct and meaningful graph based representation [17]. It offers a compact description of the pairwise connectivity between all nodes of a network. Connectivity matrix is sometimes also known as the adjacency matrix.

In connectivity matrix, each row and the corresponding column corresponds to a unique node. The matrix element positioned at the intersection of the i^{th} row and j^{th} column encodes information about the connectivity value between the nodes i and j. Thus, each matrix element represents an edge. For the case of brain network comprising of N nodes, connectivity matrix is constructed with size $N \times N$, where each row/column of this matrix represents nodes.

Diagonal elements of connectivity matrix interpreted as connectivity of each node with itself are seldom used in neuroscience. These are commonly set to a common value for all the nodes, mostly zeros. For the case of undirected network, lower (all values beneath matrix diagonal) and upper (all values above matrix diagonal) triangular elements in this matrix are equal.

2.4 Extracting functional brain networks

The emphasis of FC studies has been on the characterization of fBN structure in the brain, i.e., division of nodes into groups such that connections are dense within the groups, but sparser between groups. Each group here represents one fBN. The ability to detect such groups could be of significant practical importance. These groups are nothing but known as communities. For instance, groups within the worldwide web might correspond to sets of web pages on related topics; groups

within social networks might correspond to social units or communities; and groups in brain networks correspond to functionally coupled nodes of the brain.

Past work on methods for discovering groups (or communities) for brain networks have reported that connectivity matrix is commonly comprised of a number of communities [13, 62]. The development of methods to perform community detection on connectivity matrix has been of great interest among researchers. Various methods have been proposed to detect communities from connectivity matrix but the most popular and widely used method for community detection for brain networks is modularity maximization [63]. As the name suggests, this method maximizes modularity, a quality measure of dividing the connectivity matrix into communities. The modularity value Q is defined as the difference between the number of edges and the expected number of edges, after randomization of edges, within the community. Consider a connectivity matrix $\mathbf{A}$ of N nodes with each element as A_{ij}, modularity value Q is defined as [62]:

$$Q = \frac{1}{2m} \sum_{i,\ j} [A_{ij} - \frac{k_i k_j}{2m}] \delta(g_i, g_j), \tag{2.1}$$

where A_{ij} denotes ij^{th} entry of the connectivity matrix $\mathbf{A}$, m is the total number of edges in the network, k_i and k_j are the degree of nodes i and j, respectively, δ is the Kronecker delta and g_i is the number of the community to which node i is assigned. Thus, the complete vector $\mathbf{g}$ of community assignment specifies the division of the connectivity matrix. Modularity is optimized over possible divisions of connectivity matrix to find the one that provides the highest score. Since the number of possible divisions of a network is exponentially large, many algorithms have been tried previously [64]. Among all algorithms, Newman algorithm is the most popular algorithm for community detection in brain networks [65, 63]. This method is based on spectral partitioning. For further detail regarding this algorithm, please refer to [65, 63].

The widely used modularity method yields disjoint communities, i.e., one brain node is a part of only one community, although recent studies in fMRI [36] demonstrate that one brain node may participate in multiple communities. Although research on uncovering disjoint community structure of fBNs has attracted attention in recent years, overlapping networks are still largely unexplored. Thus, in

this thesis, we identify both disjoint and overlapping communities of nodes. Further details of method for detecting overlapping communities are provided in the subsequent chapters.

Chapter 3

Multivariate Brain Network Graph Identification in Functional MRI

Functional Connectivity (FC) is a widely used measure to quantify relationship between pairs of brain regions [1, 18]. It is utilized to identify functional brain networks (fBNs) and is useful for understanding intrinsic functional organization of human brain [3, 21]. Therefore, accurate construction of FC is one of the most essential tasks to understand the functioning of complex human brain.

Various analytical methods have been proposed for FC modeling from fMRI data such as Pearson correlation (CORR) [23, 1] and the partial correlation (PCORR) [66, 67] as discussed in previous chapters. CORR is the most-widely used method for characterizing FC between different brain regions [23, 1]. Despite its popularity, this method is limited because it reveals the pair-wise relationship between two regions without accounting for the influences of other brain regions [68]. Similarly, PCORR method, i.e., the maximum likelihood estimation of the inverse covariance matrix, only captures the pairwise information and therefore, does not fully reflect the interactions among multiple brain regions. To solve this problem, we proposed a method of estimating FC rely on linear relationships between all brain regions and delineate functional relationships by representing one region's representative time-series as a linear combination of other regions' time-series. This chapter introduces our proposed multivariate regression framework to identify FC between nodes' time-series. The section 3.1 describes existing method in the direction of multivariate regression framework followed by problem formulation of the proposed

multivariate framework in the subsequent sections.

3.1 Literature Review

Consider an fMRI data matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N] \in \mathbb{R}^{T \times N}$, where T denotes the total number of time points (or the number of brain volumes) and N denotes the number of nodes. Each column of $\mathbf{X}$, labeled as $\mathbf{x}_i \in \mathbb{R}^T$, represents the i^{th} node's representative time-series that is normalized to unit norm. Hence, each column of $\mathbf{X}$ signifies unit-normalized time-series of one brain node. The aim of FC is to generate connectivity matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, where each element $\mathbf{A}_{ij}$ signifies strength of connection between nodes i and j. In other words, given matrix $\mathbf{X}$, we require to compute connectivity matrix of dimension $N \times N$.

Multivariate method of fBN identification is an appealing but unexplored method. Recently in [69], a multivariate adaptive sparse representation (ASR) method is proposed that considers linear relationship between one node's representative time-series with all other nodes' representative time-series. Association of node i with all other nodes can be computed using ASR as [69]:

$$\min_{\mathbf{w}_i} \frac{1}{2} \left\| \mathbf{x}_i - \mathbf{X}_i \mathbf{w}_i \right\|_2^2 + \lambda \left\| \mathbf{X}_i Diag(\mathbf{w}_i) \right\|_*, \tag{3.1}$$

where $\mathbf{w}_i \in \mathbb{R}^{N-1}$ represents edges vector and $\lambda > 0$ is a regularization parameter. For each column vector $\mathbf{x}_i$ in $\mathbf{X}$, corresponding dictionary for sparse representation is $\mathbf{X}_i = [\mathbf{x}_1, ..., \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, ..., \mathbf{x}_N] \in \mathbb{R}^{T \times N-1}$. This dictionary consists of time-series of all nodes excluding i^{th} node's time-series $\mathbf{x}_i$. Stacking all $\mathbf{w}_i$ coding coefficient vectors after padding zero at the i^{th} position, results in a coefficient matrix $\tilde{\mathbf{W}} = [\mathbf{w}_1, ..., \mathbf{w}_N] \in \mathbb{R}^{N \times N}$. Next, a symmetric connectivity matrix $\mathbf{A}$ is obtained by $\mathbf{A} = (\left|\tilde{\mathbf{W}}\right| + \left|\tilde{\mathbf{W}}\right|^T)/2$.

Bivariate methods such as correlation calculates pairwise connection between two nodes without considering influence of other nodes. Owing to this, spurious association values may arise. On the other hand, ASR method helps in the consideration of all nodes while calculating connectivity with one node. Thus, ASR method is shown to perform well in comparison to correlation and partial correlation. ASR employs trace LASSO (least absolute shrinkage and selection operator)

regularization on connectivity matrix coefficients. Trace LASSO regularization combines both l^1 and l^2 norm. This allows both sparsity and grouping among spatially adjacent brain nodes. However, this method iterates over each node and hence, is computationally intensive.

3.2 Research Contributions

Motivated by [69], we propose a Multivariate Vector Regression-based Connectivity (MVRC) method for FC analysis. The proposed MVRC method regresses time-series of all nodes to those of other nodes simultaneously and estimates connectivity matrix in one run. This makes MVRC method computationally time efficient compared to the existing ASR method.

Secondly, in order to incorporate both sparsity and grouping effect on connectivity matrix coefficients as in the ASR method, we use elastic-net regularization [70] on the connectivity matrix. Elastic-net regularization is being extensively used by researchers where ever there is a grouping effect on correlated data. It combines both l^1 and l^2 norm of connectivity matrix via two regularization terms. Hence, two regularization parameters are required to be defined. Later, we show that the tuning of two regularization parameters is not a bottleneck in MVRC method as we could decide these parameters with great ease.

Finally, we show that connectivity matrix computed using the proposed method comprises collections of communities where each community is comprised of a number of functionally connected nodes. Nodes within each community are highly connected compared to other communities. We show that these communities actually represent fBNs.

3.3 Proposed Formulation

Comparing l^1- and l^2-norms, we know that l^1-norm deals with sparsity among different brain networks, while l^2-norm focuses mostly on the correlated data and ignores sparse connection between two brain networks. To accommodate these aspects, ASR method balances between l^1-norm and l^2-norm via trace LASSO regularization [69]. However, ASR method is time consuming because it solves for

each node iteratively. As a result, a higher number of nodes (N) leads to higher time complexity.

Instead of considering one node at a time, the proposed MVRC considers all nodes simultaneously. The proposed MVRC method regresses time-series of all nodes onto time-series of other nodes simultaneously and estimates the connectivity matrix in one step. This makes MVRC method computationally much more time efficient compared to the existing ASR method.

We formulate the MVRC problem as:

$$\underset{\tilde{\mathbf{W}}}{min}\ \Omega(\tilde{\mathbf{W}}) \text{ s.t. } \mathbf{X} = \mathbf{X}\tilde{\mathbf{W}} + \epsilon, \quad Diag(\tilde{\mathbf{W}}) = 0, \tag{3.2}$$

where $\tilde{\mathbf{W}}$ is the coefficient matrix, $\Omega(\tilde{\mathbf{W}})$ denotes priors on the coefficient matrix and ε is a noise. $Diag(\tilde{\mathbf{W}}) = 0$ in (3.2) ensures that the connection of a node with itself is not considered.

We reformulate above equation as below:

$$\underset{\tilde{\mathbf{W}}}{min}\ \frac{1}{2}\left\|\mathbf{X} - \mathbf{X}\tilde{\mathbf{W}}\right\|_F^2 + \Omega(\tilde{\mathbf{W}}) \text{ s.t. } Diag(\tilde{\mathbf{W}}) = 0, \tag{3.3}$$

The regularization framework in (3.3) ensures convergence to a non-trivial solution. In order to employ both l^1- and l^2-norms in this new formulation, we again re-write (3.3) as:

$$\underset{\tilde{\mathbf{W}}}{min}\ \frac{1}{2}\left\|\mathbf{X} - \mathbf{X}\tilde{\mathbf{W}}\right\|_F^2 + \mu_1\left\|\tilde{\mathbf{W}}\right\|_1 + \mu_2\left\|\tilde{\mathbf{W}}\right\|_F^2 \tag{3.4}$$

where μ_1 and μ_2 are the regularization parameters. This prior on coefficient matrix is known as the elastic-net penalty [70]. This penalty combines both LASSO (l^1-norm) and ridge (l^2-norm) operators and hence, has the ability to identify brain networks while retaining the key property of ASR. The proposed MVRC method has all these advantages in addition to being a much faster method. When μ_1 is zero, (3.4) reduces to simple ridge regression, i.e., the regression solution with only l^2-norm constraint. When μ_2 is zero, (3.4) reduces to lasso regression. During iterative learning for the solution in (3.4), $Diag(\tilde{\mathbf{W}})$ is kept to zero.

Implementation

We require to implement (3.4) that is a convex optimization problem. We utilize alternating direction multiplier method (ADMM) [71] that splits this problem into multiple sub-problems that are easier to solve. ADMM is suitable for constrained optimization problems and is being used extensively previously [72, 73, 74, 68]. This technique facilitates solution by decomposing the original objective function into multiple objective functions that are easy to solve. Next, we explain ADMM based methodology to solve this problem.

We introduce auxiliary variable $\mathbf{Z} \in \mathbb{R}^{N \times N}$ for the l^1-norm term in (3.4) [71]. It allows solving (3.4) via splitting it into sub-problems. The new objective function is:

$$\underset{\tilde{\mathbf{W}}}{min}\ \frac{1}{2}\left\|\mathbf{X}-\mathbf{X}\tilde{\mathbf{W}}\right\|_F^2+\mu_1\left\|\mathbf{Z}\right\|_1+\mu_2\left\|\tilde{\mathbf{W}}\right\|_F^2\ \text{s.t.}\ \ \mathbf{Z}=\tilde{\mathbf{W}}. \tag{3.5}$$

Following [71], the augmented Lagrangian function of problem (3.5) can be written as:

$$\begin{aligned} L(\tilde{\mathbf{W}},\mathbf{Z},\mathbf{Y},\beta)=\frac{1}{2}\left\|\mathbf{X}-\mathbf{X}\tilde{\mathbf{W}}\right\|_F^2+\mu_1\left\|\mathbf{Z}\right\|_1+\mu_2\left\|\tilde{\mathbf{W}}\right\|_F^2 \\ +\frac{\beta}{2}\left\|\mathbf{Z}-\tilde{\mathbf{W}}+\frac{\mathbf{Y}}{\beta}\right\|_F^2, \end{aligned} \tag{3.6}$$

where $\beta > 0$ is a penalty parameter and $\mathbf{Y} \in \mathbb{R}^{N \times N}$ is a Lagrangian multiplier. The above equation consists of two variables $\tilde{\mathbf{W}}$ and $\mathbf{Z}$. ADMM splits the above problem into two subproblems. Each subproblem may be treated as minimization over one variable while fixing other variables. The iterations of ADMM are described in *Algorithm* 1.

It should be noted that Step 1 has a closed form solution

$$\tilde{\mathbf{W}}_{k+1}=(2\mu_2\mathbf{I}+\mathbf{X}^T\mathbf{X}+\beta_k\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{X}+\beta_k\mathbf{Z}_k+\mathbf{Y}_k). \tag{3.7}$$

Step 2 can be solved using soft thresholding

$$(\mathbf{Z}_{k+1})_{mn}=S_{2\mu_1/\beta_k}((\tilde{\mathbf{W}}_{k+1})_{mn}-\frac{(\mathbf{Y}_k)_{mn}}{\beta_k}), \tag{3.8}$$

where 'mn' signifies individual entries in a matrix. $S_\alpha(\mathbf{.})$ is the shrinkage thresh-

olding operator defined on the entries ν of matrix $\mathbf{V}$ as [75]:

$$S_\alpha(\nu) = sign(\nu)\ max(0, |\nu| - \alpha), \tag{3.9}$$

where max denotes the maximum value. For the nonzero elements ν of $\mathbf{V}$, $sign(\nu) = \nu/\left|\nu\right|$, otherwise $sign(\nu) = 0$.

Algorithm 1 Solving problem (3.4) by ADMM

Input: data matrix $\mathbf{X}$, parameters μ_1 and μ_2.

Initialize: $\beta_0 = 0.1$, $\beta_{max} = 10^{10}$, $\rho_0 = 1.1$, $\mathbf{Y}_0 = \tilde{\mathbf{W}}_0 = \mathbf{0}$, $\mathbf{Z}_0 = \text{randn}(N, N)$, k=0.

while not converge **do**

1: Fix the other variables and update $\tilde{\mathbf{W}}$ by

$$\tilde{\mathbf{W}}_{k+1} = \underset{\tilde{\mathbf{W}}}{argmin}\ \frac{1}{2}\left\|\mathbf{X} - \mathbf{X}\tilde{\mathbf{W}}\right\|_F^2 + \mu_2\left\|\tilde{\mathbf{W}}\right\|_F^2 + \frac{\beta_k}{2}\left\|\mathbf{Z}_k - \tilde{\mathbf{W}} - \frac{\mathbf{Y}_k}{\beta_k}\right\|_F^2,$$

where solution of above equation is provided in (3.7)

2: Fix the other variables and update $\mathbf{Z}$ by

$$\mathbf{Z}_{k+1} = \underset{\mathbf{Z}}{argmin}\ \mu_1\left\|\mathbf{Z}\right\|_1 + \frac{\beta_k}{2}\left\|\mathbf{Z} - (\tilde{\mathbf{W}}_{k+1} - \frac{\mathbf{Y}_k}{\beta_k})\right\|_F^2.$$

where solution of above equation is provided in (3.8)

3: Update the multipliers by

$$\mathbf{Y}_{k+1} = \mathbf{Y}_k + \beta_k(\mathbf{Z}_{k+1} - \tilde{\mathbf{W}}_{k+1}).$$

4: Update the parameter β by β_{k+1}=$min(\beta_{max}, \rho\beta_k)$

$$\rho = \begin{cases} \rho_0 & if\ \ max(\left\|\mathbf{Z}_{k+1} - \mathbf{Z}_k\right\|_\infty, \left\|\tilde{\mathbf{W}}_{k+1} - \tilde{\mathbf{W}}_k\right\|_\infty) > 10^{-4} \\ 1 & otherwise \end{cases}$$

5: Check the convergence condition

$$\left\|\mathbf{Z}_{k+1} - \tilde{\mathbf{W}}_{k+1}\right\|_\infty < 10^{-7}$$

6: Update $k : k \leftarrow k+1$

end while

Output: $\tilde{\mathbf{W}}$

After solving for the coefficient matrix $\tilde{\mathbf{W}}$, a positive symmetric adjacency matrix $\mathbf{A}$ is obtained by $\mathbf{A} = (\left|\tilde{\mathbf{W}}\right| + \left|\tilde{\mathbf{W}}\right|^T)/2$. This is to note that averaging of $\mathbf{W}$ matrix and its transpose to extract symmetric FC matrix is a sub-optimal method and is one of the limitation of the proposed MVRC method. Ideally, the condition of symmetry should be added within the optimization formulation. In future, we would like to learn symmetric coefficient matrix within the optimization framework of MVRC method.

3.4 Functional Brain Network Identification

After computing connectivity matrix, next task is to identify communities/fBNs. Previous studies have reported that connectivity matrix used in FC analysis has some structure and it commonly comprise of a number of communities [13, 62]. These communities are functionally coupled nodes/regions of the brain. Each community is comprised of a number of nodes and represents one functional brain network. Various methods have been proposed to detect communities from connectivity matrix [76]. Among them, modularity is one of the most popular methods [62] as described in the previous chapter.

Modularity is one of the key metrics that helps in the identification of locally clustered communities [13, 62]. It divides ROI in a manner that strengthens intra-brain network compactness and inter-brain network separability. We used Newman modularity algorithm implemented in Brain Connectivity Toolbox (BCT) [62] (refer to chapter 2). This algorithm maximizes modularity Q value, where high Q value signifies prominent intra-network organization [62].

3.5 Results

Dataset

Simulated Data Description

We used publicly available "NetSim fMRI simulated dataset" generated using dynamic causal modeling (DCM) [77, 1]. This dataset is generated using realistic characteristics similar to those in real fMRI data [1]. In this chapter, we chose first

four simulated dataset, Sim1 to Sim4, having 5, 10, 15, and 50 number of nodes, respectively. Details of the same are provided in Table 3.1. For each simulated dataset, data is available for 50 subjects.

Table 3.1: **Summary of simulated dataset [1] used to check performance of proposed method**

Sim	# Nodes (N)	Time points (T)
1	5	200
2	10	200
3	15	200
4	50	200

Real fMRI Data Description

We used open source Beijing_Zang resting-state fMRI dataset. This dataset is a part of Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC) 1000 Functional Connectomes Project [1]. A set of 198 normal subjects' data (122 females, 76 males, age range: 18-26 years old) was collected in this dataset by Beijing Normal University. No history of psychiatric disorder or any neurological illness was found in all 198 subjects.

A 3.0 Tesla MR Siemens scanner was used to collect resting-state fMRI data of all subjects with their eyes closed. A gradient echo $T2^*$-weighted Echo Planar Imaging (EPI) sequence was applied to acquire functional images with Echo Time (TE) equal to 30 ms, Repetition Time (TR) equal to 2000 ms, flip angle = 80°, and Field Of View (FOV) = 240×240 mm^2. The fMRI brain data is collected over 450 seconds resulting in 225 brain volumes. Each brain volume consists of an acquisition of 33 axial interleaved ascending brain slices with dimension 64×64. A three-dimensional structural magnetization prepared rapid gradient echo (MPRAGE) $T1$-weighted images was acquired for all subjects with TE = 3.39 ms, TR = 2530 ms, flip angle = 7°, FOV = 240×240 mm^2, and the number of brain slices is equal to 128.

In addition to the above normal subjects' dataset, we also used another publicy available dataset that are described next. This dataset is a publicly available

[1] `http://fcon_1000.projects.nitrc.org/`

Autism dataset contributed by Georgetown University at the collection site of Autism Brain Image Data Exchange II (ABIDE II[2]). This dataset includes echo-planar images acquired on a Siemens 3.0 Tesla scanner using an TE equal to 30 ms and TR equal to 2000 ms. It includes fMRI data of 55 Typically Developing Controls (TDC) (age: 8.1−13.8 years) and 51 Autism Spectrum Disorder (ASD) subjects (age: 8.1−13.9 years) scanned during the resting-state. Subjects underwent a 5 min, 14 s resting-state scan with their eyes open in the awaken state. The fMRI data is collected for 152 brain volumes with 43 axial brain slices (dimension 64×64) in each volume. Table-3.2 presents the demographic information of this dataset.

Real fMRI Data Preprocessing

All real fMRI datasets are preprocessed using SPM12. Preprocessing starts with the removal of first few volumes to allow the magnetization to reach the steady state. We discarded first 10 and 5 volumes from Beijing_Zang and autism dataset, respectively. The other functional volumes are slice time corrected using the middle slice as a reference followed by motion correction. Motion correction ensures head motion below 2 mm or 2° or voxel-to-voxel correspondence across time. Functional scans are spatially normalized onto the MNI space. Further, data is smoothed with a Gaussian kernel with 6 mm FWHM.

Finally, we regress out nuisance variables (6 head motion parameters, average cerebrospinal fluid (CSF) signal from ventricular masks, and average white matter signal from white matter mask) from each voxel's time-series followed by bandpass filtering using a butterworth filter to reduce low frequency drift and high frequency noise in the frequency range of 0.01 to 0.1 Hz.

After preprocessing, whole brain data is parcellated into 90 nodes or regions of interests (ROIs) via AAL atlas [50] (refer to section 2.2.1). In order to find node-representative time series for every node, we averaged time-series of all voxels belonging to the same node. This resulted into a matrix $\mathbf{X}$ of dimension $T \times 90$, where T denotes the number of time points (or the number of brain volumes such as 215 for Beijing_Zang data) of a given fMRI data. Next, we normalized each column of $\mathbf{X}$ to have a unit normalized time-series.

[2] `http://fcon_1000.projects.nitrc.org/indi/abide/`

Table 3.2: **Summary of demographic information of the patients and controls in autism dataset**

	Dataset	GU
Characteristic	TDC (S=26)	ASD (S=35)
Gender	Male	Male
Age (years)		
Mean (SD)	10.9 (1.62)	11.17 (1.49)
Range	8.06−13.79	8.25−13.91
Full Scale IQ		
Mean (SD)	120.32 (13.53)	119.06^{γ} (14.18)
Range	91−148	95−149
ADI-R		
Social total A	−	19.74^{γ} (5.28)
Verbal total BV	−	14.97^{γ} (4.91)
RRB total C	−	5.09^{γ} (2.38)
R Onset total D	−	2.56^{γ} (1.23)
ADOS		
Total	−	10.52^{Υ} (4.61)
Communication	−	3.18^{Υ}(1.54)
Social	−	7.33^{Υ} (3.53)
Stereo Behavior	−	1.89^{Υ} (1.58)

No statistical difference is observed in the age of ASD versus TDC groups of both the dataset with the two sample t-test at the significance level of $p \leq 0.05$. $^{\gamma}$One subject's score is missing. $^{\Upsilon}$Seven ASD subjects do not have these scores. '−' signifies that these scores are not available for the TDC group.
TDC: Typically Developing Control; ASD: Autism Spectrum Disorder; ADI-R: Autism Diagnostic Interview-Revised; Social total A: Reciprocal Social Interaction Subscore A; Verbal total BV: Abnormalities in Verbal Communication Subscore; RRB total C: Restricted, Repetitive, and Stereotyped Patterns of Behavior; ADOS: Autism Diagnostic Observation; Stereo Behavior: Stereotyped Behaviors and Restricted Interest.

Methods used for Comparison

First, we list some traditional connectivity matrix construction methods used for comparison in this chapter.

- **Pearson correlation (CORR)**: This is pairwise correlation between the representative time-series of two nodes. Pearson correlation coefficients are

computed between all 90 nodes (extracted from AAL atlas), resulting in a 90×90 connectivity matrix with values lying between [-1, 1] where value absolute value 1 signifies highest correlation or connection between two nodes.

- **Partial correlation (PCORR)**: This is known as normalized correlation between two time series by regressing out influence of all other nodes. Partial correlation values are assumed to be equivalent to off-diagonal entries of the inverse covariance matrix, under the assumption of normality. Thus, first the covariance matrix is computed. Next, inverse of this matrix provides partial correlation values. However, computation of inverse is itself challenging because of low rank nature of the covariance matrix. Few studies have used regularization procedure to estimate partial correlation values. Generalized inverse or pseudo-inverse methods have also been used to compute partial correlation. These methods impose l^1-norm or LASSO penalty on the partial correlation values. More specifically, the above regularization methods are also known as sparse partial correlation estimation in the literature. We used an implementation of PCORR from [3]. This method imposes l^1-norm penalty while computing the inverse of covariance matrix.

- **ASR method** [69]: ASR method has been explained in section 3.1 and has been implementation as specified in [69]. In [69], λ required in (3.1) has been set to 0.5. Same values have been used in this chapter. We used an implementation of trace LASSO from [78]. For all 90 nodes, algorithm is run repeatedly to estimate 90×90 connectivity matrix.

Trade-off Parameters μ_1 and μ_2 in MVRC Method

In this subsection, we study parameters μ_1 and μ_2 required to be used in MVRC method in (3.4). The parameter μ_1 controls sparsity on the coefficient matrix $\tilde{\mathbf{W}}$, whereas μ_2 controls grouping among elements of the coefficient matrix. Suitable selection of these two parameters is important for the overall performance of MVRC method.

[3] `https://www.cs.ubc.ca/~schmidtm/Software/L1precision.\html`

We varied μ_1 and μ_2 as (10^{-7}, 10^{-6},..., 10^{6}, 10^{7}). Fig. 3.1 shows resulting Q values over these ranges of μ_1 and μ_2. For brevity, we present results on four random subjects, although our results are consistent on all 198 subjects of Beijing_Zang rs-fMRI dataset.

This is to note that higher value of Q signifies prominent inter-communities separability and hence, is more desirable. We observe that MVRC method performs best at $\mu_1 = 10^{-1}$ (as Q value is high only at $\mu_1 = 10^{-1}$) and over a narrow range of μ_2 varying between 10^{-1} to 10^{-7}. In addition, we also observe from Fig.3.1 that the accuracy of MVRC method is stable when μ_2 is high irrespective of μ_1 (i.e., Q is nearly constant for μ_2 lying between 10^{1} to 10^{7}). Hence, high value of μ_2 is desirable within the range specified above. With these observations, we have a shorter range to choose from for parameter selection.

To further narrow down the values of parameters, we empirically tested different values of μ_1 varying between 0.1 to 0.9 and μ_2 varying between 0.01 to 0.9. Finally, we selected $\mu_1 = 0.25$ and $\mu_2 = 0.85$ that worked satisfactorily on all real subjects' data.

3.5.1 Results on simulated fMRI Data

In this subsection, we demonstrate the performance of MVRC method on a simulated dataset of 50 subjects. To compare the ground truth adjacency matrix and the computed adjacency matrix, we use "c-sensitivity" measure on simulated dataset as used in [1, 69]. This measure signifies the ability to correctly identify presence of a connection. It is calculated as the ratio of number of true connection whose strength is greater than 95% percentile of the false positive connection strength to the total number of true connections [1]. A value of 95% is chosen as robust rank statistics as also used in [1].

We estimated c-sensitivity measure on the adjacency matrix computed using different methods. High value of this measure signifies identification of higher number of ground truth connections. We selected $\mu_1 = 0.1$ and $\mu_2 = 0.6$ in the MVRC method for the simulated fMRI data based on c-sensitivity, i.e., we obtained highest c-sensitivity index with the above values of μ_1 and μ_2.

Values of c-sensitivity is calculated for each individual subject separately and is

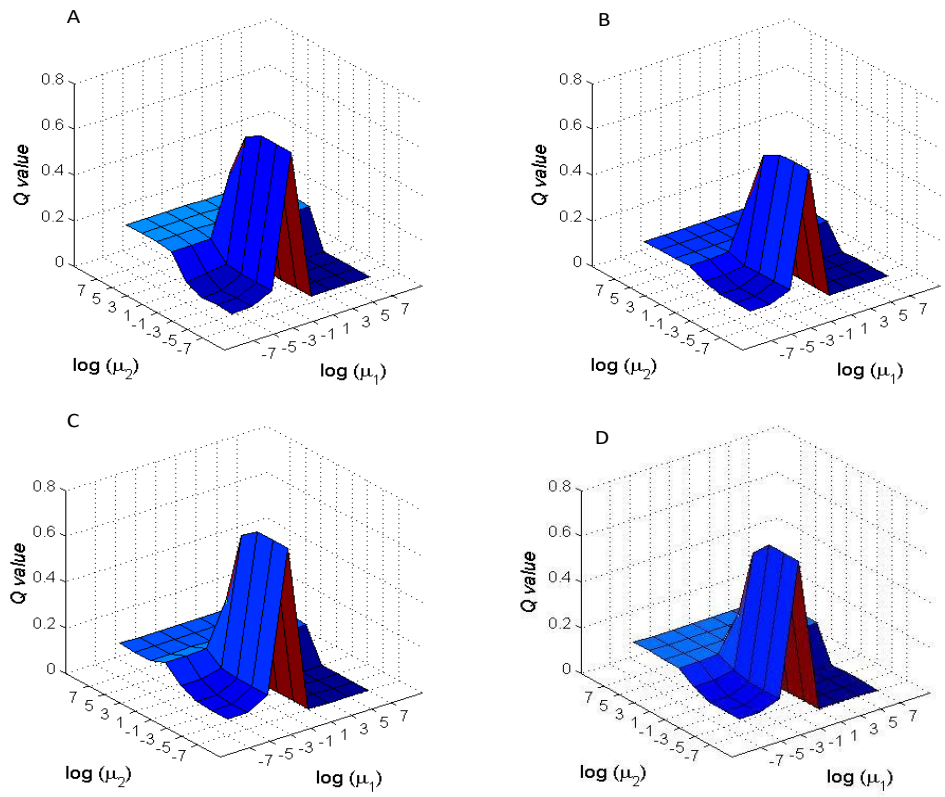


Figure 3.1: Q value vs μ_1 and μ_2 on four real fMRI subjects

averaged over all 50 subjects in all four datasets. Mean and variance of c-sensitivity values are presented in Table 3.3. Results are tabulated on all 4 simulated dataset. Higher value of c-sensitivity signifies better ability to identify connections. From Table-3.3, we note a higher mean value of c-sensitivity with the proposed MVRC method. Hence, we can infer that the proposed MVRC method consistently identifies presence of true connections. Performance of MVRC and ASR methods are comparable on all 4 dataset with reference to the mean sensitivity values. Partial correlation method provides lower value of c-sensitivity on all 4 dataset. Moreover, we also observe that Pearson correlation results are inferior compared to the proposed MVRC method.

Results of c-sensitivity obtained with MVRC have been compared statistically with other methods using the Mann-Whitney U-test, also known as the Wilcoxon rank-sum test. This test has been used instead of t-test because c-sensitivity values are not observed to follow normal distribution. This test is a two-sided test with the null hypothesis that the medians of two populations are identical and with the

alternative hypothesis that they are significantly different. The test is performed between MVRC and the other methods of CORR, PCORR and ASR. Resulting p-values are shown in brackets of Table 3.3. Value of p less than 0.05 (5% significance level) signifies that the median of c-sensitivity values of the method compared, i.e., CORR, PCORR, or ASR, is significantly different from that of MVRC.

Table 3.3: **The average results of c-sensitivity over 50 subjects.**

Dataset	**CORR** (c-sensitivity) mean±std (p-value)	**PCORR** (c-sensitivity) mean±std (p-value)	**ASR** (c-sensitivity) mean±std (p-value)	**Proposed MVRC** (c-sensitivity) mean±std
Sim1	0.80±0.18 (0.03)	0.75±0.09 (0.001)	0.86±0.14 (**0.6012**)	0.84±0.17
Sim2	0.80±0.13 (0.04)	0.72±0.09 (0.00)	0.86±0.10 (**0.44**)	0.85±0.10
Sim3	0.80±0.11 (0.04)	0.70±0.08 (0.005)	0.85±0.09 (**0.53**)	0.84±0.09
Sim4	0.83±0.03 (0.00)	0.70±0.04 (0.01)	0.89±0.03 (**0.21**)	0.88±0.04

p-values evaluated for a significance level of 0.05 using a two sided Mann-Whitney U-test between the proposed MVRC and other methods. std stands for standard deviation.

Bold p-values in Table 3.3 represent no significant difference between the medians of the c-sensitivity values of the two methods compared. From these values, we observe that MVRC and ASR methods result in statistically similar c-sensitivity values on all 4 dataset. We also observe that c-sensitivity values of CORR and PCORR are statistically different from MVRC which is inline with our mean c-sensitivity value results in Table 3.3.

3.5.2 Results on real fMRI Data

In this subsection, we first present results on Beijing_Zang rs-fMRI dataset using the proposed MVRC method and other existing methods mentioned in section 3.5. This is to show the effectiveness of the proposed method in extracting fBNs on a group of subjects. Fig.3.2 shows 90×90 connectivity matrices obtained using

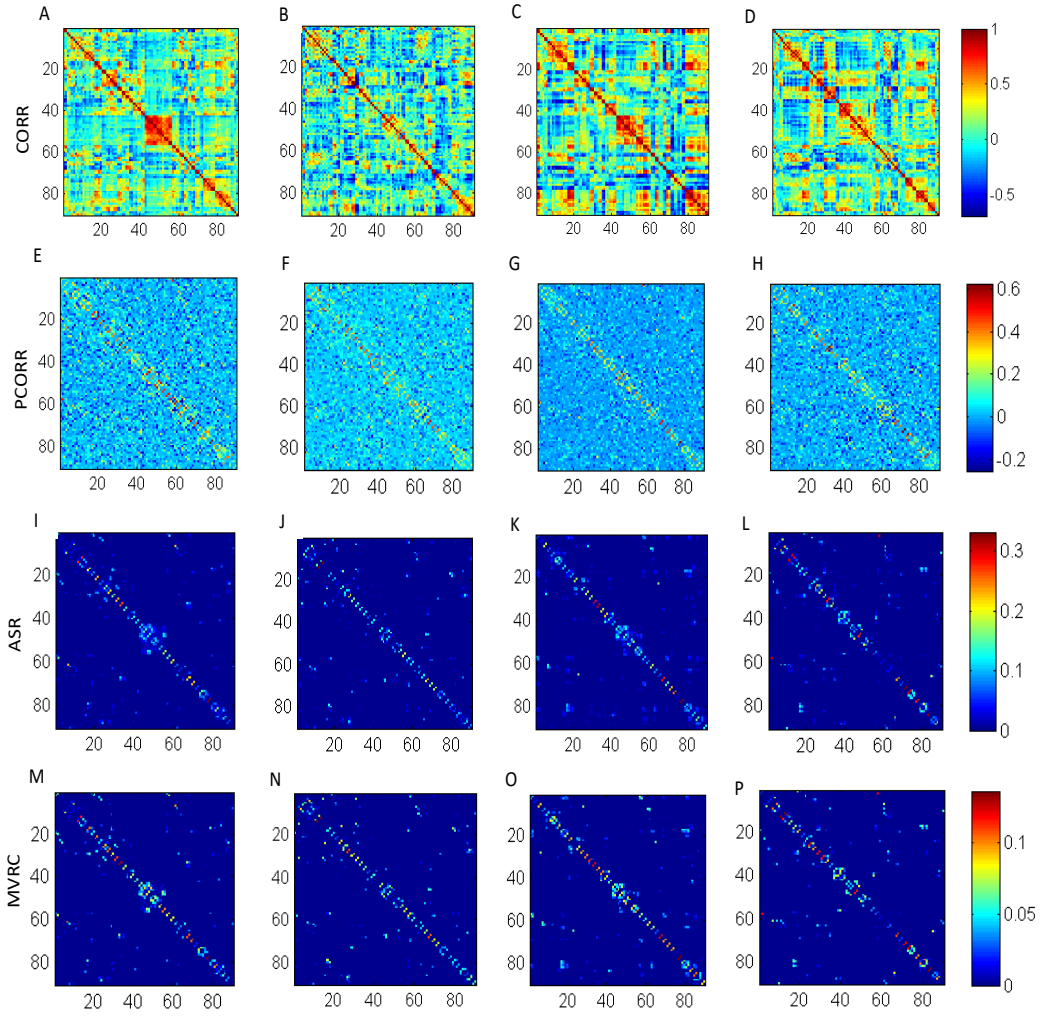


Figure 3.2: Connectivity matrices computed on four (randomly selected) subjects of subgroup-1 of Beijing_Zang fMRI dataset. (A-D): Using CORR; (E-H): Using PCORR; (I-L): Using ASR; (M-P): Using MVRC.

CORR, PCORR, ASR, and MVRC methods on four (randomly selected) subjects of subgroup-1 (refer to Table 3.1). Similarly, Fig.3.3 shows connectivity matrix on four (randomly selected) subjects of subgroup-2. We observe that the proposed MVRC is able to capture sparsity of the connectivity matrix better than the ASR method.

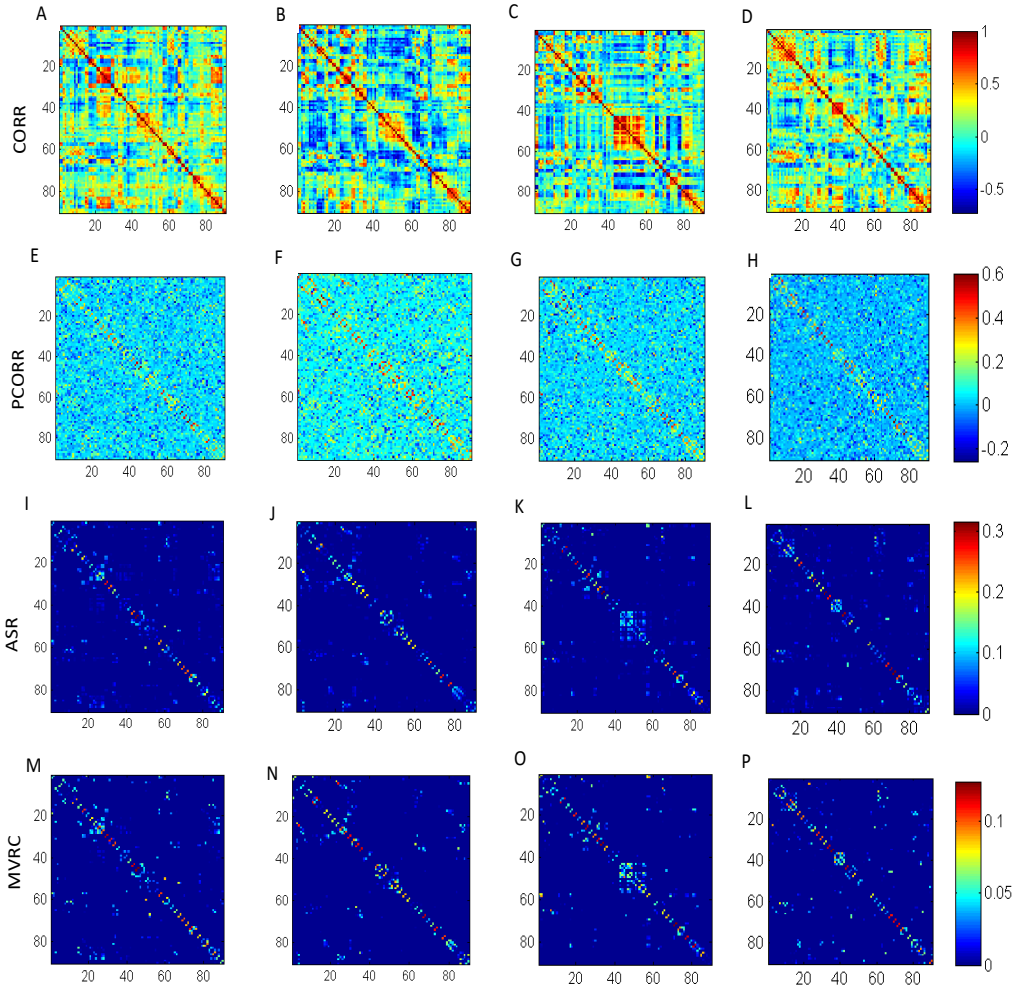


Figure 3.3: Connectivity matrices computed on four (randomly selected) subjects of subgroup-2 of Beijing_Zang fMRI dataset. (A-D): Using CORR; (E-H): Using PCORR; (I-L): Using ASR; (M-P): Using MVRC.

In the current study, communities are identified using modularity metric as aforementioned in Section-2.4. Fig.3.4 shows connectivity matrices computed before and after running the modularity algorithm on one randomly chosen subject in subgroup-1. Modularity algorithm rearranges the communities, i.e., identifies coupled regions and arranges them together. It results in a block diagonal ma-

trix with each block signifying one community of highly connected regions or one functionally connected brain network.

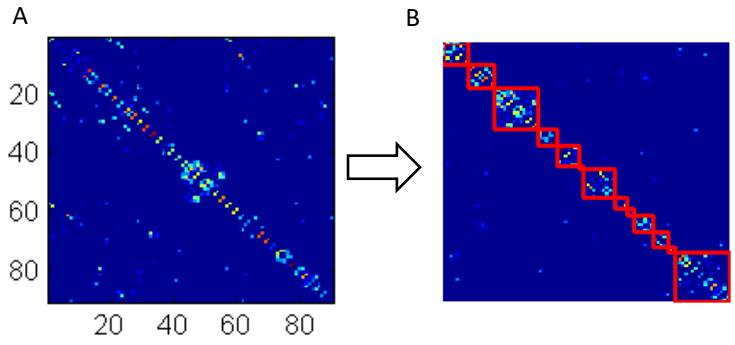


Figure 3.4: Modularity results on one subject of subgroup-1 of Beijing_Zang fMRI dataset. (A) is connectivity matrix computed using MVRC on one randomly chosen subject, (B) is connectivity matrix obtained after modularity with region number reordering where regions numbers 1 to 90 are AAL atlas region numbers as shown in Table 2.1.

From Fig. 3.4, we observe that modularity helps in extracting connected communities from the MVRC connectivity matrix. Each community represents one fBN. We discarded communities having fewer number of nodes. For comparison with other methods, we present connectivity matrices obtained after modularity on the same subject's data using CORR, PCORR and ASR method in Fig.3.5.

We obtain only three big communities with the CORR method. Moreover, no community signifies any fBN structure. For example, in one community with CORR method, we obtain node numbers 17-20, 29, 30, 41, 42, 57-60, 63, 64, 69, 70, 73-75, 79-85, 89. However, from the literature, it is known that these nodes belong to different brain networks. At the bottom of Table-2.1 we provide exhaustive list of nodes residing in any resting-state fBNs category such as visual network, auditory network, bilateral limbic network, default mode network (DMN), motor network, subcortical network, somatosensory network, and memory network. Similarly with PCORR method, we observe fewer number of communities and each community is filled with random nodes. Hence, fBNs could not be identified with CORR and PCORR methods. We observe more and valid communities with both ASR and

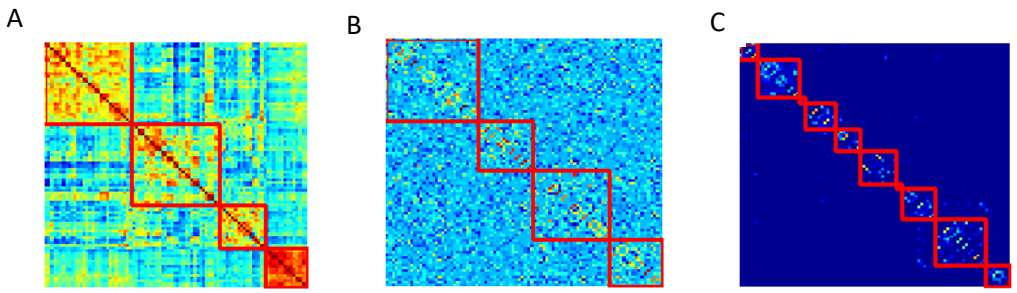


Figure 3.5: Modularity results on Beijing_Zang fMRI dataset subgroup-1. Connectivity matrix with node number reordering computed using (A) CORR, (B) PCORR and (C) ASR.

MVRC methods. This is to note that MVRC method is able to identify intrinsic brain network's organization with higher Q value (Q=0.7923) compared to the ASR method (Q=0.7638).

Comparison of computational time

In this section, we compare the run time of ASR method and the proposed MVRC method. All experiments are performed on a computer with 2.4 GHz Intel Core i7 CPU and 16GB RAM. Table 3.4 shows the run time of these methods. From Table 3.4, we note that MVRC is computationally much faster than ASR method due to the fact that ASR solves for each region iteratively, while MVRC method solves it in one step or iteration. Hence, computational time with ASR is N times higher compared to the proposed MVRC method. Moreover, trace LASSO algorithm in ASR takes more steps to converge. The computational time in ASR method also increases with the increase in number of ROIs.

Visualization of Functional brain networks

Networks in communities are visualized using "circularGraph" toolbox [4]. We present results on three randomly chosen subjects of both subgroup with ASR

[4] https://in.mathworks.com/matlabcentral/fileexchange/48576-circulargraph

Table 3.4: **Results on Computational Time (in secs)**

Dataset	ASR Method	Proposed MVRC method
Real (female)	84.56 (12755.091)	**0.125 (707.088)**
Real (male)	88.92 (8970.016)	**0.119 (402.309)**

Results are tabulated on a single subject's data. Total run time over all subjects' data (116/74 subjects' real data) is shown in brackets.

and the proposed MVRC method. Fig.3.6 and Fig.3.7 visualize networks over three random subjects of subgroup-1 and subgroup-2, respectively.

From Figs.3.6 (A) and (B), we observe that both ASR and MVRC method are able to identify visual, subcortical, default mode networks (DMNs), although ASR method detects false nodes in auditory, bilateral limbic, and motor networks that are not supposed to be the part of corresponding brain networks. For example, superior frontal gyrus, medial orbital (node 25 and 26) and anterior cingulate and paracingulate gyri (node 31 and 32) are not part of bilateral limbic network, but are identified by ASR in the limbic network. Similarly node 83 to 86 are detected as part of auditory network, although they are not. Hence, auditory network identified by ASR network is not accurate. Similarly, we observe many false positive network nodes in motor network detected by ASR method. Thus, we note that MVRC has greater accuracy than ASR.

From Figs.3.6 (C) and (D), we observe that MVRC method is able to identify visual, subcortical, DMNs, bilateral limbic, memory, and motor network. ASR method could identify these but could not identify memory network on this subject. Also, MVRC method could identify more number of active regions in DMN network compared to the ASR identified DMN network.

From Figs.3.6 (E) and (F), we observe that visual, motor, subcortical, and bilateral limbic networks are identified on this subject using both the ASR and MVRC method. In addition, DMN network with two active regions and the auditory network are also found to be active using the MVRC method, though they are not found with the ASR method. However, rolandic operculum (regions 17 and 18) and insula (regions 29 and 30) regions are also found to be simultaneously activated with the auditory regions, although these are not part of the auditory

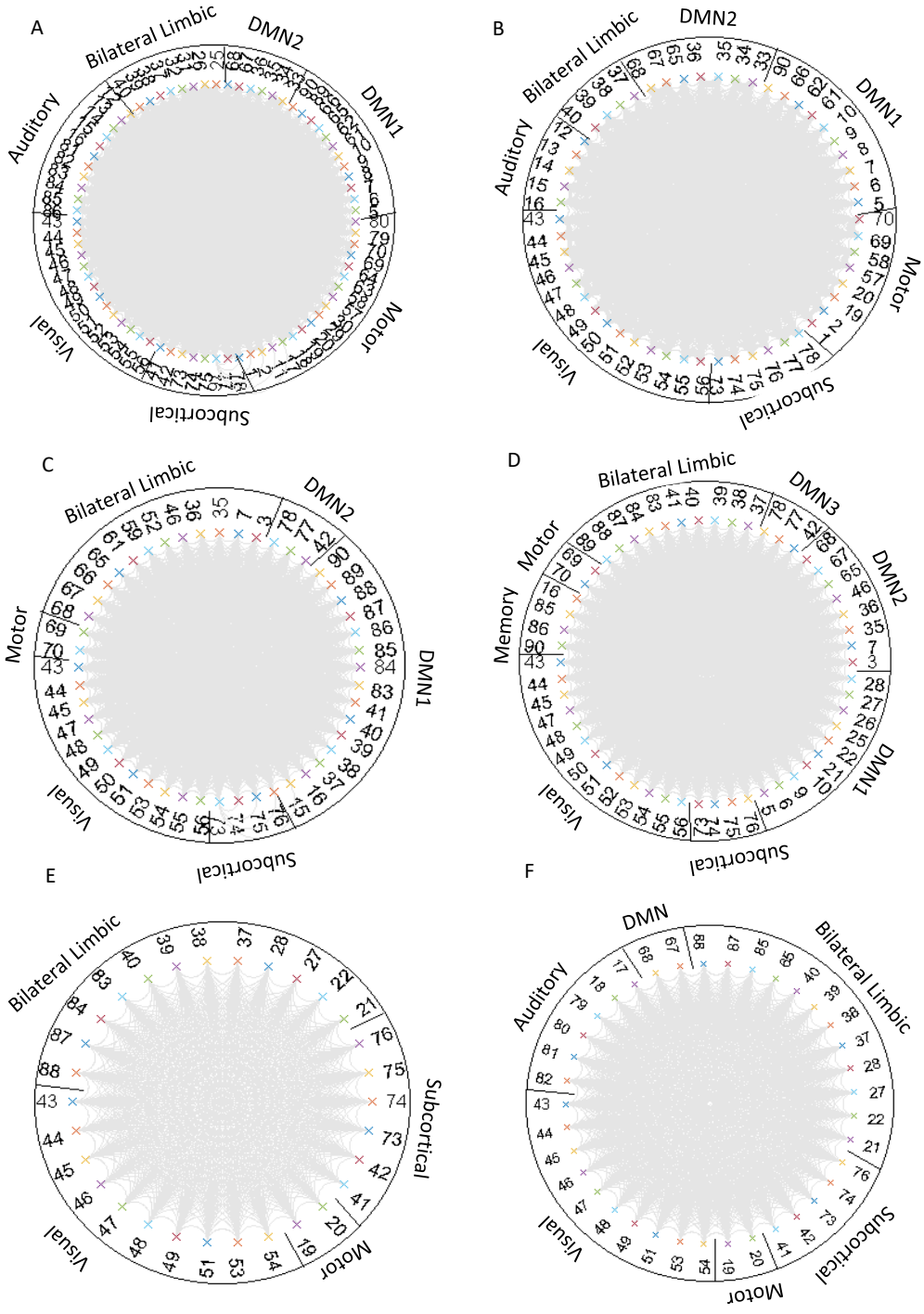


Figure 3.6: Brain networks identified on three randomly chosen subjects of subgroup-1 of Beijing_Zang datset, where each row corresponds to results on one subject. Left column represents network identified using the ASR method and the right column represents network identified using the proposed MVRC method.

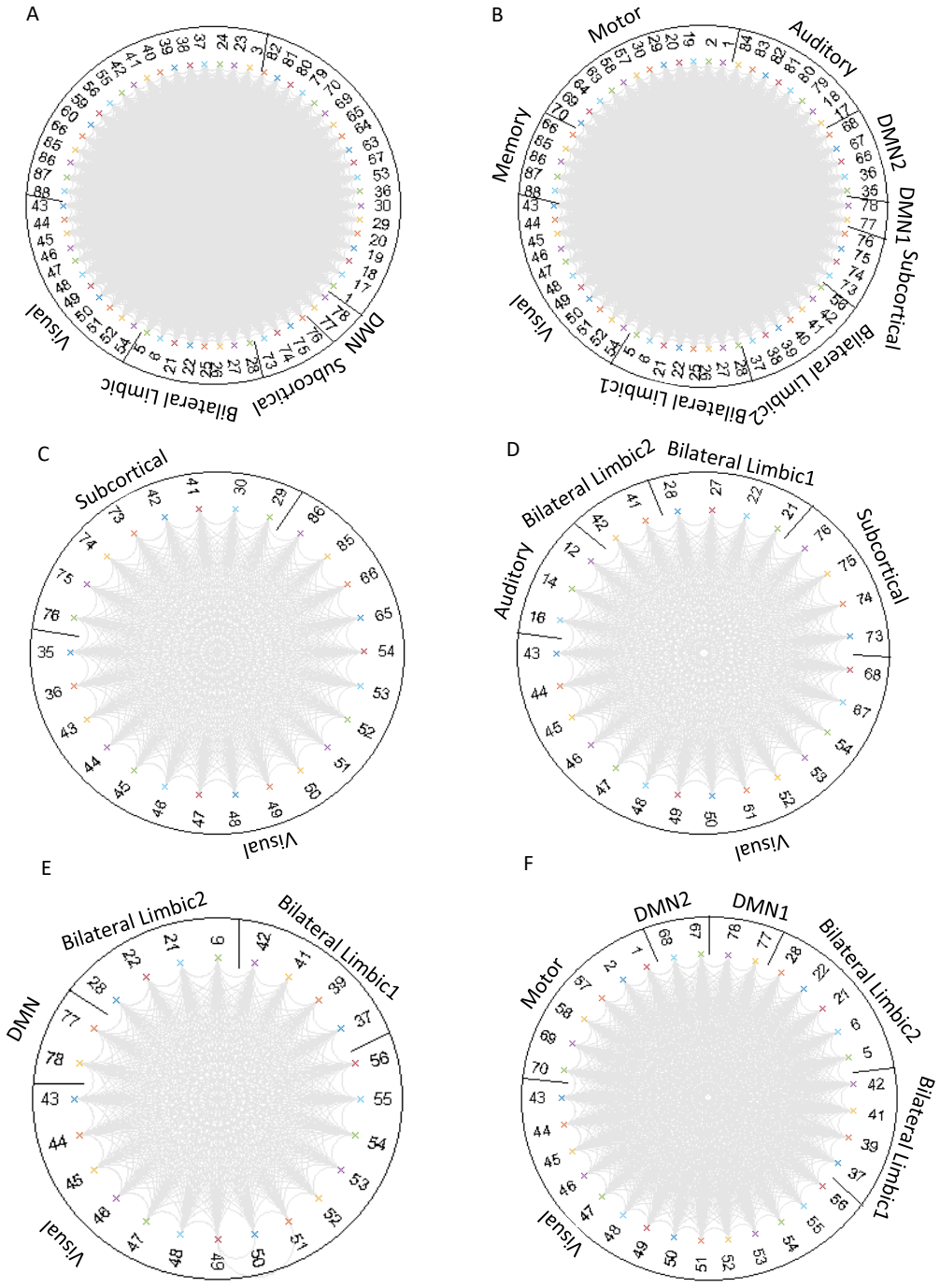


Figure 3.7: Brain networks identified on three randomly chosen subjects of subgroup-2 of Beijing_Zang datset where each row corresponds to results on one subject. Left column represents network identified using ASR method and right column represents network identified using the proposed MVRC method.

network.

From Figs.3.7 (A) and (B), we observe that MVRC could identify visual, bilateral limbic, subcortical, DMN, auditory, motor, and memory network, while ASR could only identify visual, bilateral limbic, subcortical, and DMN network. This is due to the less number of identified communities in the case of ASR (=14) compared to MVRC (=18). Hence, similar to the CORR and PCORR methods, larger and spurious number of regions are clubbed in one community with ASR as shown in Fig. 3.7(A), although this effect is less in ASR compared to the CORR and PCORR bivariate methods. Compared to the MVRC method that has two regularization parameters and hence, is more flexible, ASR method has only one regularization parameter that constrains its performance. With the MVRC method, we are able to control sparsity and denseness terms separately while calculating the adjacency matrix. Moreover, this is to note that within bilateral limbic and DMN networks, ASR identifies less number of regions compared to the MVRC method. This shows poor (or false negative) network identification of the ASR method.

From Figs.3.7 (C) and (D), we again observe false negative network identification with ASR. MVRC could identify visual, subcortial, bilateal limbic and auditory network on this subject, while ASR could identify only visual and subcortical network. In addition, false positive spurious regions are identified in both visual and subcortical networks with ASR. For example, posterior cingulate gyrus (region 35 and 36) and angular gyrus (region 65 and 66) and middle temporal gyrus (region 85 and 86) are not part of visual network but are identified by ASR. Similarly region 29, 30, 41, and 42 are not part of subcortical network. Hence, subcortical network identified by ASR network is not accurate.

From Figs.3.7 (E) and (F), we observe that fewer number of functional brain networks are observed to be active on this subject compared to the other subjects. Visual, bilateral limbic, motor, and two communities of DMN networks are identified using the proposed MVRC method. This is to note that ASR method could not identify motor network and one DMN network on this subject. Rest other networks could be identified by the ASR method.

From above brain networks analysis, we can state that MVRC method is able to identify robust brain networks compared to the ASR method. This is due to

flexible design of MVRC method in terms of two regularization terms. We are able to control sparsity and denseness terms separately while calculating the adjacency matrix. On the other hand, ASR method has only one regularization parameter that constrains the design. Compared to the ASR method, less no. of spurious false positive regions are observed with the proposed MVRC method. Moreover, we observe that ASR method is not able to identify some networks on few subjects. Also, the proposed MVRC method is computationally much faster compared to the ASR method.

Reproducibility of results

In this subsection, we test the reproducibility of brain networks identified using MVRC method in each subgroup. Figs.3.8 A and B show the histogram of each functional network in a 116 female subgroup subgroup-1 and 74 male subgroup subgroup-2, respectively. From this figure, we observe that visual network is identified in 113 subjects of SG-1 and 72 subjects of SG-2. This finding is in inline with the literature that visual network is generally found to be active in rs-fMRI.

In addition, MVRC could identify default mode network in 104 subjects of subgroup-1 and 66 subjects of subgroup-2. Default mode network is the one of main functional network during resting-state. We observe that our MVRC method could identify DMN on a large population in both subgroups. Another network which MVRC identified on large population is subcortical network. Subcortical network is identified in 104 subjects of subgroup-1 and 67 subjects of subgroup-2. Furthermore, bilateral limbic network is identified in 109 subjects of SG-1 and 67 subjects of SG-2. Motor network is identified in 66 subjects of SG-1 and 55 subjects of SG-2. This implies that motor network in males may be more predominant than females. Memory network is identified in 29 subjects of SG-1 and 10 subjects of SG-2. This indicates that perhaps females memorize things more inside the scanner compared to men. Auditory network is identified in 31 subjects of SG-1 and 10 subjects of SG-2 and Somatosensory network is identified in 21 subjects of SG-1 and 5 subjects of SG-2.

From the above findings, we observe that functional networks identified with MVRC method are reproducible across subjects.

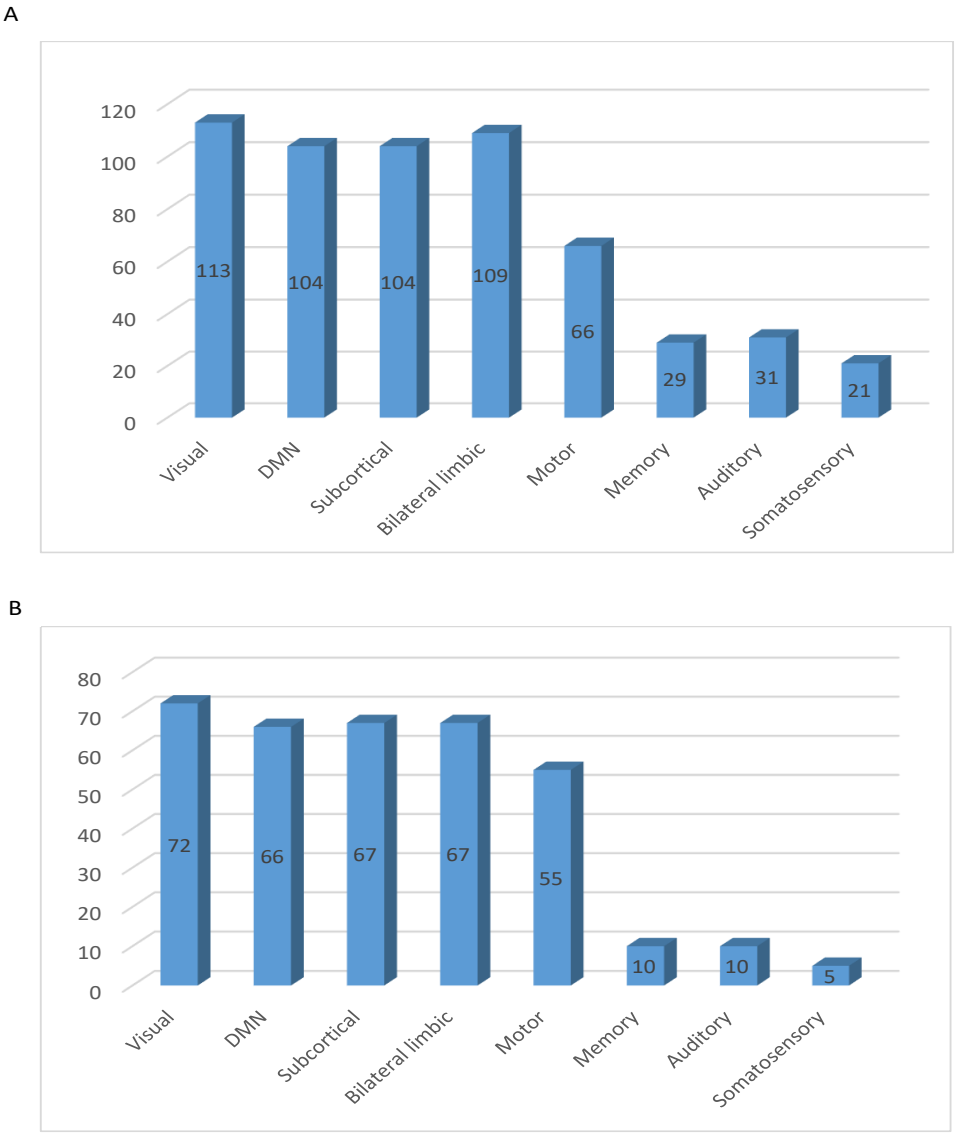


Figure 3.8: Number of subjects in each subgroup reporting presence of functional network A) subgroup-1; B) subgroup-2 of Beijing_Zang datset.

3.5.3 Statistical validation of the proposed MVRC method

While functional connectivity offers insights of the brain networks, it is also critical to determine whether the connectome we observe differs significantly from the expected null model using an appropriate statistical parameter. Therefore, FC assessment requires an appropriate statistical testing of connectivity where we can define and test an appropriate statistical null test.

In this section, we discuss an appropriate statistical test where various graph theoretical measures [62] extracted from the MVRC based connectivity matrices are assessed and compared against a null distribution to probe the true connectivity pattern. We utilize a group comparison framework that focuses on consistency of functional hubs or key nodes within and between both the groups. In other words, this method compares the consistency of hub organization between the normal and patient groups of autism dataset.

To do so, first, top 5% key nodes or hubs are identified for each subject of both the groups. These key nodes are identified based on various node-based graph theoretical measures. Measures signifying functional brain network integration such as participation coefficient and betweenness centrality are used to identify important key nodes (hubs). This results in two key-nodes sets. The first set C is corresponding to the normal control group and the second set A is corresponding to the autism patient group. Further, to check the similarity between these two sets, we utilize Jaccard index (JI), a metric to represent similarity between sets, as $\frac{|C \cap A|}{|C \cup A|}$, where $|C \cap A|$ is the total number of key nodes that overlap between two groups and $|C \cup A|$ is the total number of key distinct nodes in both groups. JI ranges between 0 (no overlap between groups) to 1 (perfect overlap indicating high similarity across two groups).

We test the statistical significance of obtained JI value using a permutation test [79]. This test helps to generate null hypothesis distributions of JI under the condition of no difference among groups. This distribution is generated by randomly permuting each subject's label across the two groups and re-calculating the JI over a large number of permutations. In other words, this process starts with calculating JI based on original group labels and than re-calculating again and again over randomly permuted group labels. These randomly permuted labels based JI values (JI^{perm}) are used to generate empirical distribution of JI. Fig. 3.9

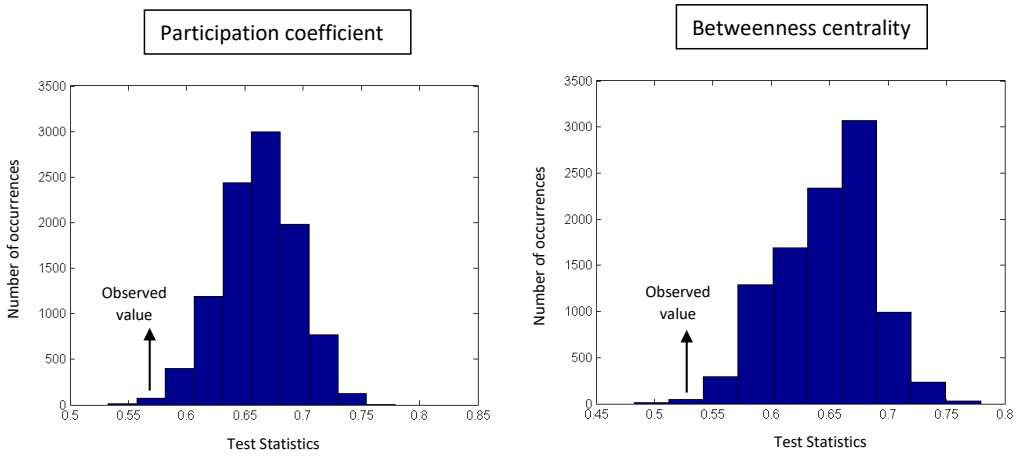


Figure 3.9: Empirical distribution generated by the permutation process for the Jaccard values corresponding to participation coefficient (left) and betweenness centrality (right).

presents this empirical distributions corresponding to both PC and BC measures.

By comparing original JI value to this empirical distribution, we can determine the significance of the group difference. In particular, the p-value for this comparison can be computed as $p = \frac{|JI^{perm} > JI|)}{Total\ no.\ of\ permutation}$ where, $|JI^{perm} > JI|$ is the number of permutations for which JI^{perm} is greater than original JI value. If the p-value is small, it indicates that the observed JI value did not occur by chance, signifying strong evidence against null hypothesis. In our work, we generated JI distribution based on 10000 permutations on both PC and BC measures. The results of our experiment indicated statistically different JI with p-value$<$0.05.

3.6 Conclusions

In this chapter, we present multivariate vector regression-based connectivity (MVRC) method. This proposed method considers linear relationship among all nodes representative time series simultaneously. On the other hand bivariate methods, such as correlation and partial correlation, calculate pairwise association between two nodes without considering influence of other nodes. Owing to this, spurious association values may arise. MVRC employs both l^1- and l^2-norms on connectivity matrix coefficients. While l^1-norm helps in identification of sparse brain networks, l^2-norm helps in the identification of correlated brain networks. This allows both

sparsity and grouping among spatially adjacent brain nodes. Results on fMRI data demonstrate that MVRC method is more robust in identification of fBNs compared to the existing methods.

Chapter 4

Low Rank and Sparsity Constrained Method for Identifying Overlapping Functional Brain Networks

Apart from considering sparsity and denseness constraints in MVRC method explained in Chapter-3, an informative FC graph is essentially low rank [80, 81]. This implies that columns and rows of connectivity matrix are lying in the lower dimensional space, where rows and columns signify brain nodes. However, none of the existing methods utilize low-rank assumption in their formulation. Inspired by this insight, we propose a new multivariate method for learning connectivity matrix with an assumption that this matrix is reasonably low rank. In addition, we also impose sparsity constraint to learn sparse fBNs. We name this proposed method as Low Rank-Multivariate Vector Regression-based Connectivity (LR-MVRC) [82].

Moreover, we propose to extract overlapping networks. In many instances, the communities are characterized as combination of disjoint brain nodes, although recent studies indicate that nodes may participate in more than one community. In this chapter, large-scale overlapping fBNs are identified on rs-fMRI data by employing non-negative matrix factorization. Our findings support the existence of overlapping fBNs. The section 4.1 describes the problem formulation of the proposed method followed by implementation details in the subsequent sections.

4.1 Proposed Formulation

Consider a matrix $\mathbf{X}$ of dimension $T \times N$, where T denotes the number of time points and N denotes the number of region of interests (ROIs). Each column of $\mathbf{X}$ signifies unit normalized time-series of one brain region. Given this matrix, we require to compute the functional connectivity matrix of dimension $N \times N$. Recently proposed Multivariate Vector Regression-based Connectivity (MVRC) method of identifying FC regresses time series of all regions (i.e. columns of $\mathbf{X}$) onto the time series of other regions multiplied by an FC matrix as $\mathbf{X} = \mathbf{X}\tilde{\mathbf{W}}$ [68]. In addition, this method employs elastic-net penalty onto the FC matrix comprising of both l^1 and l^2-norms. However, to incorporate modular structure of fBNs into FC formulation, we consider nuclear norm constraint along with l^1-norm constraint in the computation of this matrix and name the propose method as Low-Rank Multivariate Vector Regression-based Connectivity (LR-MVRC). We formulate the proposed LR-MVRC objective function as:

$$\underset{\tilde{\mathbf{W}}}{min}\ \frac{1}{2}\left\|\mathbf{X} - \mathbf{X}\tilde{\mathbf{W}}\right\|_F^2 + \mu_1 \left\|\tilde{\mathbf{W}}\right\|_1 + \mu_2 \left\|\tilde{\mathbf{W}}\right\|_*,\ \text{s.t.}\ \ diag(\tilde{\mathbf{W}}) = 0, \tag{4.1}$$

where μ_1 and μ_2 are the regularization parameters associated with l^1 and nuclear norms terms, respectively. Nuclear norm minimization implies l^1 penalty on singular values of matrix $\tilde{\mathbf{W}}$ that supports this matrix to be low-rank. $diag(\tilde{\mathbf{W}}) = 0$ term ensures no self connections in the matrix $\tilde{\mathbf{W}}$. Finally, symmetric adjacency matrix from $\tilde{\mathbf{W}}$ is computed as $\mathbf{A} = (|\tilde{\mathbf{W}}| + |\tilde{\mathbf{W}}|^T)/2$.

Implementation of LR-MVRC

Next, we introduce the algorithm to solve LR-MVRC problem. We utilize ADMM [71] that splits eq. (4.1) into multiple subproblems that are easier to solve. We introduce two auxiliary variables $\mathbf{Z}_1$ and $\mathbf{Z}_2$ for the l^1-norm and nuclear norm terms in eq. (4.1) as:

$$\begin{aligned}\underset{\tilde{\mathbf{W}}}{min}\ \frac{1}{2}\left\|\mathbf{X} - \mathbf{X}\tilde{\mathbf{W}}\right\|_F^2 + \mu_1 \left\|\mathbf{Z}_1\right\|_1 + \mu_2 \left\|\mathbf{Z}_2\right\|_* \\ \text{s.t.}\ \ \mathbf{Z}_1 = \tilde{\mathbf{W}}, \mathbf{Z}_2 = \tilde{\mathbf{W}}, diag(\tilde{\mathbf{W}}) = 0.\end{aligned} \tag{4.2}$$

The augmented Lagrange function for eq. (4.2) can be written as:

$$\begin{aligned}\mathcal{L}(\tilde{\mathbf{W}}, \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Y}_1, \mathbf{Y}_2) = \frac{1}{2}\left\|\mathbf{X} - \mathbf{X}\tilde{\mathbf{W}}\right\|_F^2 + \mu_1 \left\|\mathbf{Z}_1\right\|_1 \\ +\mu_2 \left\|\mathbf{Z}_2\right\|_* + \frac{\beta_1}{2}\left\|\mathbf{Z}_1 - \tilde{\mathbf{W}} + \frac{\mathbf{Y}_1}{\beta_1}\right\|_F^2 + \frac{\beta_2}{2}\left\|\mathbf{Z}_2 - \tilde{\mathbf{W}} + \frac{\mathbf{Y}_2}{\beta_2}\right\|_F^2,\end{aligned} \tag{4.3}$$

where $\beta_1, \beta_2 > 0$ are the penalty parameters and $\mathbf{Y}_1$, $\mathbf{Y}_2$ are the Lagrangian multiplier matrices. The above equation consists of three variables $\tilde{\mathbf{W}}$, $\mathbf{Z}_1$ and $\mathbf{Z}_2$. This is to note that during the iterative learning for the solution of $\tilde{\mathbf{W}}$ in eq. (4.3), $diag(\tilde{\mathbf{W}})$ is kept to zero.

ADMM splits eq. (4.3) into three subproblems as described below. Each subproblem may be treated as minimization over one variable while fixing the other two variables.

Subproblem 1: Solving $\tilde{\mathbf{W}}$:

$$\underset{\tilde{\mathbf{W}}}{min}\ \frac{1}{2}\left\|\mathbf{X} - \mathbf{X}\tilde{\mathbf{W}}\right\|_F^2 + \frac{\beta_1}{2}\left\|\mathbf{Z}_1 - \tilde{\mathbf{W}} + \frac{\mathbf{Y}_1}{\beta_1}\right\|_F^2 + \frac{\beta_2}{2}\left\|\mathbf{Z}_2 - \tilde{\mathbf{W}} + \frac{\mathbf{Y}_2}{\beta_2}\right\|_F^2. \tag{4.4}$$

Update of $\tilde{\mathbf{W}}$, while other variables are fixed, is performed by solving the above equation as below:

$$\tilde{\mathbf{W}} = (\mathbf{X}^T\mathbf{X} + (\beta_1 + \beta_2)\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{X} + \beta_1\mathbf{Z}_1 + \mathbf{Y}_1 + \beta_2\mathbf{Z}_2 + \mathbf{Y}_2), \tag{4.5}$$

where $\mathbf{I}$ is an identity matrix. The diagonal elements of $\tilde{\mathbf{W}}$ obtained from eq. (4.4) are replaced with zeros.

Subproblem 2: Solving $\mathbf{Z}_1$:

$$\underset{\mathbf{Z}_1}{min}\ \mu_1 \left\|\mathbf{Z}_1\right\|_1 + \frac{\beta_1}{2}\left\|\mathbf{Z}_1 - \tilde{\mathbf{W}} + \frac{\mathbf{Y}_1}{\beta_1}\right\|_F^2. \tag{4.6}$$

Update of $\mathbf{Z}_1$, while other variables are fixed, can be done using soft thresholding as:

$$\mathbf{Z}_1 = Soft_{2\mu_1/\beta_1}(\tilde{\mathbf{W}} - \frac{\mathbf{Y}_1}{\beta_1}), \tag{4.7}$$

where $Soft$ is the shrinkage thresholding operator defined as [75]:

$$Soft_{\alpha}(\nu) = sgn(\nu)\max(0, |\nu| - \alpha), \tag{4.8}$$

where *sgn* denotes the signum value and *max* denotes the maximum value.
Subproblem 3: Solving $\mathbf{Z_2}$:

$$\underset{\mathbf{Z_2}}{min}\ \mu_2 \left\|\mathbf{Z}_2\right\|_* + \frac{\beta_2}{2}\left\|\mathbf{Z}_2 - \tilde{\mathbf{W}} + \frac{\mathbf{Y}_2}{\beta_2}\right\|_F^2. \tag{4.9}$$

Global minimum of convex nuclear norm minimization is obtained by soft thresholding on singular values, known as singular value thresholding (SVT) [83]. Hence, update of $\mathbf{Z}_2$ while other variables are fixed, can be carried out using soft thresholding on singular values of $\mathbf{Z}_2$ as:

$$\mathbf{Z}_2 = \mathrm{SVT}_{2\mu_2/\beta_2}(\tilde{\mathbf{W}} - \frac{\mathbf{Y}_2}{\beta_2}), \tag{4.10}$$

where SVT is defined as:

$$\mathrm{SVT}_{\alpha}(\nu) = \mathbf{U} \times diag(SOFT_{\alpha}(\nu)) \times \mathbf{V}^T, \tag{4.11}$$

and singular value decomposition of $\mathbf{Z}_2$ is defined as $\mathbf{U} \times diag(\nu) \times \mathbf{V}^T$.

The iterations of LR-MVRC, with update of variables, is described in *Algorithm* 2.

4.2 Functional Brain Network Identification

After computing the adjacency matrix, next task is to identify communities or fBNs. Various methods have been proposed to detect communities from FC matrix [76]. Among them, modularity is one of the most popular methods [62]. The widely used modularity method implemented in Brain Connectivity Toolbox [62] yields disjoint communities, i.e., one brain region is part of only one community, although recent studies in fMRI [36] demonstrate that one brain region may participate in multiple communities. Thus, we identify overlapping communities of ROIs. In the

Algorithm 2 LR-MVRC problem

Input: Data matrix $\mathbf{X}$ and parameters μ_1, μ_2.
Initialize: $\beta_1 = \beta_2 = 0.1$, $\beta_{max} = 10^{10}$, $\rho = 1.1$, $\mathbf{Y}_1 = \mathbf{Y}_2 = \tilde{\mathbf{W}} = \mathbf{Z}_1 = \mathbf{Z}_2 = \mathbf{0}$.
while convergence criteria not met **do**

1: Fix the other variables and update $\tilde{\mathbf{W}}$ by Eq (4.4)
2: Fix the other variables and update $\mathbf{Z}_1$ by Eq (4.6)
3: Fix the other variables and update $\mathbf{Z}_2$ by Eq (4.9)
4: Update the multipliers by

$$\mathbf{Y}_1 = \mathbf{Y}_1 + \beta_1(\mathbf{Z}_1 - \tilde{\mathbf{W}}), \mathbf{Y}_2 = \mathbf{Y}_2 + \beta_2(\mathbf{Z}_2 - \tilde{\mathbf{W}})$$

5: Update β_1=$min(\beta_{max}, \rho\beta_1)$ and β_2=$min(\beta_{max}, \rho\beta_2)$.

end while
Output: $\mathbf{A} = (|\tilde{\mathbf{W}}| + |\tilde{\mathbf{W}}|^T)/2$.

next section, we describe the method for detecting overlapping communities.

Detection of Overlapping Communities

We utilize NMF technique to obtain overlapping communities. NMF is a feature extraction and dimensionality reduction method of machine learning, which has been adapted to community detection recently [84, 85, 86]. This technique factorizes a non-negative input FC matrix $\mathbf{A}$, approximately, into a product of two non-negative matrices as:

$$\mathbf{A} \simeq \mathbf{PQ}. \tag{4.12}$$

The above factorization is carried out with a particular rank K so that $\mathbf{P}$ is of dimension $N \times K$ and $\mathbf{Q}$ is of dimension $K \times N$. This factorization can be viewed as a representation of data in a lower (K) dimensional space. In NMF, matrices $\mathbf{P}$ and $\mathbf{Q}$ are updated iteratively to improve the approximation to $\mathbf{A}$, while maintaining non-negative matrix entries throughout. For a given K, the algorithm runs iteratively until it finds a good approximate factorization or the stop criterion is met. For a symmetric data matrix $\mathbf{A}$, the factors $\mathbf{P}$ and $\mathbf{Q}$ can be considered as $\mathbf{Q} = \mathbf{P}^T$.

In particular, NMF algorithm minimizes the cost function, representing the approximation error between the actual data $\mathbf{A}$ and the reduced dimension recon-

struction of the data, i.e., $\mathbf{PQ}$. One of the popular cost function is the squared Euclidean distance as described below:

$$L(\mathbf{P}, \mathbf{Q}) = \frac{1}{2} \sum_i \sum_j (A_{ij} - [\mathbf{PQ}]_{ij})^2. \tag{4.13}$$

The minimization of the above cost function w.r.t. $\mathbf{P} \geq 0$ and $\mathbf{Q} \geq 0$ is a nonconvex problem. This approach can be interpreted as a maximum likelihood estimation with additional non-negativity constraints. However, the implementation details which are necessary to ensure decreasing cost function under the non-negativity constraints seem to be rather complex [87]. In addition, without any other constraint or prior to this problem, solution may lead to unstable convergence, and therefore, good initial values are necessary for more sophisticated NMF algorithms [88, 89]. Different iterative algorithms have been proposed to solve this problem [88]. In general, at each iteration of these algorithms, one matrix is considered to be fixed and the other one is estimated. This process continues until the convergence is achieved. The main idea behind these approaches is that by fixing one matrix, the estimation of the other matrix becomes a convex problem that can be solved easily.

Since NMF, in general, has no unique solution, it is necessary to introduce some additional constraints reflecting some prior knowledge. To this end, recently a Bayesian approach of NMF is proposed that imposes prior distributions on the matrices and leads to unique convergence to solution [90]. In recent overlapping communities detection work [84], it is assumed that the pairwise interaction A_{ij} is generated by a Poisson distribution with rate $\hat{A}_{ij} = \sum_k P_{ik} Q_{kj}$ and shrinkage prior on P_{ik} and Q_{kj} with hyper-parameter β_k. These additional constraints help to induce some kind of uniqueness to the NMF solution. Further, this Bayesian NMF approach maximizes posterior density under non-negative constraints. Keeping in view the success of this Bayesian NMF approach for overlapping communities detection, we have utilized algorithm presented in [84] to solve for NMF in this work. This algorithm starts with random initial matrices $\mathbf{P}$ and $\mathbf{Q}$, with non-negative weights chosen from a uniform random distribution on the interval [0, 1].

In brief, these matrices are updated iteratively as

$$\begin{aligned} \mathbf{Q} &\leftarrow (\frac{\mathbf{Q}}{\mathbf{P}^T\mathbf{1} + \mathbf{B}\mathbf{Q}}).\ [\mathbf{P}^T(\frac{\mathbf{A}}{\mathbf{P}\mathbf{Q}})], \\ \mathbf{P} &\leftarrow (\frac{\mathbf{P}}{\mathbf{1}\mathbf{Q}^T + \mathbf{P}\mathbf{B}})\ .\ [(\frac{\mathbf{A}}{\mathbf{P}\mathbf{Q}})\mathbf{Q}^T], \end{aligned} \tag{4.14}$$

where $\mathbf{B} \in \mathbb{R}^{K \times K}$ is a matrix with hyper-parameters β_k on the diagonal. For more details of this algorithm, one may refer to [84]. Code of this algorithm is available online at [1]. We have set the number of maximum iterations to 100 as has been done in [84]. This is to note that we ran the algorithm ten times with random initial matrices. The solution converged to the same output matrices with some columns exchanged.

Most experiments show that the matrix $\mathbf{P}$ represents the final clustering partition. In our work, we mainly utilized matrix $\mathbf{P}$ to determine overlapping communities of the input FC matrix. This is to note that each column of matrix $\mathbf{P}$ defines one community and hence, K denotes the total number of communities. Each column of $\mathbf{P}$ indicates the extent to which any brain region i belongs to that community. Thus, we compared each column of $\mathbf{P}$ with a threshold and if the value of region i is more than the threshold, region i is considered to be a part of that community. Similarly, if a region's value is less than the threshold, we delete that region from the corresponding community. Further details about selection of this threshold is provided in the Results Section.

Thresholding of each column of $\mathbf{P}$ independently allows brain regions to be a part of multiple communities simultaneously and hence, detects overlapping communities. The NMF approach of detecting communities is already used in the literature of social networks. However, their use in brain networks is limited so far.

Choosing Number of K Communities

NMF algorithm requires a given value of dimensionality K that is an important input used in matrix factorization. However, determining the value of K is a challenge in most community discovery algorithms because the number of communities

[1] `https://github.com/ipsorakis/commDetNMF`

K is not known in advance. If K is too small, some communities will be very large (in terms of more number of regions) with random grouping of ROIs. On the contrary, if K is too large, communities will be randomly scattered and will be very small.

To determine the number of communities K, we used NMF for two different numbers of communities 8 and 15. We chose 8 number of communities because, in general, the following eight brain networks are observed in the resting-state fMRI data: Visual Network (VN), Somato-motor Network (SMN), Auditory Network (AN), Cognitive Control Network (CCN), Bilateral Limbic Network (BLN), Language Network (LN), Subcortical Network (SCN), and Default Mode Network (DMN). We chose to extract 15 number of communities because we would like to ascertain if choosing a number larger than the expected number of communities yields scattered networks.

4.3 Results

We utilized $\mathbf{X}$ and computed adjacency matrix using eq. (4.1) for all subjects. We used $\mu_1 = 0.25$ in eq. (4.1) as was utilized previously in [68] and the value of μ_2 is decided empirically to be equal to 0.1. This value is decided based on the correspondence of identified fBNs with the ground truth atlas labels. We averaged LR-MVRC FC matrices of all subjects and utilized NMF to obtain communities.

Comparison of FC methods

In this section, we evaluate the feasibility and the robustness of the proposed LR-MVRC method w.r.t. traditional methods such as CORR, PCORR and MVRC method explained in chapter 3.

We averaged each method's FC matrices across all subjects and compared them w.r.t. various graph theoretical measures [91]. Degree was calculated as the sum of weighted edges connecting to a node i as $D_i = \sum_j \mathbf{A}_{ij}$. Similarly, participation coefficient (PC) arising from modularity community assignment was computed as $P_i = 1 - \sum_{s=1}^{N_m} (\frac{D_{is}}{D_i})^2$, where D_{is} is the number of edges of node i to nodes in module s, D_i is the degree of node i, and N_m is the total number of modules in

the graph. We also computed another graph theoretical measure, i.e., betweenness centrality (BC), averaged over all 90 AAL brain regions, whose high value signifies the network to be highly central or modular. In fact, it is a network centrality measure that represents the fraction of all shortest paths in the network that pass through a given node [62].

LR-MVRC matrix in this study is extracted using l^1 and low rank constraints, while MVRC FC matrix in [68] was extracted using l^1 and l^2 constraints. As a result, we noticed that while MVRC had more connections due to the denseness imposed by l^2 norm constraint, LR-MVRC had fewer higher magnitude connections compared to MVRC due to the low rank constraint. However, these fewer connections of LR-MVRC weighed higher than those of MVRC. In addition, we observed same degree in both the matrices, LR-MVRC FC and MVRC FC, since we did not apply any thresholding on FC matrices. Participation coefficients (of size 90 that is equal to the number of AAL brain regions) arising from the modularity community assignment using both the methods are statistically compared using the two-sample t-test with significance level of 0.05. Altered PC were observed between both the matrices ($p<0.05$). Further, we observed BC of LR-MVRC to be more (=128) compared to that of MVRC (=96). In conclusion, these findings indicate that the LR-MVRC FC matrix is more modular or central compared to MVRC FC.

Moreover, on comparing LR-MVRC method on simulated dataset mentioned in chapter 3, values of c-sensitivity is calculated for each individual subject separately and is averaged over all 50 subjects in all four datasets. Mean and variance of c-sensitivity values are observed to be 0.87 ± 0.13, 0.90 ± 0.05, 0.89 ± 0.07 and 0.92 ± 0.06 for all the four dataset respectively. Higher value of c-sensitivity compare to MVRC signifies better ability to identify connections. Hence, we can infer that the proposed LR-MVRC method consistently identifies presence of true connections.

Further, on comparing LR-MVRC method with the CORR and PCORR methods, we observed same degree with both the matrices, since we did not apply any thresholding on FC matrices. In addition, we observed PC using CORR and PCORR methods to be statistically different w.r.t. the proposed LR-MVRC method, suggesting that the FC matrices obtained by the three methods are sta-

tistically different ($p<0.05$). These results are consistent with the findings on LR-MVRC and MVRC methods as discussed previously, although this is to note that we obtained fewer number of communities using modularity with CORR and PCORR methods with no community signifying any functional brain network structure. Further details regarding this comparison is presented in the next section. Furthermore, analysis of the average BC showed lesser values among the CORR and PCORR methods compared to the proposed LR-MVRC method. These results indicated that there are structural differences among the networks constructed by the three methods.

Competing Methods

Modularity optimization is a leading method aimed at finding optimal non-overlapping partitions of fBNs based on modularity function [65]. Unlike NMF that finds overlapping communities, modularity optimization results in non-overlapping communities. We performed modularity optimization on the averaged LR-MVRC adjacency matrices implemented in the Brain Connectivity Toolbox [62].

Principal and independent component analysis (PCA, ICA) are some other matrix factorization methods commonly used for projecting data into lower dimensional representation, from which the overlapping communities structure could be identified [92]. These factorization methods impose different constraints or priors in order to obtain a solution, e.g., components must be orthogonal in PCA and must be independent in ICA. The ICA method is being extensively used in fMRI and it computes independent components directly on fMRI time-series and identifies overlapping communities. In this chapter, fMRI time-series of all subjects are concatenated temporally and the dimensionality is reduced by using PCA [93]. PCA is generally used before ICA for dimensionality reduction. The resulting fMRI time-series data is fed into the fastICA algorithm [94] implemented by the FastICA software (`https://research.ics.aalto.fi/ica/fastica/`).

To further test the performance of the proposed LR-MVRC FC identification method, all fBNs identification methods, i.e., NMF and modularity optimization are also carried out on the FC matrix constructed using the Pearson correlation (CORR). Specifically, for each pair of ROIs, Pearson correlation is used as a metric to compute the FC. Next, we averaged CORR adjacency matrices of all subjects

and utilized it to obtain communities.

Results on fMRI dataset

Beijing Zang resting-state fMRI data from 1000 Functional Connectome Project is utilized in this chapter. More details about this dataset is being provided in Chapter-3. Based on the analyses of LR-MVRC and CORR FC matrices, the overall community structure obtained by NMF and modularity are shown in Fig. 4.1. The matrices $\mathbf{P}$ obtained using LR-MVRC and CORR matrices with K=8 and 15 in NMF are shown in first two columns of Fig. 4.1. Visually, LR-MVRC matrices yield sparse overlapping community structures, while the ones derived using CORR matrices are also overlapping but are extremely dense. More sparsity, or in other words, fewer high valued coefficients in each column of matrix $\mathbf{P}$ results in accurate fBN identification by applying thresholding as described in the previous section. We considered mean plus standard deviation of each column of matrix $\mathbf{P}$ as a threshold value for that column or community. Modularity optimization method (refer to third column of Fig 4.1) achieves greatest sparsity among all methods, but this method results in non-overlapping communities. This is to note that this method does not require pre-defined number of communities unlike the NMF method. Using this method, CORR based FC matrix could identify only 4 communities compared to 10 communities detected with LR-MVRC based FC matrix.

In the second row of Fig 4.1, we show the community structure derived by ICA. The first and the second columns of this row display the overall community structures identified with K=8 and 15, respectively. For ICA method, group-level community structure is derived by concatenating fMRI time-series of all subjects. Here, the number of communities K for ICA methods were set according to the values chosen for NMF method in this chapter. Visually, the communities derived by ICA show an overlapping structure, but they tend to produce a much denser result with only a few high or negative coefficients values. It reveals that ICA method achieves a moderate similarity to NMF method since this method works directly on the time-series instead of first extracting FC matrix as is done in the NMF method.

To depict sparsity differences between NMF results of LR-MVRC, CORR and

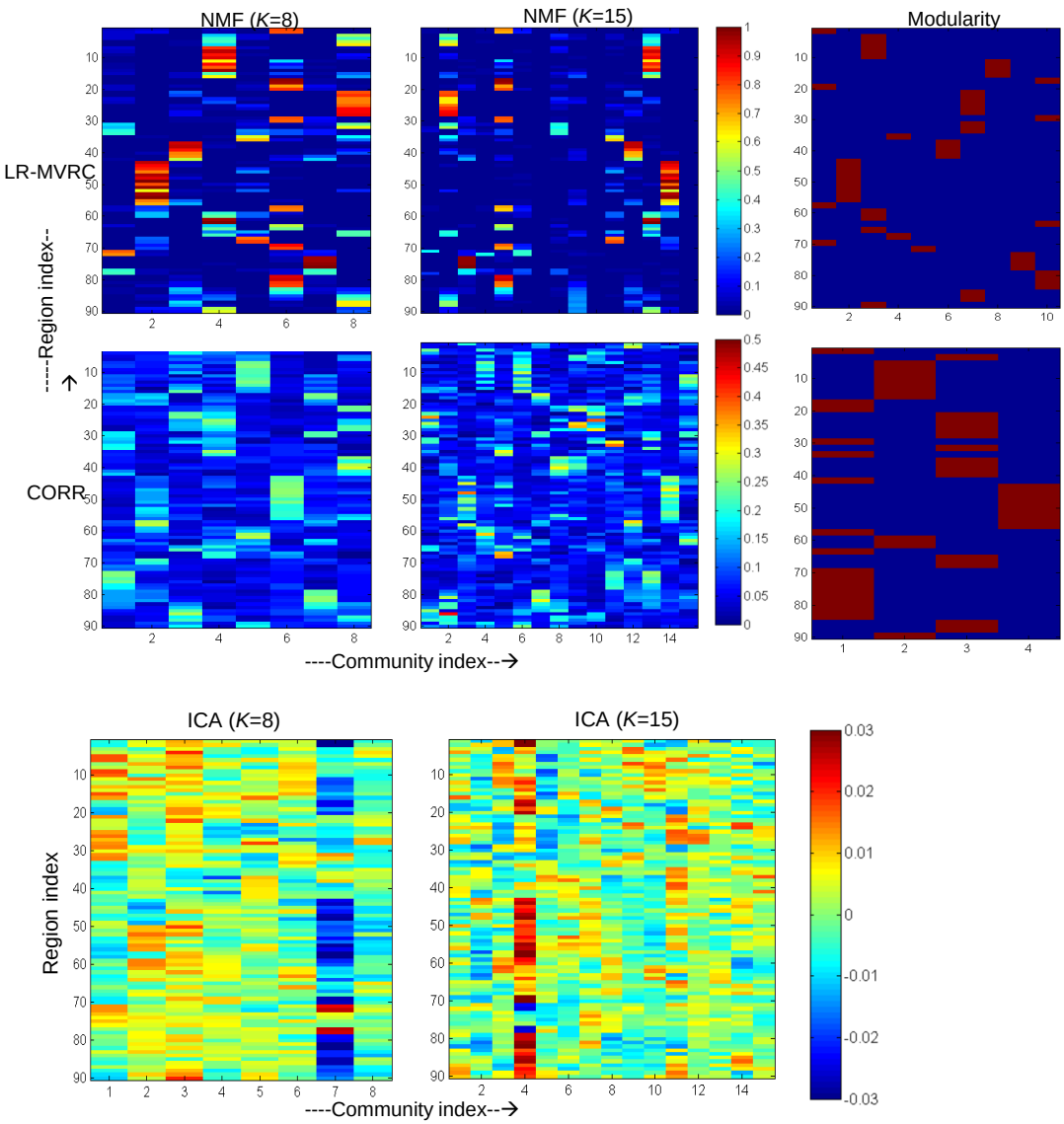


Figure 4.1: First row: Community structure obtained using LR-MVRC and CORR matrices with NMF and modularity optimization methods. Second row: Community structure obtained using ICA with K=*8* and 15 in the first and second columns, respectively. In the first row, first two columns represent the matrix $\mathbf{P}$ obtained from LR-MVRC and CORR with K=8 and 15 in NMF, respectively. The third column represents modularity optimization results on both LR-MVRC and CORR based adjacency matrices.

PCORR more accurately, we plot sorted coefficients of matrix $\mathbf{P}$ in Fig 4.2. We plot results for both 8 and 15 value of K in NMF. From this figure, we observe

that for each value of K, LR-MVRC based results are more sparse with only a few non zero coefficients compared to those extracted from CORR and PCORR FC matrices.

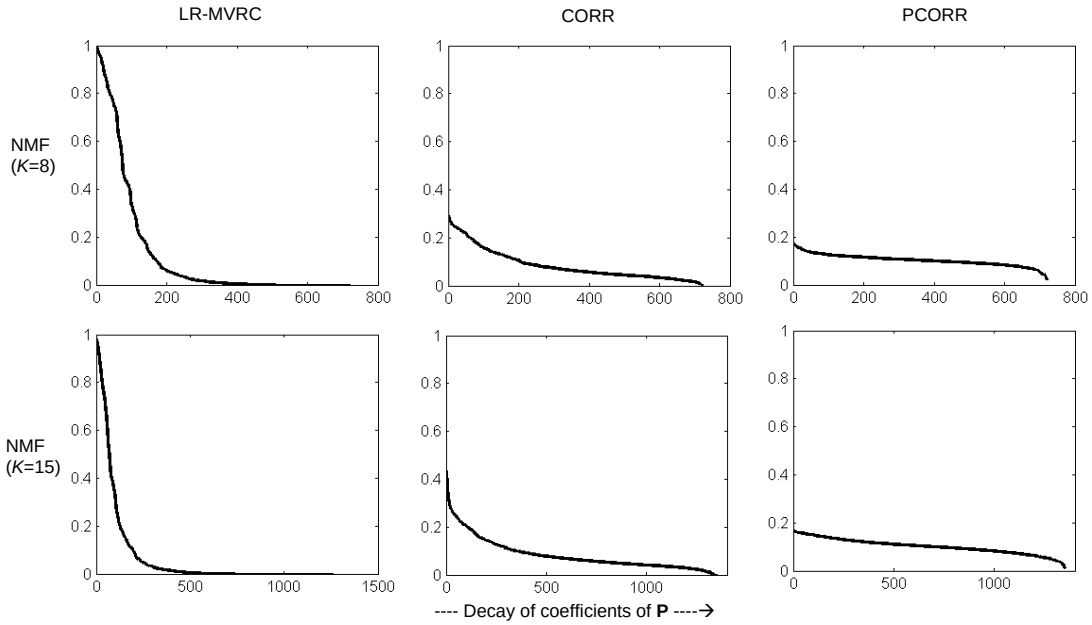


Figure 4.2: Illustration of sparsity of matrix $\mathbf{P}$ obtained with NMF. Plot of sorted coefficients of matrix $\mathbf{P}$, obtained using LR-MVRC, CORR and PCORR matrices with different values of K in NMF.

Further, to explore the neuro-physiological interpretations of identified communities with LR-MVRC (refer to Table 4.1) and CORR, we mapped the identified communities corresponding to K=8 onto the human brain, as shown in Figs. 4.3 and 4.4. Fig. 4.3 displays the communities derived by LR-MVRC based on the thresholding on matrix $\mathbf{P}$. Eight communities derived by LR-MVRC shown in this figure refer to default mode and subcortical (C1), visual (C2), bilateral limbic (C3), cognitive control and default mode (C4), default mode and visual (C5), auditory and motor (C6), subcortical (C7), and default mode and bilateral limbic (C8) networks. These networks are highly consistent with several well-recognized resting-state fBNs discovered by previous studies [95, 96, 56, 39]. In essence, two communities (C2, C5) are related to the visual information processing, one community (C6) is related to the auditory information processing, four communities (C1, C4, C5 and C8) are associated with the well-known default mode network

(DMN), and one (C6) is associated with the motor network. C1 and C7 correspond to the subcortical network. In addition, bilateral limbic network (C3 and C8) and cognitive control network (C4), associated with high-order brain functions, are readily identified. On the other hand, it is noted from Fig. 4.4 that ROIs associated with different fBNs are clubbed in a single community, showing random grouping of ROIs. This causes difficulty in the interpretation of identified communities as valid fBNs.

Table 4.1: **Brain Networks identified using LR-MVRC with K=8 in NMF. Fourth column represent AAL atlas ROI indices belonging to the community K, whereas second and third columns represent the associated brain networks and number of ROIs in that community, respectively.**

Community	**Networks**	**No. of nodes**	**AAL Regions**
1	default mode and subcortical	8	31, 32, 33, 34, 71, 72, 77, 78
2	visual	14	43-56
3	bilateral limbic	8	37-42, 83, 87
4	cognitive control and default mode	13	7-14, 16, 61, 62, 89, 99
5	default mode and visual	7	33, 35, 36, 46, 65, 67, 68
6	motor and auditory	17	1, 2, 17-20, 29, 30, 57, 58, 63, 69, 70, 79-82
7	subcortical	7	42, 73-78
8	default mode and bilateral limbic	17	3, 5, 6, 21-28, 31, 32, 65, 86-88

Non overlapping communities derived from LR-MVRC (see Fig. 4.5) refer to motor (C1), visual (C2), default mode and cognitive control (C3), default mode (C4), subcortical (C5), bilateral limbic (C6), default mode and bilateral limbic (C7), cognitive control (C8), subcortical (C9) and auditory networks, which are consistent with the overlapping communities detected with LR-MVRC method (see Fig. 4.3).

Apart from considering K=8 in NMF with LR-MVRC, we also present results with K=15 in Table 4.2. Fifteen communities derived by LR-MVRC shown here refer to default mode and subcortical (C1), default mode and bilateral limbic (C2), bilateral limbic and subcortical (C3), default mode and subcortical (C4), motor and auditory (C5), motor and subcortical (C6), default mode and subcor-

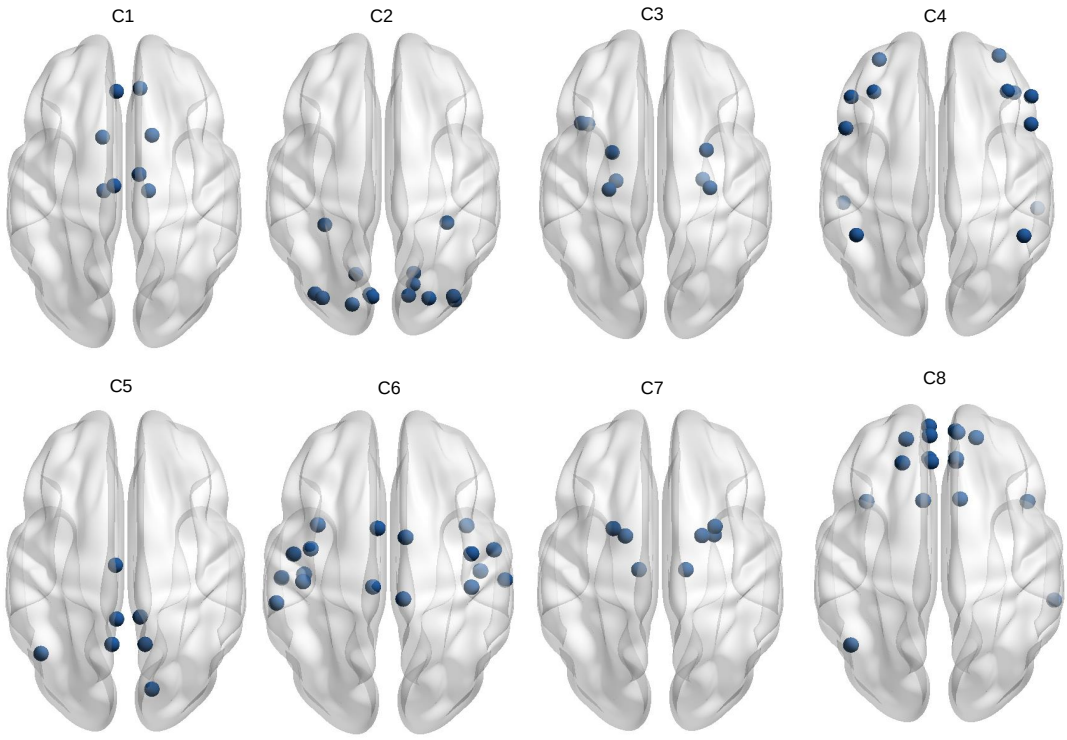


Figure 4.3: **Overlapping Communities derived from LR-MVRC.** This figure displays the communities derived from the LR-MVRC based adjacency matrix via thresholding of matrix **P** obtained with NMF (K=8). Eight communities derived from LR-MVRC matrix shown in this figure refer to default mode and subcortical (C1), visual (C2), bilateral limbic (C3), cognitive control and default mode (C4), default mode and visual (C5), auditory and motor (C6), subcortical (C7), and default mode and bilateral limbic (C8) networks.

tical (C7), default mode and subcortical (C8), coginitive control, bilateral limbic and default mode (C9), default mode and subcortical (C10), default mode (C11), bilateral limbic (C12), cognitive control and default mode (C13), visual (C14), and default mode and subcortical (C15) networks, which are consistent with several well-recognized fBNs discovered by previous studies. This is to note that the similarity observed in the four communities (C4, C7, C10 and C15) in Table 4.2 points to the fact that lesser value of K should have been used in the method. A value larger than required shows repeated communities, signifying that no better partitioning can happen. This observation can work as an initial criterion to decide the parameter K in the NMF method.

We further looked regions belonging to more than one community in LR-MVRC

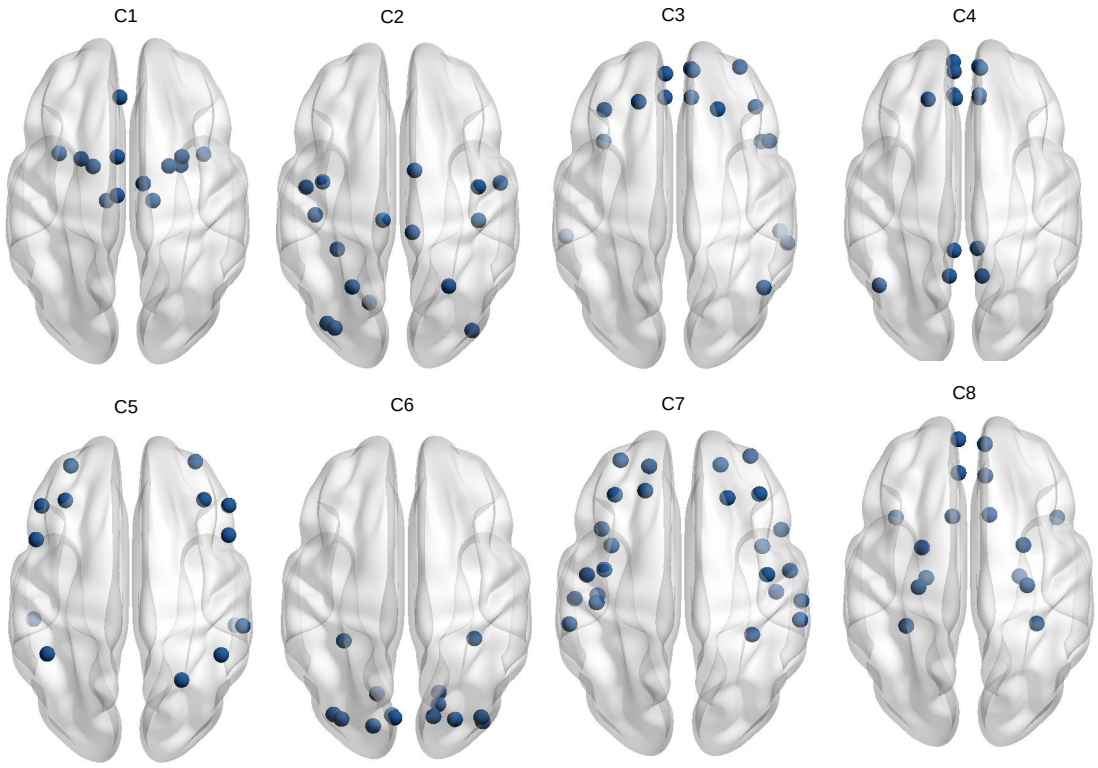


Figure 4.4: **Overlapping Communities derived from CORR.** This figure displays the communities derived using CORR based matrix via thresholding of matrix **P** obtained with NMF (K=8).

based NMF results. It reveals that regions involved in more than one community span over default mode, subcortical, bilateral limbic and visual networks. By contrast, regions related to sensori-motor and auditory participate in fewer communities.

Effect of Community Number *K*

Parameter K is the number of communities in the NMF method. On varying the value of K, different fBN structures/communities are observed. Increasing the number K allows us to represent more and more functional communities formed in the brain (refer to Table 4.2 with K=15). However, as the number of communities increases, we move from underfitting to overfitting community structures, i.e., we face the trade-off between approximating complex brain structures and overfitting them, potentially capturing noise or redundant brain networks.

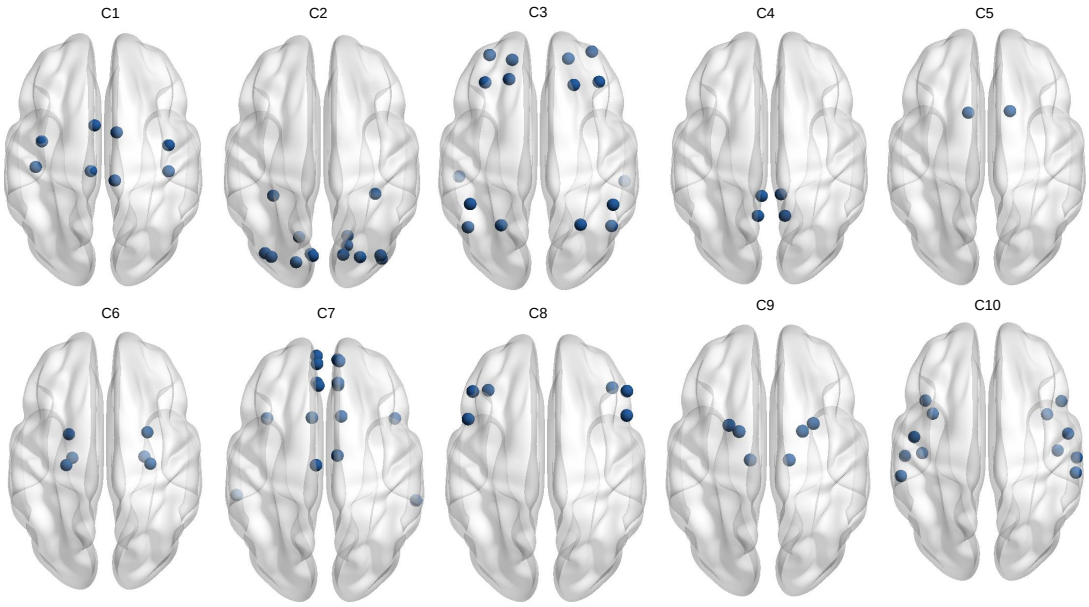


Figure 4.5: **Non Overlapping Communities derived from LR-MVRC.** This figure displays the communities derived using LR-MVRC based matrix via modularity optimization. Ten communities derived by LR-MVRC shown in this figure refer to motor (C1), visual (C2), default mode and cognitive control (C3), default mode (C4), subcortical (C5), bilateral limbic (C6), default mode and bilateral limbic (C7), cognitive control (C8), subcortical (C9) and auditory networks.

By starting with a large K (say 15, which is possibly double the number of communities in human brain), the effect of considering a higher number of communities can be accounted by ignoring redundant number of communities that correspond to similar fBNs. For example, four communities in Table 4.2 are identical and correspond to default mode and subcortical networks.

4.4 Conclusions

In this chapter, we propose a Low Rank Multivariate Vector Regression-based Connectivity (LR-MVRC) method for estimating FC matrix. Proposed method utilizes both sparsity and low rank constraints while estimating the FC matrix. Most of the previous studies extract disjoint communities or functional brain networks, we extract overlapping communities via NMF that may be biologically more correct for understanding human brain's functional organization. Experi-

Table 4.2: **Brain Networks identified using LR-MVRC with K=15 in NMF. Fourth column represent AAL atlas ROI indices belonging to the community K, whereas second and third columns represent their corresponding associated brain networks and number of ROIs in one community, respectively.**

Community	**Networks**	**No. of nodes**	**AAL Regions**
1	default mode and subcortical	3	33, 71, 78
2	default mode and bilateral limbic	15	3, 5, 6, 21-28, 31, 32, 87, 88
3	bilateral limbic and subcortical	7	42, 73-78
4	default mode and subcortical	3	34, 72, 77
5	motor and auditory	17	1, 2, 17-20, 29, 30, 57, 58, 63, 69, 70, 79-82
6	motor and subcortical	10	1, 2, 57, 58, 69-72, 77, 78
7	default mode and subcortical	3	34, 72, 77
8	default mode and subcortical	8	3, 23, 31-34, 77, 78
9	cognitive control, bilateral limbic and default mode	12	15, 16, 56, 66, 83-90
10	default mode and subcortical	3	34, 72, 77
11	default mode	6	35, 36, 46, 65, 67, 68
12	bilateral limbic	9	37-42
13	cognitive control and default mode	12	7-16, 61, 62, 90
14	Visual	14	43-56
15	default mode and subcortical	3	34, 72, 77

mental results suggest that the proposed framework can better characterize the brain networks' organization at the group level. In conclusion, we believe that the proposed method and its potential applications could provide new insights into the functional networks' organization of the human brain.

Chapter 5

Group-Fused Multivariate Regression Modeling for Group-Level Brain Networks

Availability of huge data, collected over multiple subjects, on publicly available 1000 functional Connectome Project (FCP) [97] and Human Connectome Project (HCP) [98], has made it possible to study the complex organization of fBNs, thereby, pressing the need to identify brain networks at the group-level. Recent works have identified large scale fBNs by analyzing a group of subjects [99, 100, 101].

Of various group-level fBN detection methods, the most widely used method carries out the averaging of connectivity matrices of all subjects [57, 91, 102, 103]. This approach is simple but has the limitation of not accounting the inter-subject variability [104]. Recently, in [105, 106], group-level fBNs are identified by using common single group-level matrix. In [106], group-level matrix is computed using joint diagonalization approach. In [105], group-level matrix is computed with variable weighting to each subject's connectivity matrix, which is further utilized to obtain group-level fBNs. Standard spectral clustering algorithm is applied to the group-level connectivity matrix in both the works above.

However, above methods compute connectivity matrices separately for each subject before combining them to obtain group-level connectivity matrix. In order to extract effective group-level fBNs, two issues must be addressed, namely, how to

obtain group-fused connectivity matrix results from all subjects' data simultaneously and second, how to combine these group-fused results to identify group-level fBNs. In this chapter, based on a newly proposed objective function, we include information of all subjects via group-fused penalty term. We also impose sparsity and denseness penalties on connectivity matrix coefficients as in Multivariate Vector Regression-based Connectivity (MVRC) method discussed in chapter 3 and name the proposed formulation as group-fused MVRC (GF-MVRC). The GF-MVRC method determines subject-level weights and group-embedded lower dimensional space iteratively by utilizing iterative spectral clustering. It is then extended to extract group-level fBNs. This approach not only identifies group-level networks, but also identifies the weighting of each subject's data in explaining group-level networks, i.e., the proposed work relies on determining subject-level weights responsible for group-level networks.

This work differs from the existing methods as follows: 1) Existing methods do not integrate the raw data of subjects, but focus on integrating the information at higher level only. Thus, their performance may degrade due to the heterogeneity associated with FC of each subject. On the other hand, this work computes group-fused FC using the raw data of all subjects in one framework; 2) The sparsity and denseness constraints imposed on each subject's FC explicitly, in our proposed formulation, consider the inter-fBNs sparsity and intra-fBNs denseness, leading to better network determination; and 3) we introduce weighted-subject framework via iterative spectral clustering that includes weight of each subject for extracting group-level networks. Thus, this approach not only identifies group-level fBNs, but also utilizes each subject's data appropriately for explaining these networks.

5.1 Proposed Formulation

GF-MVRC

Consider a group of S subjects with $\mathbf{X} = \left\{\mathbf{X}^s \in \mathbb{R}^{T\times N}\right\}_1^S$ set of matrices. To compute group-fused adjacency matrix $\mathbf{A}^s \in \mathbb{R}^{N\times N}$, we present the proposed GF-MVRC algorithm with l^1 and l^2 regularization terms on connectivity matrices as used in the MVRC method. Our objective is to find a group-fused adjacency

matrices $\mathbf{A}^s$ that balances the agreement on regional FC across different subjects, while at the same time promotes sparsity (via l^1-norm) and denseness (via l^2-norm) in the solution of connectivity matrices. We formulate a novel objective function that forces adjacency matrices, across different subjects via regularization, towards a common consensus as

$$\begin{aligned} \underset{\tilde{\mathbf{W}}^1,\tilde{\mathbf{W}}^2,...,\tilde{\mathbf{W}}^S}{min} \sum_{s=1}^{S} \Bigg(\frac{1}{2} \left\| \mathbf{X}^s - \mathbf{X}^s\tilde{\mathbf{W}}^s \right\|_F^2 + \mu_1 \left\| \tilde{\mathbf{W}}^s \right\|_1 + \\ \mu_2 \left\| \tilde{\mathbf{W}}^s \right\|_F^2 \Bigg) + \sum_{1\leq s,v\leq S, s\neq v} \lambda^s \left\| \tilde{\mathbf{W}}^s - \tilde{\mathbf{W}}^v \right\|_F^2 \\ s.t. \quad diag(\tilde{\mathbf{W}}^s) = 0, \quad s = 1, ..., S \end{aligned} \tag{5.1}$$

where μ_1, μ_2, and λ^s are non-negative regularization parameters. These parameters define the trade-off between sparsity and denseness constraints and the agreement across different subjects, respectively. First three terms in the above equation represents data regularization term, l^1-norm on adjacency matrix and l^2-norm on adjacency matrix respectively. This is to note that the proposed method considers group-fused regularization term (fourth term in above equation) in order to extract group-level brain networks. If we remove this term, the above optimization problem reduces to MVRC on single subject. The last term in the above objective function prevents self loops by making diagonal terms to zero. Since regularization term in (5.1) is non-differentiable, it is not easy to solve the above GF-MVRC formulation. Thus, (5.1) is reformulated in the next section that helps in solving this problem.

Implementation

In this subsection, we design an algorithm to solve (5.1) using the ADMM [71]. With all but one $\tilde{\mathbf{W}}^s$ fixed, (5.1) can be written as:

$$\begin{aligned} \underset{\tilde{\mathbf{W}}^s}{min} \frac{1}{2} \left\| \mathbf{X}^s - \mathbf{X}^s\tilde{\mathbf{W}}^s \right\|_F^2 + \mu_1 \left\| \tilde{\mathbf{W}}^s \right\|_1 + \mu_2 \left\| \tilde{\mathbf{W}}^s \right\|_F^2 \\ +\lambda^s \sum_{1\leq v\leq S, s\neq v} \left\| \tilde{\mathbf{W}}^s - \tilde{\mathbf{W}}^v \right\|_F^2 \quad s.t. \quad diag(\tilde{\mathbf{W}}^s) = 0. \end{aligned} \tag{5.2}$$

Following [71], three auxiliary matrices $\tilde{\mathbf{W}}_1$, $\tilde{\mathbf{W}}_2$, and $\tilde{\mathbf{W}}_3$ are introduced in

(5.2) as

$$\begin{aligned}\min_{\tilde{\mathbf{W}}^s,\tilde{\mathbf{W}}_1^s,\tilde{\mathbf{W}}_2^s,\tilde{\mathbf{W}}_3^s} \frac{1}{2}\left\|\mathbf{X}^s-\mathbf{X}^s\tilde{\mathbf{W}}^s\right\|_F^2+\mu_1\left\|\tilde{\mathbf{W}}_1^s\right\|_1+\mu_2\left\|\tilde{\mathbf{W}}_2^s\right\|_F^2 \\ +\lambda^s\sum_{1\le v\le S,s\ne v}\left\|\tilde{\mathbf{W}}_3^s-\tilde{\mathbf{W}}^v\right\|_F^2 \\ s.t.\quad diag(\tilde{\mathbf{W}}^s)=0,\tilde{\mathbf{W}}_1^s=\tilde{\mathbf{W}}^s,\tilde{\mathbf{W}}_2^s=\tilde{\mathbf{W}}^s,\tilde{\mathbf{W}}_3^s=\tilde{\mathbf{W}}^s.\end{aligned} \tag{5.3}$$

The augmented Lagrangian of the above equation is:

$$\begin{aligned}\mathcal{L}(\tilde{\mathbf{W}}^s,\left\{\tilde{\mathbf{W}}_i^s\right\}_{i=1}^3,\{\mathbf{Y}_i^s\}_{i=1}^3)=\frac{1}{2}\left\|\mathbf{X}^s-\mathbf{X}^s\tilde{\mathbf{W}}^s\right\|_F^2+ \\ \mu_1\left\|\tilde{\mathbf{W}}_1^s\right\|_1+\mu_2\left\|\tilde{\mathbf{W}}_2^s\right\|_F^2+\lambda^s\sum_{1\le v\le S,s\ne v}\left\|\tilde{\mathbf{W}}_3^s-\tilde{\mathbf{W}}^v\right\|_F^2 \\ +\frac{\beta_1}{2}\left\|\tilde{\mathbf{W}}^s-\tilde{\mathbf{W}}_1^s+\frac{\mathbf{Y}_1}{\beta_1}\right\|_F^2+\frac{\beta_2}{2}\left\|\tilde{\mathbf{W}}^s-\tilde{\mathbf{W}}_2^s+\frac{\mathbf{Y}_2}{\beta_2}\right\|_F^2 \\ +\frac{\beta_3}{2}\left\|\tilde{\mathbf{W}}^s-\tilde{\mathbf{W}}_3^s+\frac{\mathbf{Y}_3}{\beta_3}\right\|_F^2,\quad s.t.\quad diag(\tilde{\mathbf{W}}^s)=0,\end{aligned} \tag{5.4}$$

where $\{\beta_i>0\}_{i=1}^3$ are the penalty parameters that need to be tuned and $\{\mathbf{Y}_i>0\}_{i=1}^3$ are Lagrange dual variables used to enforce equality between the original and auxiliary matrices. ADMM updates variables in the above equation alternately. The minimization over one variable in an iteration assumes the other variables to be fixed. Therefore, the above function can be alternately optimized over each variable separately. This allows splitting of (5.4) into different subproblems with new objective functions stated as below:

$$\begin{aligned}P1:\quad arg\ \min_{\tilde{\mathbf{W}}^s}\ \frac{1}{2}\left\|\mathbf{X}^s-\mathbf{X}^s\tilde{\mathbf{W}}^s\right\|_F^2 \\ +\frac{\beta_1}{2}\left\|\tilde{\mathbf{W}}^s-\tilde{\mathbf{W}}_1^s+\frac{\mathbf{Y}_1}{\beta_1}\right\|_F^2+\frac{\beta_2}{2}\left\|\tilde{\mathbf{W}}^s-\tilde{\mathbf{W}}_2^s+\frac{\mathbf{Y}_2}{\beta_2}\right\|_F^2 \\ +\frac{\beta_3}{2}\left\|\tilde{\mathbf{W}}^s-\tilde{\mathbf{W}}_3^s+\frac{\mathbf{Y}_3}{\beta_3}\right\|_F^2,\quad s.t.\quad diag(\tilde{\mathbf{W}}^s)=0.\end{aligned} \tag{5.5}$$

$$P2:\quad arg\ \min_{\tilde{\mathbf{W}}_1^s}\ \mu_1\left\|\tilde{\mathbf{W}}_1^s\right\|_1+\frac{\beta_1}{2}\left\|\tilde{\mathbf{W}}^s-\tilde{\mathbf{W}}_1^s+\frac{\mathbf{Y}_1}{\beta_1}\right\|_F^2. \tag{5.6}$$

$$P3: \quad arg \min_{\tilde{\mathbf{W}}_2^s} \ \mu_2 \left\| \tilde{\mathbf{W}}_2^s \right\|_F^2 + \frac{\beta_2}{2} \left\| \tilde{\mathbf{W}}^s - \tilde{\mathbf{W}}_2^s + \frac{\mathbf{Y}_2}{\beta_2} \right\|_F^2. \tag{5.7}$$

$$\begin{aligned} P4: \quad arg \min_{\tilde{\mathbf{W}}_3^s} \ \lambda^s \sum_{1 \le v \le S, s \ne v} & \left\| \tilde{\mathbf{W}}_3^s - \tilde{\mathbf{W}}^v \right\|_F^2 + \\ & \frac{\beta_3}{2} \left\| \tilde{\mathbf{W}}^s - \tilde{\mathbf{W}}_3^s + \frac{\mathbf{Y}_3}{\beta_3} \right\|_F^2. \end{aligned} \tag{5.8}$$

Above four subproblems are solved iteratively along with the update of Lagrange variables. The complete algorithm is summarized in Algorithm 3, while the solution of each subproblem is explained in the next few subsections.

Algorithm 3 GF-MVRC

Input: $\mathbf{X} = \left\{ \mathbf{X}^s \in \mathbb{R}^{T \times N} \right\}_1^S$, μ_1, μ_2, $\{\lambda^s\}_1^S$,$\{\beta_i\}_1^3$, β^{max}, ρ

Output: $\left\{ \tilde{\mathbf{W}}^s \in \mathbb{R}^{N \times N} \right\}_1^S$, $\left\{ \mathbf{A}^s \in \mathbb{R}^{N \times N} \right\}_1^S$

Intialize:$\{\mathbf{W}_i^s = \mathbf{0}\}_{i=1}^3$, $\{\mathbf{Y}_i^s = \mathbf{0}\}_{i=1}^3$
while convergence criteria not met **do**
    for $s = 1$ to S do
        Fix others and update $\tilde{\mathbf{W}}^s$ by solving (5.5)
        Fix others and update $\mathbf{W}_1^s$ by solving (5.6)
        Fix others and update $\mathbf{W}_2^s$ by solving (5.7)
        Fix others and update $\mathbf{W}_3^s$ by solving (5.8)
        Fix others and update Lagrange dual variables by

$$\begin{aligned} \mathbf{Y}_1^s &= \mathbf{Y}_1^s + \beta_1(\tilde{\mathbf{W}}_1^s - \tilde{\mathbf{W}}^s) \\ \mathbf{Y}_2^s &= \mathbf{Y}_2^s + \beta_2(\tilde{\mathbf{W}}_2^s - \tilde{\mathbf{W}}^s) \\ \mathbf{Y}_3^s &= \mathbf{Y}_3^s + \beta_3(\tilde{\mathbf{W}}_3^s - \tilde{\mathbf{W}}^s) \end{aligned}$$

    end for
    Update $\beta_i = min(\rho\beta_i, \beta^{max})$, $i = 1, \cdots, 3$
end while

P1 Subproblem The update rule of $\tilde{\mathbf{W}}^s$ is obtained directly by setting the partial derivative of problem $P1$ to zero. $\tilde{\mathbf{W}}^s$ is updated using the following update rule:

$$\begin{aligned} \tilde{\mathbf{W}}^s = [(\mathbf{X}^s)^\intercal \mathbf{X}^s + (\beta_1 + \beta_2 + \beta_3)\, \mathbf{I}]^{-1} \times \\ ((\mathbf{X}^s)^\intercal \mathbf{X}^s + \beta_1 \mathbf{W}_1^s + \beta_2 \mathbf{W}_2^s + \beta_3 \mathbf{W}_3^s - \mathbf{Y}_1^s - \mathbf{Y}_2^s - \mathbf{Y}_3^s), \end{aligned} \tag{5.9}$$

where $\mathbf{I}$ denotes an identity matrix. The matrix $diag(\tilde{\mathbf{W}}^s)$ is kept to zero in (5.9)

in every iteration after solving for $\tilde{\mathbf{W}}^s$.

P2 Subproblem This subproblem is a l^1 minimization problem. The minimization of (5.6) provides the following update rule for matrix $\tilde{\mathbf{W}}_1^s$ [107]:

$$\mathbf{W}_1^s = Soft(\tilde{\mathbf{W}}^s + \frac{\mathbf{Y}_1}{\beta_1}, 2\frac{\mu_1}{\beta_1}), \tag{5.10}$$

where '*Soft*' denotes soft thresholding operator applied entry wise to $(\tilde{\mathbf{W}}^s + \frac{\mathbf{Y}_1}{\beta_1})$.

P3 Subproblem The update rule of $\tilde{\mathbf{W}}_2^s$ directly obtained by setting partial derivative of problem $P3$ to zero. $\tilde{\mathbf{W}}_2^s$ is updated by the following update rule:

$$\tilde{\mathbf{W}}_2^s = [(\mu_2 + \beta_2)\,\mathbf{I}]^{-1} \times \left(\beta_2 \tilde{\mathbf{W}}^s + \mathbf{Y}_2\right) \tag{5.11}$$

P4 Subproblem The update rule of $\tilde{\mathbf{W}}_3^s$ directly obtained by setting partial derivative of problem $P4$ to zero as follows:

$$\begin{aligned}\frac{\partial(P4)}{\partial \tilde{\mathbf{W}}_3^s} = [2\lambda^s(S-1) + \beta_3]\tilde{\mathbf{W}}_3^s - 2\lambda^s \sum_{1\le v\le S, s\neq v} \tilde{\mathbf{W}}^v \\ -\beta_3 \tilde{\mathbf{W}}^s - \mathbf{Y}_3\end{aligned} \tag{5.12}$$

Setting the partial derivative in (5.12) to zero, $\tilde{\mathbf{W}}_3^s$ is updated as

$$\begin{aligned}\tilde{\mathbf{W}}_3^s = [2\lambda^s(S-1) + \beta_3]^{-1}(2\lambda^s \sum_{1\le v\le S, s\neq v} \tilde{\mathbf{W}}^v \\ +\beta_3 \tilde{\mathbf{W}}^s + \mathbf{Y}_3).\end{aligned} \tag{5.13}$$

After solving for above four subproblems iteratively, the symmetric adjacency matrix of each subject is computed by $\mathbf{A}^s = (\left|\tilde{\mathbf{W}}^s\right| + \left|\tilde{\mathbf{W}}^s\right|^{\intercal})/2$. In algorithm 3, iterations are stopped, if the following conditions are satisfied in iteration j: $\left\|\tilde{\mathbf{W}}^s - \tilde{\mathbf{W}}_1^s\right\|_\infty \le \varepsilon$, $\left\|\tilde{\mathbf{W}}^s - \tilde{\mathbf{W}}_2^s\right\|_\infty \le \varepsilon$, $\left\|\tilde{\mathbf{W}}^s - \tilde{\mathbf{W}}_3^s\right\|_\infty \le \varepsilon$, and $\left\|\tilde{\mathbf{W}}_j^s - \tilde{\mathbf{W}}_{j-1}^s\right\|_\infty \le \varepsilon$ for $s = 1, 2, \cdots, S$, or with the maximum number of iterations reached. We use the same initial values of $\{\lambda^s\}_1^S$,$\{\beta_i\}_1^3$, β^{max}, and ρ for all S subjects because this leads to a significantly reduced number of variables for optimization. However, $\{\beta_i\}_1^3$ are updated after completion of one iteration of all subjects.

Choice of parameters: For the cases, where we do not have a prior information

about which subject is more important than the others, λ^s is not dependent on subject s. Therefore, the same value of $\lambda^s = 0.1$ is used for all subjects $s = 1, \cdots, S$. This value is decided empirically based on the identified group-fused brain networks. Values of μ_1 and μ_2 are utilized as decided in [68].

5.2 Functional Brain Network Identification

Iterative Spectral Clustering

After computing group-fused adjacency matrices of all subjects, we determine lower dimensional subspace $\mathbf{U} \in \mathbb{R}^{N \times k}$ shared by all subjects by using spectral clustering with trace optimization problem as

$$\underset{\mathbf{U} \in \mathbb{R}^{N \times k}}{min}\ tr(\mathbf{U}^T \mathbf{L} \mathbf{U}) \quad s.t.\ \mathbf{U}^T \mathbf{U} = \mathbf{I}, \tag{5.14}$$

where $\mathbf{L} = \mathbf{I} - \mathbf{A}_n$ denotes the normalized graph Laplacian matrix and $\mathbf{A}_n$ denotes the normalized undirected adjacency matrix ($\mathbf{D}^{\frac{-1}{2}} \mathbf{A} \mathbf{D}^{\frac{-1}{2}}$) with $\mathbf{D} = \sum_i^N \mathbf{A}(i)$ as the diagonal degree matrix. N represents the number of ROIs and k denotes the number of communities (or brain networks). The solution of (5.14), according to Ky Fan theorem [108], is obtained by choosing k eigenvectors corresponding to the lowest eigenvalues of $\mathbf{L}$ or the highest eigenvalues of $\mathbf{A}_n$. Problem in (5.14) can also be written as [109]:

$$\underset{\mathbf{U} \in \mathbb{R}^{N \times k}}{max}\ tr(\mathbf{U}^T \mathbf{A}_n \mathbf{U}) \quad s.t.\ \mathbf{U}^T \mathbf{U} = \mathbf{I}. \tag{5.15}$$

After estimating lower dimensional space $\mathbf{U}$, communities are determined by applying K-means clustering to the matrix $\mathbf{U}$. The optimization problem in (5.15) for one subject can be extended to identify group-fused lower dimensional space using weighted combination of adjacency matrices of all subjects as below:

$$\begin{aligned} \underset{\mathbf{U}, \mathbf{w}}{max}\ \sum_{s=1}^{S} w^s\ tr(\mathbf{U}^T \mathbf{A}_n^s \mathbf{U}) = tr(\mathbf{U}^T (\sum_{s=1}^{S} w^s \mathbf{A}_n^s) \mathbf{U}) \\ s.t.\ \mathbf{U}^T \mathbf{U} = \mathbf{I}, (w^s)_1^S \geq 0, \|\mathbf{w}\|_F = 1, \end{aligned} \tag{5.16}$$

where $\mathbf{w} \in \mathbb{R}^{S \times 1}$ represents weights of all subjects for determining the com-

mon lower dimensional space $\mathbf{U}$. Equal weights to all subjects can be provided in the lower dimensional space $\mathbf{U}$ by setting w^s to one for all subjects in (5.16). However, this might not be true in real scenario as fBNs change from one subject to another. Therefore, we consider a general case of unequal weightage to all subjects. This results into two unknowns in (5.16), $\mathbf{U}$ and $\mathbf{w}$. We iteratively solve for these two variables and hence, name the proposed method as iterative spectral clustering. Notably, column space of $\mathbf{U}$ in (5.16) represents dominant eigenspace of $\sum_{s=1}^{S} w^s \mathbf{A}_n^s$. Above formulation of identifying common lower dimensional space $\mathbf{U}$ has also been utilized previously for the case of multiview clustering in text-based data [109]. The pseudo code for the solution of (5.16) is reproduced in Algorithm 4 as provided in [109]. Here, $P(\mathbf{U})$ is denoted as $(trace(\mathbf{U}^T\mathbf{A}_n^1\mathbf{U}), trace(\mathbf{U}^T\mathbf{A}_n^2\mathbf{U}), \cdots, trace(\mathbf{U}^T\mathbf{A}_n^S\mathbf{U}))$ and $\mathbf{U}$ is initialized to dominant eigenspace of averaged matrix $\sum_{s=1}^{S} \mathbf{A}_n^s$. In the end, we apply K-means clustering on matrix $\mathbf{U}$ to identify fBNs common to the group of subjects.

Algorithm 4 Iterative Spectral Clustering

Input: $\left\{\mathbf{A}_n^s \in \mathbb{R}^{N\times N}\right\}_1^S$, k
Output: $\mathbf{U} \in \mathbb{R}^{N\times k}, \mathbf{w}$

```
while convergence criteria not met do
    Obtain P(U)
    Calculate w by scaling P(U) to unit-norm
    Compute Σ_{s=1}^{S} w^s A_n^s and obtain U as dominant eigenspace of this matrix
end while
```

5.3 Results

Beijing Zang resting state fMRI data from 1000 Functional Connectome Project is utilized in this chapter to validate the proposed framework. More details about this dataset has been provided in Chapter-3. Experimental results indicate that the proposed GF-MVRC method outperforms the existing state-of-the-art group-level identification methods [105, 106]. The contribution of different subjects is also identified effectively in the proposed work. Experiments demonstrate that the proposed method has good performance in discovering group level functional brain networks compared to the state-of-the-art methods.

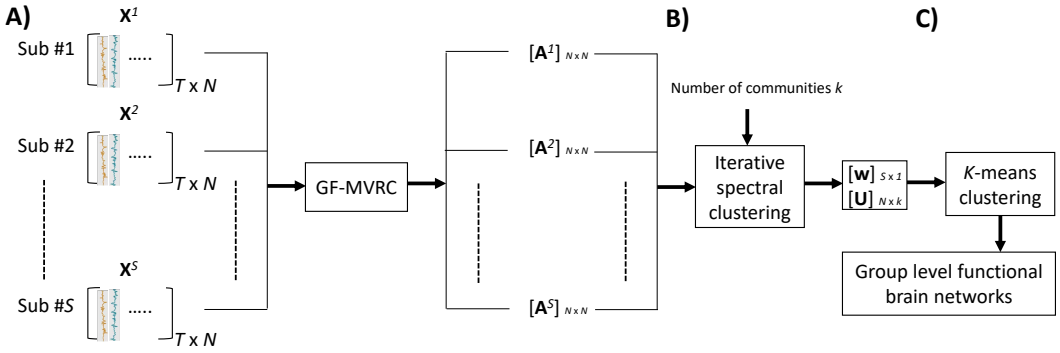


Figure 5.1: A schematic diagram of the proposed method to find Group level fBNs A) GF-MVRC is utilized to identify group fused functional connectivity $\{\mathbf{A}^s \in \mathbb{R}^{N\times N}\}_1^S$ across all S subjects, where N denotes the number of region of intersts (90 for the case of AAL atlas) and T denotes number of time points; B) Iterative spectral clustering is utilized to identify lower dimensional space $\mathbf{U} \in \mathbb{R}^{N\times k}$ and the corresponding weighing vector $\mathbf{w} \in \mathbb{R}^{S\times 1}$ of each subject; C) K-means clustering is used on $\mathbf{U}$ to identify group level functional brain networks.

Methods of Comparison

We compare the proposed GF-MVRC method with several existing group-level fBNs identification methods. Existing methods compute FC from each subject separately, which are later utilized to estimate group-embedded lower-dimensional matrix. These methods adopt different weighting approach of combining subject-level adjacency matrices.

1. **Multiple View Spectral Clustering (MVSC) [105]**: This method assigns equal weights to the adjacency matrices of all subjects and hence, computes group-embedded adjacency matrix via averaging over all subjects.

2. **Multiple View Spectral Clustering with Weights (MVSCW) [105]**: This method first computes the Laplacian matrix corresponding to each subject's adjacency matrix and chooses weights corresponding to k smallest eigenvalues $\lambda_1, \cdots, \lambda_k$ as

$$w_s = \frac{[\sum_{i=1}^{k} \lambda_i^{(s)}]^{-1}}{\sum_{l=1}^{S}[\sum_{i=1}^{k} \lambda_i^{(l)}]^{-1}}, \quad s = 1, \cdots, S, \tag{5.17}$$

where S here denotes number of subjects.

3. **Affinity Aggregation for Spectral Clustering (AASC) [110]**: This method determines weights in an optimization framework using line search technique.

4. **Joint Diagonalization of Laplacians (JDL) [106]**: This method identifies an orthogonal matrix that approximates embedding of all subjects' normalized Laplacian matrices.

For generating results with the above existing methods, we utilize Pearson correlation to build each subject's adjacency matrix as was used previously in [105]. Correlation values are Fisher z-transformed and negative coefficients are set to zero [105]. We utilize the code provided by [105] for the implementation of these existing methods.

Number of communities k

The proposed GF-MVRC method and existing methods require the specification of the number of communities k. Since there is no apriori knowledge of determining the number of communities, we identify the number k using modularity [65]. Modularity is a largely used method to extract fBNs from single subject. This method is based on maximizing the modularity quality function Q and does not require number of communities k at the input. Therefore, we utilized modularity implemented in Brain Connectivity Toolbox to capture the number of communities of each subject and finally, considered k as a mean value of the number of communities determined for each subject.

Experimental results using the proposed GF-MVRC on 116 female and 74 male participants are shown in Fig. 5.2. We utilized implementation[1] from [111] to visualize brain networks on slice grid. We could identify 8 and 7 number of fBNs in female and male subgroups, respectively. These extracted communities with K-means clustering closely resemble to different brain networks available in AAL atlas. These includes visual, auditory, default mode, bilateral limbic, subcortical, cognitive control networks, etc.

[1] `https://github.com/premgopalan/svinet`

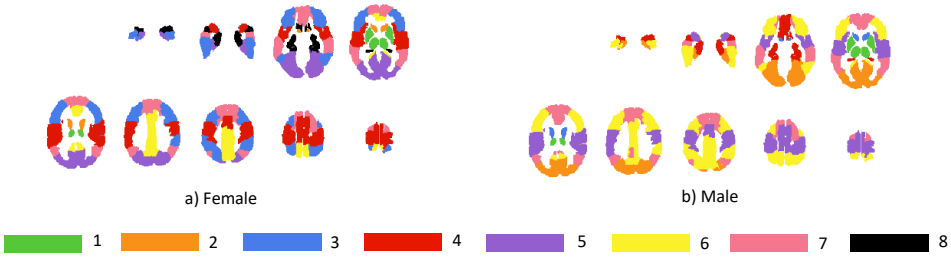


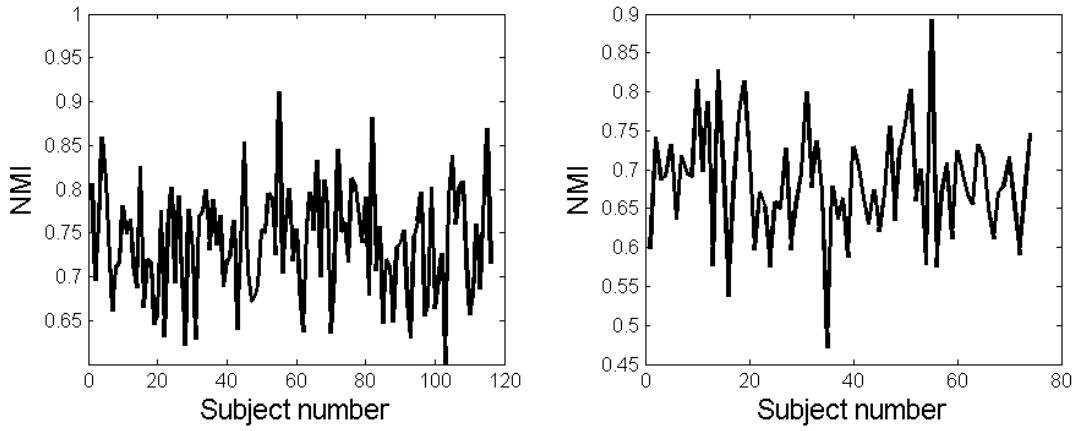
Figure 5.2: Group-wise fBNs using proposed GF-MVRC in A) female subgroup (with k=8) and B) male subgroup (with k=7). Communities 1 and 2 of female subgroup corresponds to SCN, 3 corresponds to CCN, DMN and VN, 4 corresponds to AN, MN and BLN, 5 corresponds to VN, 6 corresponds to DMN, 7 corresponds to DMN with one LN related region, 8 corresponds to BLN; B) male subgroup (with k=7): communities 1 and 3 corresponds to SCN, 2 corresponds to VN, 4 corresponds to BLN, 5 corresponds to AN, 6 corresponds to CCN, DMN, VN and LN, 7 corresponds to DMN with one LN related region. Abbreviations: VN- Visual Network; SMN- Somato-Motor Network; AN- Auditory Network; CCN- Cognitive Control Network; DMN- Default Mode Network; BLN - Bilateral Limbic Network; SCN - Subcortical Network; LN - Language Network.

Assessment of individual subject's community structures

In order to check community-wise variability across individuals, we identified individual subject's communities using equ.(5.14) for all subjects. To quantify similarity between the group-wise and subject-level community structures, Normalized Mutual Information (NMI) [112, 113, 114] is computed for every pair of the group-based community and individual subject's communities. NMI is one of the most popular similarity measure used to compute similarity between two communities. The basic idea behind NMI is that if two communities are similar, less information is necessary to infer one from the other community structure [112]. If the similarity between communities is high, NMI is also high.

The NMI was calculated for every pair of the group-based community structure and the individual's communities (116 subjects of female and 74 subjects of males). To evaluate the NMI performance, we set the number of clusters for both individual's and group data to be the same, i.e. K=8 and 7 for female and male groups, respectively. The obtained NMI values for each individual are depicted in Fig.5.3 for both the groups. The averaged NMI values and the standard deviation were observed to be 0.7379±0.0621 and 0.6821±0.0754 for the female and male subgroups, respectively. A high value of averaged NMI indicates consistency

in communities structure between the individuals' communities to those of the respective subgroup's communities identified at the group-level. The number of subjects showing greater similarity with the group-based community is observed to be 58 and 36 for the female and male subgroups, respectively. This high value of similarity of individuals communities to the group-level communities pushes the need for identifying group-level networks instead of individual networks for better human brain understanding. Analyzing individuals' communities may be less meaningful for studies on large neuroimaging data, when population specific findings are to be drawn. Instead, it may be more useful to identify common group-based community structures.



(a) (b)

Figure 5.3: Individual to group level community assessment using NMI a) female and b) male subgroups, respectively.

All the existing methods yield four communities for both the subgroups (see Figs. 5.4 and 5.5), whereas the proposed method results in 8 and 7 communities for female and male subgroups, respectively (see Fig. 5.2). From Fig. 5.2, we note that the proposed GF-MVRC method identifies multiple fBNs including Visual Network (VN), Somatomotor Network (SMN), Auditory Network (AN), Cognitive Control Network (CCN), Language Network (LN), Subcortical Network (SCN), and Default Mode Network (DMN). We observe that our method significantly outperforms existing methods including MVSC, MVSCW, AASC, and JDL. These methods are not able to clearly separate various brain networks. Since these methods compute adjacency matrices independently, it might have led to the loss

of potential information about group-level fBNs. On the other hand, our method is capable of finding "genuine" ground truth fBNs.

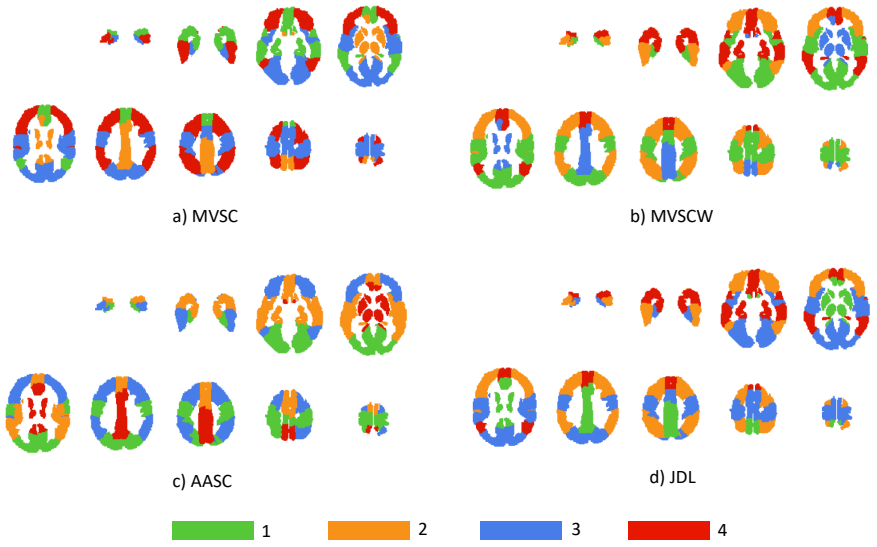


Figure 5.4: Group-wise fBNs (with k=4) in female subgroup using
a) **MVSC**: community 1 corresponds to BLN, DMN and LN, community 2 corresponds to DMN and SCN, community 3 corresponds to VN, SMN and AN, community 4 corresponds to DMN, CCN, VN and LN;
b) **MVSCW**: community 1 corresponds to AN, SMN and VN, community 2 corresponds to DMN, CCN, VN and LN, community 3 corresponds to DMN and CCN, community 4 corresponds to BLN, DMN and LN;
c) **AASC**: community 1 corresponds to SMN and VN, community 2 corresponds to AN, SMN, DMN, BLN and LN, community 3 corresponds to DMN, CCN, VN and LN, community 4 corresponds to DMN and SCN;
d) **JDL**: community 1 corresponds to DMN and SCN, community 2 corresponds to DMN, CCN, VN and LN, community 3 corresponds to SMN, AN and VN, community 4 corresponds to DMN and BLN.

5.4 Conclusions

Availability of huge open source data has enabled research on identifying networks at the group-level. Most of the work in the literature considers averaging method to quantify group-level networks, although averaging does not account for inter-subject variability. To overcome this problem, we have proposed a novel GF-MVRC framework that generates group-fused functional connectivity for each

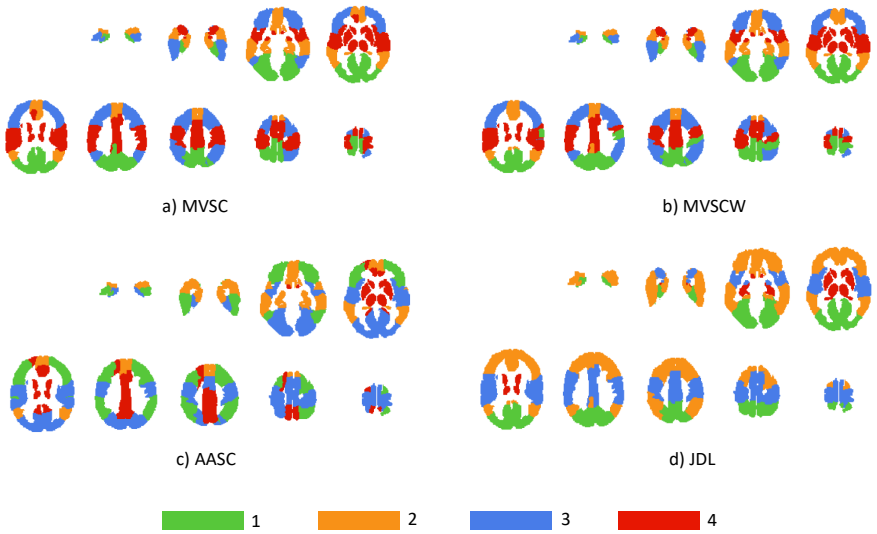


Figure 5.5: Group-wise fBNs (with k=4) in male subgroup using
a) **MVSC**: community 1 corresponds to DMN and VN, community 2 corresponds to DMN, BLN and LN, community 3 corresponds to VN, MN and AN, community 4 corresponds to SMN, AN and BLN;
b) **MVSCW**: community 1 corresponds to VN and DMN, community 2 corresponds to DMN, BLN and LN, community 3 corresponds to DMN, CCN, VN and LN, community 4 corresponds to SMN, AN, and SCN;
c) **AASC**: community 1 corresponds to DMN, CCN, VN and LN, community 2 corresponds to BLN, DMN and LN, community 3 corresponds to SMN, AN and VN, community 4 corresponds to SCN and DMN;
d) **JDL**: community 1 corresponds to VN and DMN, community 2 corresponds to DMN, CCN, VN and LN, community 3 corresponds to SMN, AN, LN and SCN, community 4 corresponds to SCN and BLN.

subject using sparsity, denseness, and group fused constraints (see Section-5.1).

Next, we determined consensus networks among all the subjects by using iterative spectral clustering. This allowed us to capture inter-subject variations by weighting subject-level matrices and detected lower dimensional subspace common for all subjects. In this step, we optimized both weights and lower-dimensional space simultaneously in the proposed framework. This is an important advancement by the proposed method over existing methods since it not only extracts group-level networks, but also accounts for the individual contribution of each subject. Finally, fBNs are captured by applying clustering on this lower-dimensional representation. Clustering with K-means is utilized with l^1 distance and 500 iterations. It resulted in k number of communities that are fBNs at the group-level.

Experimental results on openly available dataset show the efficacy of the proposed approach compared to the existing methods. Proposed method could identify multiple brain networks including visual, auditory, default mode, cognitive control, subcortical in both the male and female subgroups. These findings are inline with the literature, suggesting for a large-scale functional brain networks in the resting state.

Chapter 6

Multivariate Graph Learning for Detecting Aberrant Connectivity of Dynamic Brain Networks in Autism

Typically, functional brain networks are studied using static connectivity analysis that assumes networks to be static during the entire scan session. However, it has been reported that brain networks change with time and are inherently non-stationary over the duration of a single scan [30]. Hence, static functional connectivity may not fully capture these time-varying brain networks.

Till now, we have shown the effectiveness of MVRC algorithm on capturing static functional connectivity. In this chapter, we use the MVRC to identify dynamic evolving brain's networks. This chapter first discusses proposed multivariate graph learning based method for detecting dynamic functional connectivity followed by proposed framework of dynamic brain networks identification. The proposed work is also utilized to infer altered functional brain networks in brain disorders at the group level. As an illustration, we apply the proposed method to the publicly available fMRI dataset of autistic and healthy subjects discussed in Section-3.5. To the best of our knowledge, there are only a few studies on identifying aberrant dynamic functional brain networks in Autism [115].

6.1 Learning Multivariate Regression-based dynamic Connectivity

Dynamic functional connectivity (dFC) based analysis is emerging as a promising technique complementary to static functional connectivity [30]. Recent methods in dFC have relied upon sliding window based analysis [30], where a temporal window of a certain time duration is decided and slided by one time point, in general, to capture dynamic brain networks. This process is repeated until all T time points are covered via sliding windows. Inside each single window, functional connectivity (FC) is estimated independently of the data of other windows and networks are assumed to be static within the duration of each window length. This leads to multiple window-indexed functional connectivity adjacency matrices.

We have utilized MVRC method [68] in every window to estimate the functional connectivity matrix $\mathbf{G}_w \in \mathbb{R}^{N \times N}$, where the subscript w denotes the w^{th} window. We computed the adjacency matrix in each time window and formed a tensor $\mathcal{T}$ of dimension $N \times N \times L$, where N denotes the number of ROIs and L denotes the total number of windows.

The dynamic FC requires predefined window length in sliding window approach. There is always a dilemma in choosing an appropriate window length because results are strongly dependent on the chosen window length [30]. Shorter windows may provide spurious changes in dFC leading to difficulties in distinguishing signals of interest from other confounding factors [33]. On the other hand, larger duration windows may not capture time-varying functional networks. In a recent work, it has been shown that window length is dependent on the preprocessing step of bandpass filtering of BOLD time series [116]. Lowest frequency f_{min} of the filter provides an appropriate window length that is equal to $1/f_{min}$ [116]. It is a simplified relationship between the window length and the frequency content of a signal. We utilized this approach to select the window duration. We considered a window length of 100 seconds corresponding to the lowest frequency of 0.01Hz used in our pre-processing step of filtering. With TR=2 seconds of the Autism dataset (refer to section 3.5), this corresponds to 50 time points in one window (50 TR). This window length helps us in capturing one full cycle of the slowest bandpass filtered frequency component.

Group-level Summarized dFC

For every subject of a group, we computed tensors $\mathcal{T}$ with dimension $N \times N \times L$ (region $\times$ region $\times$ window). We carry out group-level dFC summarization that helps us in comparing TDC versus ASD group-level differences/alterations in dynamic brain networks.

Generally, adjacency matrices are averaged over subjects to form a single group-averaged adjacency matrix [117, 118]. Since this averaging considers equal contribution of every subject's data to the resultant adjacency matrix, it would result in loss of inter-subject variability. Hence, in this chapter, we identified subject-summarized adjacency matrix in each time window by estimating activation associated with each subject.

We used dFC matrices of all subjects (region $\times$ region $\times$ subject) in one time window w to identify group-level FC matrix (with dimension region $\times$ region). The collection of adjacency matrices across all subjects $s = 1, 2, ..., S$ in one time window forms a three-mode spatio-subject tensor $\mathcal{X} \in \mathbb{R}^{N \times N \times S}$, where N denotes the number of ROIs and S denotes the total number of subjects. The value of S may differ in the TDC and ASD groups.

We utilized Tucker decomposition of $\mathcal{X} \in \mathbb{R}^{N \times N \times S}$ [119], to obtain $N \times N$ adjacency matrix in one time window. It computes orthonormal subspaces corresponding to each mode of the tensor and is defined as below:

$$\mathcal{X} = \mathcal{C} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3, \tag{6.1}$$

where $\mathcal{C} \in \mathbb{R}^{N \times N \times S}$ is the core tensor, $\times_k$ represent mode k product, and $\mathbf{U}_1 \in \mathbb{R}^{N \times N}$, $\mathbf{U}_2 \in \mathbb{R}^{N \times N}$ and $\mathbf{U}_3 \in \mathbb{R}^{S \times S}$ denote the orthonormal mode matrices along all the three modes. The columns of $\mathbf{U}_3$ provide connectivity information about subjects. The first left vector ($\mathbf{u}_{3,1} \in \mathbb{R}^{S \times 1}$) of matrix $\mathbf{U}_3$ represents the most significant vector capturing highest functional connectivity energy of subjects in time window w [120, 121]. Hence, we utilized $\mathbf{u}_{3,1}$ to obtain the subject-summarized adjacency matrix $\mathbf{D}(w) \in \mathbb{R}^{N \times N}$ of time window w as:

$$\mathbf{D}(w) = \frac{\sum_{s=1}^{S} \mathbf{u}_{3,1}(s) \mathbf{G}_w(s)}{\sum_{s=1}^{S} \mathbf{u}_{3,1}(s)}, \tag{6.2}$$

where $\mathbf{G}_w(s)$ is the MVRC adjacency matrix of subject s in the w^{th} time window. Similar processing across all L windows results in a three-mode spatio-temporal tensor $\mathcal{T}$ of dimension $N \times N \times L$ (region x region x window), summarized over all subjects as mentioned above.

6.2 Extracting Dynamic Brain Networks

Next, we partitioned these estimated whole-brain dFC into non-overlapping and overlapping networks as explained below. Each network contains several densely interconnected regions, while there are relatively few connections between regions belonging to different networks.

Determination of Non-Overlapping Networks by Modularity

We used utilized modularity [63] to determine non-overlapping communities as presented in chapter 2. Next, we calculated the within-module degree z-score and the participation coefficient (PC) for each brain network [122]. Module degree z-score describes the within-module connectivity strength, whereas PC describes the connectedness of regions across different modules. For region i, within-module degree is defined as:

$$z_i = \frac{k_i - < k_i >}{\sigma_{k_i}}, \tag{6.3}$$

where k_i denotes the degree of region i in a network, $< k_i >$ denotes the average degree of all the regions in the same network, and σ_{k_i} denotes the standard deviation of all regions' degrees.

PC describes the extent to which a region belonging to one network c is connected to regions of other networks. It is defined as:

$$P_i = 1 - \sum_c (\frac{k_{ic}}{k_i})^2, \tag{6.4}$$

where k_{ic} denotes the number of connections between region i and other regions within network c and k_i denotes the degree of region i. A threshold of mean plus one standard deviation of z-values are considered for identifying "connector" hubs in each time window, while a threshold of mean plus one standard deviation of PC

are considered for identifying "provincial" hubs in each time window. Regions with values higher than the threshold and reproducible with 90% or higher percentile in all time windows are considered as "connector" hubs for the case of module-degree z-score and "provincial" hubs for the case of PC.

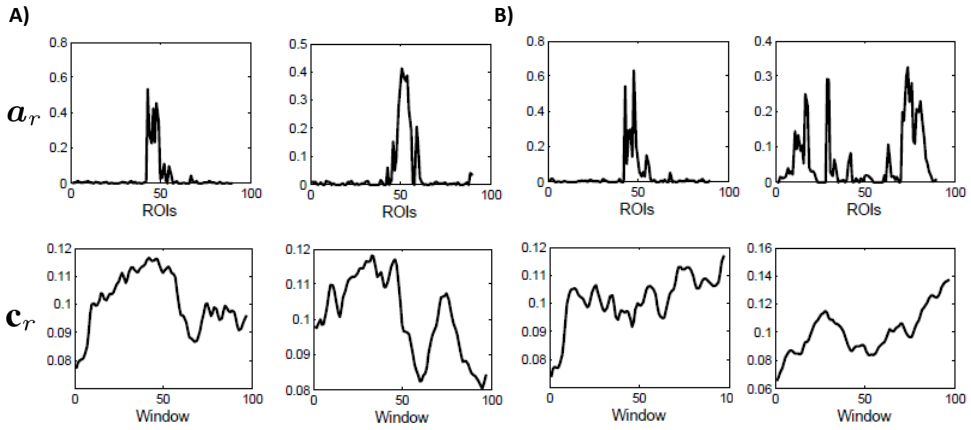


Figure 6.1: **Results on GU dataset**: We present two randomly selected loading vectors, obtained with R=10, for both A) TDC and B) ASD groups. Vector $\boldsymbol{a}_r$ represents ROI loading vectors containing membership values of each region to the r^{th} network and $\mathbf{c}_r$ presents information about the temporal loading vector corresponding to the r^{th} network.

Identifying overlapping networks

A dFC is partitioned into non-overlapping networks using modularity, i.e., each region belongs to one network only. Since certain brain regions may be involved in more than one network, it suggests for the identification of overlapping networks [36]. There are various algorithms that estimate overlapping networks, such as mixed membership algorithm proposed in [36] and non-negative tensor factorization (NNTF) proposed in [123]. In this work, we utilized NNTF to extract overlapping networks. To the best of our knowledge, this is the first study that compares overlapping networks in a neuro-development disorder.

We performed Parallel Factor (PARAFAC) decomposition [124] of dFC tensor $\mathcal{T}$ as estimated in Section 6.1. Given a rank R, it decomposes input tensor $\mathcal{T}$ into a sum of R number of rank-one tensors, generated by the outer product of column

vectors as below:

$$\mathcal{T} \approx \sum_{r=1}^{R} \lambda_r \boldsymbol{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \tag{6.5}$$

where '$\circ$' denotes the outer product, $\boldsymbol{a}_r \in \mathbb{R}^{N\times 1}$, $\mathbf{b}_r \in \mathbb{R}^{N\times 1}$, and $\mathbf{c}_r \in \mathbb{R}^{L\times 1}$ denote the loading vectors corresponding to each mode of tensor $\mathcal{T}$ and R denotes the number of rank-one tensors (also called as the number of components). The decomposition in (6.5) is expressed in the matrix form as:

$$\mathcal{T} \approx \mathcal{D} \times_1 \mathbf{A}_R \times_2 \mathbf{B}_R \times_3 \mathbf{C}_R, \tag{6.6}$$

where $\mathcal{D}$ is the diagonal core tensor and $\mathbf{A}_R = [\boldsymbol{a}_1, ..., \boldsymbol{a}_R] \in \mathbb{R}^{N\times R}$, $\mathbf{B}_R = [\mathbf{b}_1, ..., \mathbf{b}_R] \in \mathbb{R}^{N\times R}$, and $\mathbf{C}_R = [\mathbf{c}_1, ..., \mathbf{c}_R] \in \mathbb{R}^{L\times R}$ are the component matrices containing R loading vectors. $\times_{(.)}$ denotes the tensor product. This is to note that the column dimension of each of the component matrix is R in the PARAFAC model in (6.6). Here, components one and two provide loading vectors corresponding to regions and component three provides temporal loading vector. Since the adjacency matrix is symmetric in the first two modes of the tensor $\mathcal{T}$, components one and two are similar.

Modes one and two of tensor $\mathcal{T}$ contain the connectivity matrices with non-negative values. Hence, we add non-negative constraint on both $\mathbf{A}_R$, $\mathbf{B}_R$, and also on the third mode $\mathbf{C}_R$ in (6.6). Similarly, as aforementioned, matrices $\mathbf{A}_R$ and $\mathbf{B}_R$ are identical because the adjacency matrix is symmetric. There are many algorithms for carrying out non-negative PARAFAC decomposition. We applied the block principal pivoting algorithm[1] because of its good performance in fMRI studies [123]. Further, we utilized the core consistency method for determining R values for both TDC and ASD groups [125]. One may refer to [125] for further details of this method.

Identifying networks from loading vectors: We utilized each column of $\mathbf{A}_R$ component matrix to identify R number of networks. Loading vectors $\boldsymbol{a}_r$ or $\mathbf{b}_r$, for $r = 1, 2, \cdots, R$, provide the membership of regions to the component or network r, whereas $\mathbf{c}_r$ provides their time-varying or the temporal profile. A high value of i^{th} element of loading vector $\boldsymbol{a}_r$ represents greater membership of the corresponding

[1] `www.cc.gatech.edu/~hpark/nmfsoftware.php`

i^{th} region to that component r. Hence, thresholding on loading vectors $\boldsymbol{a}_r$ would capture networks as collection of ROI with higher membership. Fig. 6.1 depicts a few loading vectors $\boldsymbol{a}_r$ and $\mathbf{c}_r$ for both TDC and ASD groups considering R=10. Fig.6.1 shows two randomly selected loading vectors of both TDC (Fig.6.1A) and ASD (Fig.6.1B) groups. ROI loading vectors $\boldsymbol{a}_r$ represent membership values of 1 to N ROI. Similarly, temporal loading vectors $\mathbf{c}_r$ represent time varying profiles.

From Fig.6.1, we observe that every loading vector $\boldsymbol{a}_r$ has a few regions with higher membership values and rest with very less values in both the TDC and ASD groups. Hence, we carry out thresholding of these vectors that results in networks of ROI. We considered mean plus one standard deviation as the threshold on each loading vector. Regions with membership values exceeding the above specified threshold in vector $\boldsymbol{a}_r$ are considered as a part of the r^{th} network. This leads to the detection of overlapping networks. This is essential because one region can be a part of multiple networks owing to the relationship between different networks. For example, an auditory stimulus may trigger recollection of past events. In such a case, auditory and cognitive control networks may have some connected/overlapping regions.

Strength of networks: Further, strength of each network is computed from its ROI loading vector and temporal loading vector as $\mathbf{s}_r = \mathbf{c}_r \sum_i \boldsymbol{a}_r(i)$. This formulation was first proposed to compute the strength of communities in social networks [125]. We utilized strength vectors of similar networks in TDC and ASD groups to identify altered functional brain networks.

Fig.6.2 shows the complete process of extracting dynamic functional brain networks.

Statistical Analysis

We have done a thorough statistical analysis of dynamic networks. We looked at both the overlapping and non-overlapping networks. For overlapping networks, we tested for statistically significant differences in similar networks, identified using strengths of networks, of TDC and ASD groups. Further, for subject-summarized dynamic connectivity, we investigated statistical differences between modularity quality function. Given 90 ROIs, we also assessed differences between connector

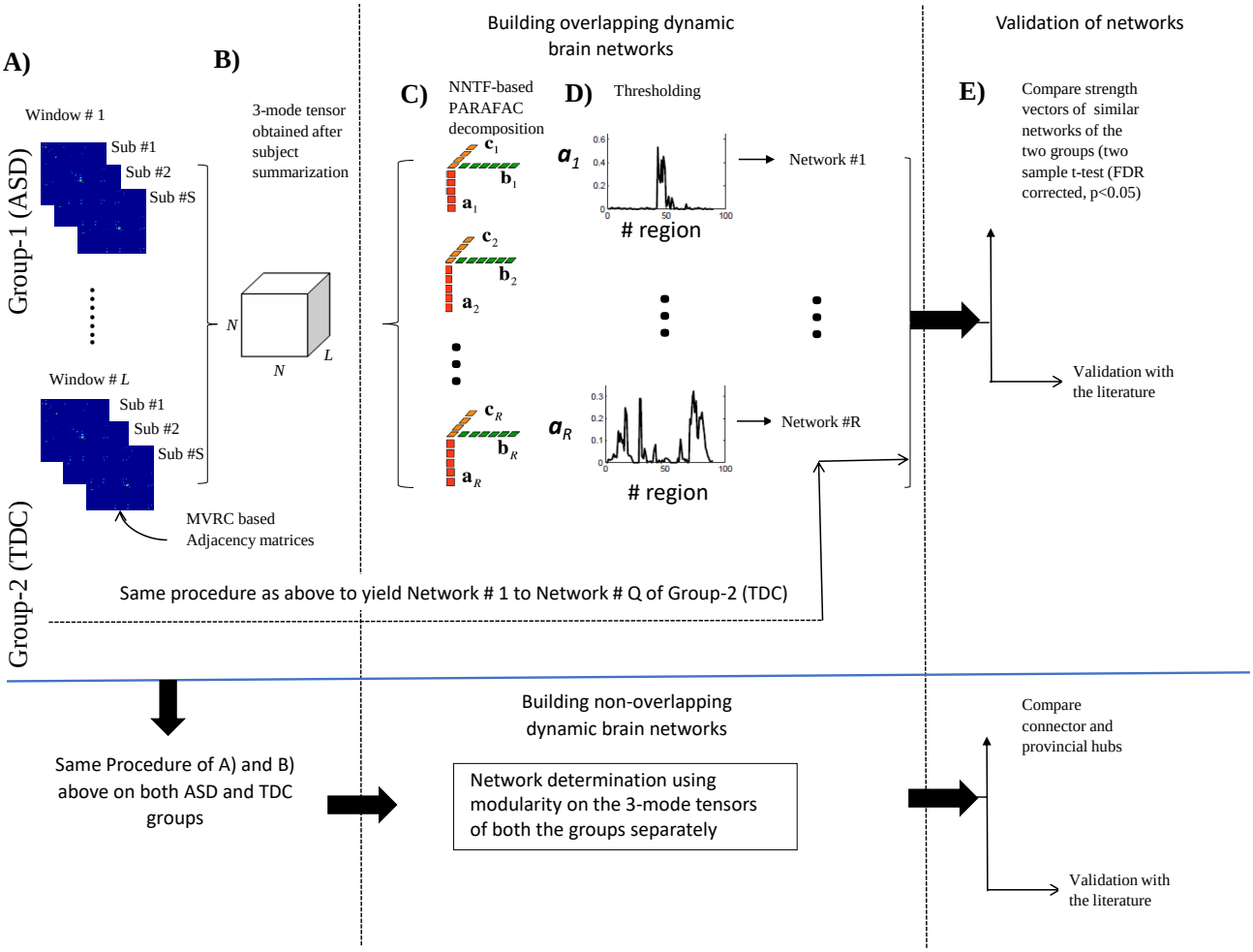


Figure 6.2: Determination of overlapping and non-overlapping dynamic functional brain networks A) MVRC based adjacency matrices are computed for all subjects in each time window for both the groups, B) 3-mode tensor of size $N \times N \times L$ is obtained after subject summarization in both the groups; top half C) Non-negative tensor factorization (NNTF) based PARAFAC decomposition with R factors where the first and second components represent region loading vectors, and the third component represents the time loading vector. D) overlapping networks are obtained by thresholding each region's loading vectors. bottom half Non-overlapping network determination of both the groups using modularity, E) Validation of results.

and provincial non-overlapping hubs extracted using modularity. Results were considered statistically significant for a value of $p < 0.05$. We used two-sample t-test for testing the statistical significance of differences between two groups. We also corrected for multiple comparison using false discovery rate (FDR) [126] (with $p < 0.05$), whenever needed.

6.3 Results

Our primary interest is to report disease-driven changes in functional brain networks. We computed dFC matrices using MVRC. Further, we studied both overlapping and non-overlapping dynamic functional brain networks. We would like to examine if the proposed approach can report changes in FC induced by ASD via comparing the results with the TDC group. In this study, we have used two publicy available dataset that are described below.

Data Description

Dataset-1 (GU dataset): This is the publicly available Autism dataset contributed by Georgetown University at the collection site of Autism Brain Image Data Exchange II (ABIDE II[2]). More details about this dataset has been provided in Section-3.5.

Dataset-2 (KKI dataset): This dataset is the pediatric Autism datasest contributed by Kennedy Krieger Institute (KKI) at the collection site of ABIDE II. This dataset is chosen due to similar age subjects as that with GU dataset. This dataset consists of 155 TDC (age: 8−13 years) and 56 ASD subjects (age: 8−13 years). Images are acquired on a Philips 3-T scanner using an Echo Time (TE) equal to 30 ms and Repetition Time (TR) equal to 2500 ms. Subjects underwent a 6 min, 40 s resting-state scan with fMRI data collected with 47 axial brain slices (dimension 64×64) in each volume.

Data Preprocessing

All BOLD images are pre-processed using SPM12 (Statistical Parametric Mapping[3]). First five brain volumes, accounting for T1 equilibration effect, are discarded in both dataset. The other functional volumes are slice time corrected, motion corrected, spatially normalized onto the MNI space resulting into functional images of dimension $53 \times 63 \times 52$ (3-mm isotropic voxels) that are smoothed with a Gaussian kernel of 6 mm full width half maximum. Finally, nuisance variables (6

[2] http://fcon_1000.projects.nitrc.org/indi/abide/
[3] http://www.fil.ion.ucl.ac.uk/spm/software/spm12/

Table 6.1: **Summary of TDC versus ASD Subjects' data.**

	Dataset-1	(GU)	Dataset-2	(KKI)
Characteristic	TDC (S=26)	ASD (S=35)	TDC (S=85)	ASD (S=39)
Gender	Male	Male	Male	Male
Age (years)				
Mean (SD)	10.9 (1.62)	11.17 (1.49)	10.44 (1.49)	10.49 (1.53)
Range	8.06−13.79	8.25−13.91	8.02−13	8.01−12.99
Full Scale IQ				
Mean (SD)	120.32 (13.53)	119.06$^{\gamma}$ (14.18)	114.9 (11.04)	101.8$^{\gamma}$ (15.93)
Range	91−148	95−149	94−140	84−148
ADI-R				
Social total A	−	19.74$^{\gamma}$ (5.28)	−	20.29 (4.08)
Verbal total BV	−	14.97$^{\gamma}$ (4.91)	−	15.75 (3.15)
RRB total C	−	5.09$^{\gamma}$ (2.38)	−	8.41 (2.45)
R Onset total D	−	2.56$^{\gamma}$ (1.23)	−	6.07 (2.03)
ADOS				
Total	−	10.52$^{\Upsilon}$ (4.61)	−	−
Communication	−	3.18$^{\Upsilon}$(1.54)	−	−
Social	−	7.33$^{\Upsilon}$ (3.53)	−	−
Stereo Behavior	−	1.89$^{\Upsilon}$ (1.58)	−	−

$^{\gamma}$One subject's score is missing. $^{\Upsilon}$Seven ASD subjects do not have these scores. '−' signifies that these scores are not available for the TDC group.
TDC: Typically Developing Control; ASD: Autism Spectrum Disorder; ADI-R: Autism Diagnostic Interview-Revised; Social total A: Reciprocal Social Interaction Subscore A; Verbal total BV: Abnormalities in Verbal Communication Subscore; RRB total C: Restricted, Repetitive, and Stereotyped Patterns of Behavior; ADOS: Autism Diagnostic Observation; Stereo Behavior: Stereotyped Behaviors and Restricted Interest.

head motion parameters, average cerebro-spinal fluid signal from ventricular mask, and average white matter signal from white matter mask) are regressed out from each voxel's time series followed by bandpass filtering in the frequency range of 0.01 to 0.1 Hz.

Framewise displacement (FD) is used to identify temporal artifact in each subject by calculating the derivatives of the six rigid-body realignment parameters estimated during motion correction preprocessing step [127]. Subjects having more than 30% of total brain volumes with FD>0.5mm were excluded from further analysis. This led to the rejection of 4 out of 55 TDC subjects' data and 12 out of 51 ASD subjects' data from dataset-1 and 14 out of 99 TDC subjects' data and 2 out of 41 ASD subjects' data from dataset-2. This is to note that, in dataset-1, head motion was minimized using foam cushions placed in the space between the participant's head and the headcoil. We included only male subjects in both dataset. in the study to control for gender differences associated with functional brain networks. After quality check, a total of 26 TDC and 35 ASD subjects remained from Dataset-1 (GU) and a total of 85 TDC and 39 ASD subjects remained from Dataset-2 (KKI) for our analysis. A two-sample t-test with unequal variance showed no significant difference (at $p < 0.05$ significance level) in the age of two groups of both the dataset (Table 6.1).

Preprocessed data is parcellated into 90 anatomical predefined ROIs via AAL atlas (refer to Table 2.1) [50]. We averaged the time-series of all voxels belonging to the same ROI to find the region-representative time series of every ROI. This resulted into a matrix $\mathbf{X}$ of dimension $T \times N$, where T denotes the number of time points and N=90 denotes the number of ROIs obtained with AAL atlas. Network labels corresponding to each of the 90 ROIs are listed at the bottom of Table-2.1, based on the information provided in [128].

Next, we first present results on GU dataset. We will use KKI dataset to demonstrate the reproducibility of results in further Section. In each time window, we used MVRC method to construct $N \times N$ adjacency matrix with parameters $\mu_1 = 0.1$ and $\mu_2 = 0.85$ (see equ.(3.4)), as used previously in [68]. In order to balance the sparsity level and the window length in windowed MVRC formulation (used here for dynamic time-varying networks compared to static network formulation of MVRC [68]), we used $\mu_1 = 0.1$ that is smaller compared to $\mu_1 = 0.25$ used in [68].

The collection of adjacency matrices of all subjects $s = 1, 2, ..., S$ in each group forms a three-mode tensor $\mathcal{X} \in \mathbb{R}^{N \times N \times S}$ in each time window, where the value of S in control and ASD group is 26 and 35, respectively (see Table 6.1). Next, as described in Section 6.1, we carried out subject-summarization using Tucker decomposition and formed a spatio-temporal tensor $\mathcal{T} \in \mathbb{R}^{N \times N \times L}$ with window numbers as the indices of the third mode.

Non-Overlapping Dynamic Functional Brain Networks Identified using Modularity

To have a better understanding of dFC states, we estimated modularity based networks that play a pivotal role in the organization of overall networks. We estimated non-overlapping networks using modularity from tensor $\mathcal{T}$, as explained in the previous section, in both TDC and ASD groups. We could identify 3 to 6 number of networks across all time windows in the TDC group. The number of networks for ASD group was 4 to 6. With less number of detected networks, regions belonging to multiple functional brain networks were found to club randomly. With 6 numbers of detected networks in both TDC and ASD groups, we could largely locate the following functional networks: network 1- SMN, AN; network 2- BLN, DMN; network 3- CCN, LN; network 4- BLN, SCN; network 5- VN; network 6- DMN.

The value of the modularity quality function Q is ranged between 0.5479 to 0.7821 for the TDC and between 0.6038 to 0.7720 for the ASD group. Significant differences were observed in the Q values of both the groups ($p < 0.05$, two sample t-test) that indicates that the quality of partitioning of dynamic functional brain networks using modularity is distinguishable across the two groups, although we did not observe consistency in networks of all time windows as mentioned above. In addition, we observed random grouping of regions in networks. At times, modularity is assessed on the FC matrix averaged across all time windows. However, this procedure results in loss of temporal characteristics associated with dynamic networks and hence, we did not employ averaging on adjacency matrix for our analyses.

To have further understanding of the hubs' architecture in dFC networks, we

estimated the connector and provincial hubs' architecture using the module-degree and the PC, respectively. Higher module-degree connector hubs in the TDC group was shown by nine regions: bilateral rolandic operculum, bilateral medial superior frontal gyrus, bilateral medial orbital superior frontal gyrus, left superior occipital gyrus, and bilateral putamen. In the ASD group, connector hubs were formed by nine regions: left dorsolateral superior frontal gyrus, left rolandic operculum, left medial superior frontal gyrus, right anterior cingulate cortex, right parahippocampal gyrus, left superior occipital gyrus, left postcentral gyrus, right precuneus, and right putamen. Among these regions, four common regions (left rolandic operculum, left medial superior frontal gyrus, left superior occipital gyrus, right putamen) were located in the hubs of both ASD and TDC groups. In the literature on sFC in autism [129], it is stated that the presence of group-specific regions in the hubs of ASD group indicates an alteration of hub organization. In this chapter, we have studied altered hubs' organization in dynamic functional connectivity compared to that studied for static functional connectivity in [129].

Regions showing higher PC in both the TDC and ASD groups comprised of left parahippocampal gyrus, right amygdala, bilateral supramarginal gyrus, left angular gyrus, bilateral temporal pole superior temporal gyrus, and bilateral inferior temporal gyrus. Hence, suggesting for no altered provincial hubs in both the groups.

Overlapping Dynamic Functional Brain Networks Identified using PARAFAC Decomposition

As presented in (6.6), the spatio-temporal tensor $\mathcal{T}$ is decomposed into R number of rank-one tensors, where R represents the number of networks. The value of R is decided according to the core consistency [125]. We considered initial value of R from 2 to 15 and for each value of R, we computed the core consistency value and the plot of these values as a function of R is shown in Fig.6.5. From this figure, we observe an abrupt change in the slope at R=10 for both TDC and ASD groups. Moreover, core consistency value above 0.5 is considered acceptable [125] and corresponding R values range are considered to be as optimal range. Therefore, we considered R=10 for both the groups.

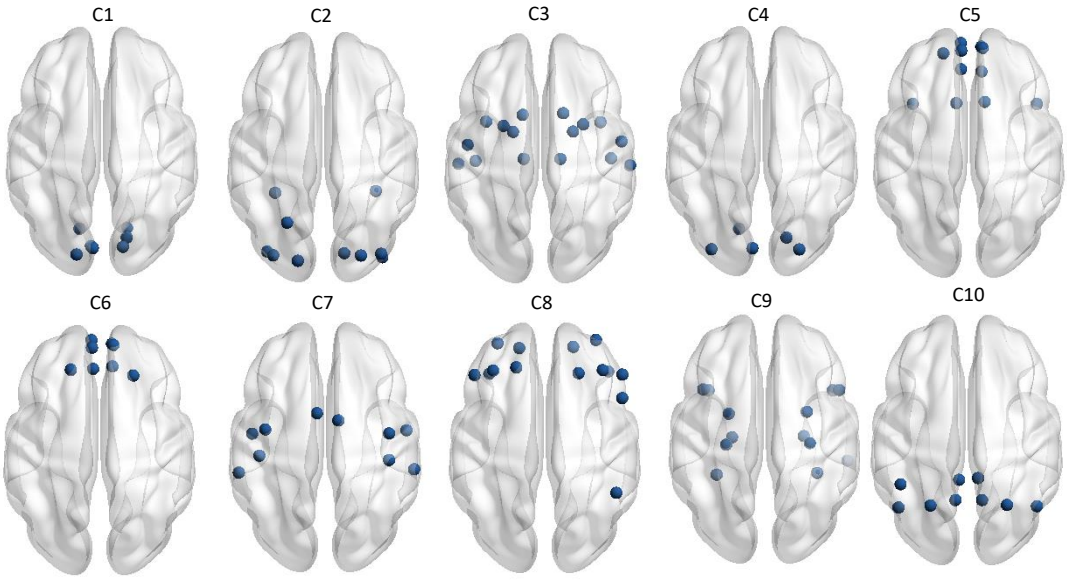


Figure 6.3: Group-level dFC networks (denoted as community 1 (C1) to community 10 (C10) identified on TDC subjects of GU dataset.

After estimating the value of R, we factorized tensor $\mathcal{T}$ into R components. We identified ten overlapping dFC networks in both TDC and ASD groups as shown in Fig.6.3 and Fig.6.4 respectively. From these figures, we note that the spatio-subject tensor decomposition results in multiple functional brain networks in both TDC and ASD groups. We could identify Visual Network (VN), Auditory Network (AN), Bilateral Limbic Network (BLN), Default Mode Network (DMN), Somato-Motor Network (SMN), Subcortical Network (SN), Language Network (LN), and Cognitive Control Network (CCN).

Comparison of Networks in ASD and TDC Groups

To compare the strength of identified dFC networks among both the groups, we compared the strength of similar networks. We computed the strength vectors $\mathbf{s}_r$ (see Section 6.2) of similar networks in both the groups and examined the statistically significant differences by using two-sample t-test at 0.05 significance level as shown in Fig.6.6. Column 4 and 9 of this figure shows the obtained p-values.

Compared to the TDC group, the ASD group showed altered connectivity in 56

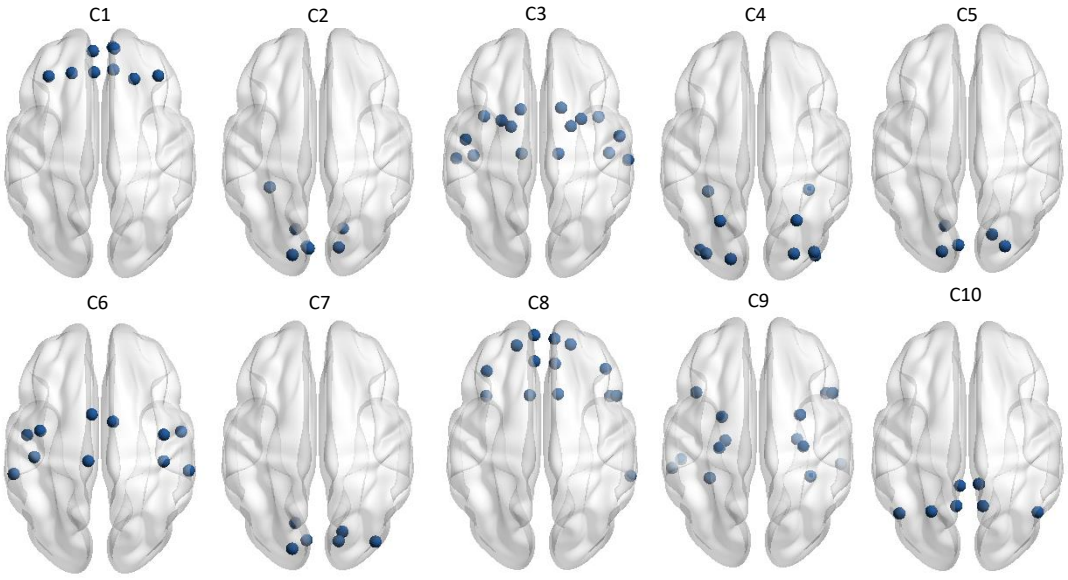


Figure 6.4: Group-level dFC networks (denoted as community 1 (C1) to community 10 (C10) identified on ASD subjects of GU dataset.

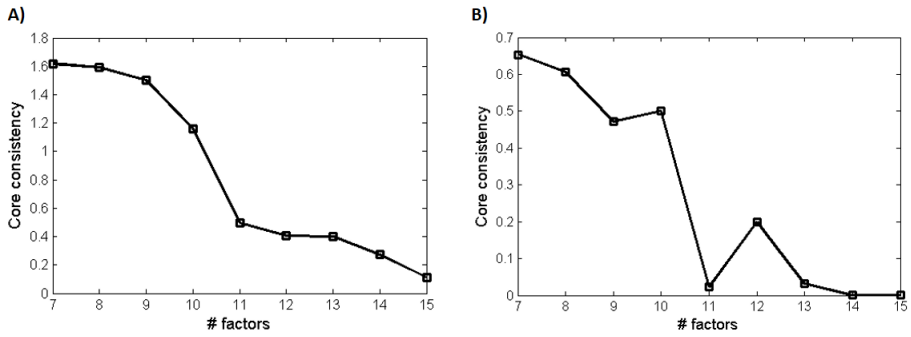


Figure 6.5: Plot of core consistency with varying number of components in A) TDC and B) ASD groups of GU dataset.

regions (24 regions in both right and left hemisphere) (two-sample t-test, $p < 0.05$), as shown in Fig.6.6: i) the bilateral rolandic operculum, bilateral insula, bilateral caudate nucleus, bilateral putamen, bilateral pallidum, bilateral thalamus, bilateral heschl gyrus, bilateral superior temporal gyrus (network 3 of both the groups), ii) the right calcarine fissure, left cuneus, left lingual gyrus, right superior occipital gyrus (networks 4 and 5 of TDC and ASD groups, respectively), iii) the left orbital superior frontal gyrus, bilateral olfactory cortex, bilateral medial orbital superior frontal gyrus, bilateral gyrus rectus, bilateral temporal pole middle temporal gyrus

	TDC	ASD	*p(h)*			TDC	ASD	*p(h)*
VN	 C1	 C2	0.9550 (0)		DMN CCN	 C6	 C1	**1.5507e-04 (1)**
VN	 C2	 C4	0.2128 (0)		SMN AN LN	 C7	 C6	0.59 (0)
AN CCN SCN	 C3	 C3	**7.2806e-04 (1)**		BLN	 C9	 C9	**4.3543e-25 (1)**
VN	 C4	 C5	**0.00048 (1)**		DMN VN	 C10	 C10	**8.9224e-16 (1)**
DMN BLN CCN	 C5	 C8	**1.9707e-04 (1)**					

Figure 6.6: **Results on GU dataset**: Statistical difference of strength vectors $\mathbf{s}_r$ of similar networks identied for TDC and ASD groups (refer to Fig.6.3). Vectors $\mathbf{s}_r$ are statistically tested using two-sample t-test with 0.05 significance level (column 2). The value of $h = 1$ indicates statistically significant difference in the dFC networks of ASD and control groups (p<0.05). VN: Visual Network; SMN: Somato-Motor Network; AN: Auditory Network; CCN: Cognitive Control Network; DMN: Default Mode Network; SCN: Subcortical Network; LN: Language Network; BLN: Bilateral Limbic Network.

(networks 5 and 8 of TDC and ASD groups, respectively), iv) the bilateral dorsolateral superior frontal gyrus, bilateral medial superior frontal gyrus, right medial orbital superior frontal gyrus, bilateral anterior cingulate cortex (networks 6 and 1 of TDC and ASD groups, respectively), v) the bilateral hippocampus, bilateral parahippocampal gyrus, bilateral amygdala, bilateral fusiform gyrus, bilateral temporal superior temporal gyrus, right inferior temporal gyrus (network 9 of both the groups), and vi) the bilateral posterior cingulate cortex, left superior parietal gyrus, bilateral angular gyrus, bilateral precuneus (network 10 of both the groups). The distribution of these regions is largely in accordance with subcortical, cognitive control, default mode, visual, bilateral limbic and auditory networks. These findings are inline with the literature as discussed in the next Section and suggest alterations in large scale dynamic functional brain networks in ASD.

Comparison with Static functional Conenctivity

We computed sFC using MVRC method, resulting in an $N \times N$ size adjacency matrix for each subject of both the groups. Next, we applied two-sample t-test on each FC value of the adjacency matrix, separately, to observe alteration of static FC across the two groups. This approach allowed the determination of regions where sFC differed between the TDC and ASD groups. Compared to the TDC group, the ASD group showed no significant FC alterations across the two groups (FDR corrected at $p \leq 0.05$).

Validation of the advanced MVRC method on existing studies

In this section, we validate the use of the advanced MVRC method based adjacency matrices in our framework compared to the correlation based matrices on the existing method of extracting network states.

Conventionally, brain regions states are identified using correlation matrices computed in sliding windows [115]. Here, first, connection strengths from the symmetric correlation matrices are stacked for all subjects and windows ($= LS \times C$ where L denotes the number of windows, S denotes the number of subjects, and C is $N \times (N-1)/2$ connections from $N \times N$ sized symmetric correlation matrix) in a group. Next, k-means clustering based on Manhattan distance is performed on the columns of this matrix. This produces multiple dynamic brain states as shown in Fig. 6.7 and 6.8. Results in these states are depicted with reordered regions according to the underlying networks (refer to [115]). From each identified states, we computed dwell time for each individual subject in both the groups and carried out two sample t-test ($p<0.05$) to statistically compare dwell times between the two groups. We observed dwell times of states 1, 2 and 3 to be statistically different using MVRC adjacency matrices, whereas only two states 1 and 3 are found to be different using CORR adjacency matrices. Thus, MVRC based adjacency matrices are able to reveal better differences between the ASD and TDC groups.

On further comparing the statistically different states in Figs. 6.7 and 6.8, results with both the methods show altered connectivity in cognitive control (CCN), visual (VN), default mode networks (DMN), and other networks. Pearson correla-

tion based connectivity in Fig.6.8 shows somewhat enhanced (read colors according to color bar) networks in ASD in states 1 and enhanced in TDC in states 3. On the other hand, MVRC based results show enhanced networks in TDC compared to ASD in both the states 1 and 3. In general, it is known that there is a reduced coordination among brain regions in autism. Inline with this belief, from the above results, we note that MVRC based connectivity matrices show enhanced connectivity in TDC compared to ASD, while the Pearson correlation based matrices show some enhanced networks in TDC and some in ASD that may lead to confused findings.

Further, MVRC based connectivity matrices show enhanced connections largely along the block diagonal connections, while Pearson correlation based adjacency matrices show connectivity in almost the entire 90x90 matrix. Thus, in a way, Pearson Correlation based adjacency matrices are showing connectivity largely across the whole brain that may be misleading. Compared to this, MVRC based results can be interpreted much easily.

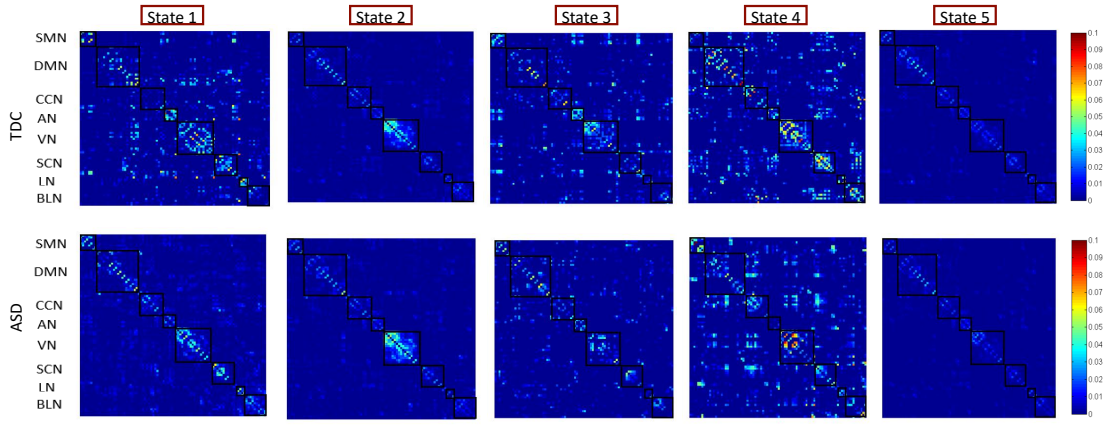


Figure 6.7: Results on brain regions' states using the adjacency matrices generated using MVRC.

In summary, this method by Yao et al. studies brain regions in states via applying k-means clustering on the adjacency matrices, but does not extract *dynamic overlapping brain networks*. Hence, it is a significantly different work compared to the presented study that first identifies overlapping communities of ROIs in both the groups separately and then statistically compare the strength of these communities across the two groups. Thus, these two works are important contributions

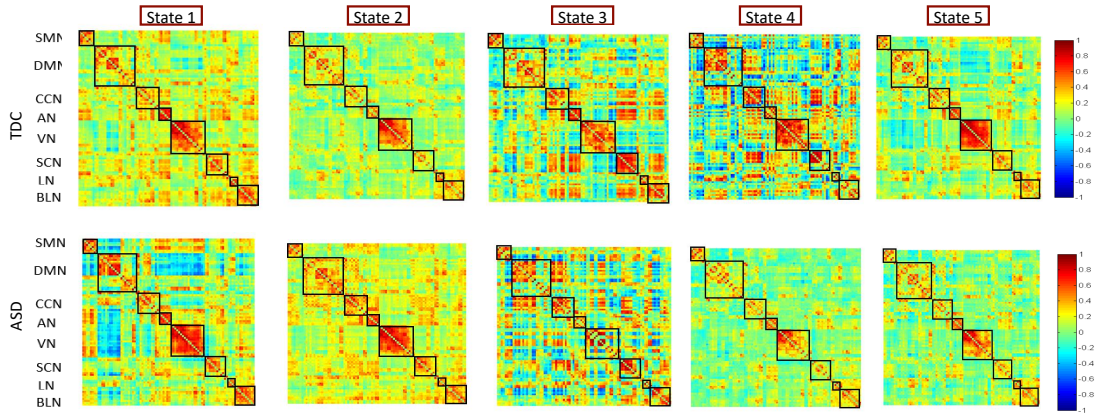


Figure 6.8: Results on brain regions' states using the adjacency matrices generated using Pearson correlation.

to the literature on fMRI studies in different ways.

Reproducibility analysis

In order to check the reproducibility of results, the proposed method has also been tested on KKI dataset. Obtained communities from both TDC and ASD groups are being shown in Fig.6.9. First and second row of this figure show obtained communities, largely lying in frontal, occipital, limbic, parietal and temporal parts of the brain, in both the groups. Similar brain activation has been observed in this dataset as was obtained with the GU dataset. In addition, we also tested the strengths of communities between the two groups. We identified a number of altered communities in ASD compared to TDC as shown in the third row of Fig. 6.9

Discussion

This chapter utilizes resting-state fMRI to find aberrant dynamic brain networks in ASD. We utilized open source autism dataset with age-matched male TDC and ASD subjects. The proposed method utilized recent multivariate graph learning method over multiple windows of fMRI data to identify dFC matrices. This provides a three mode spatio-subject tensor (region x region x subject) in each

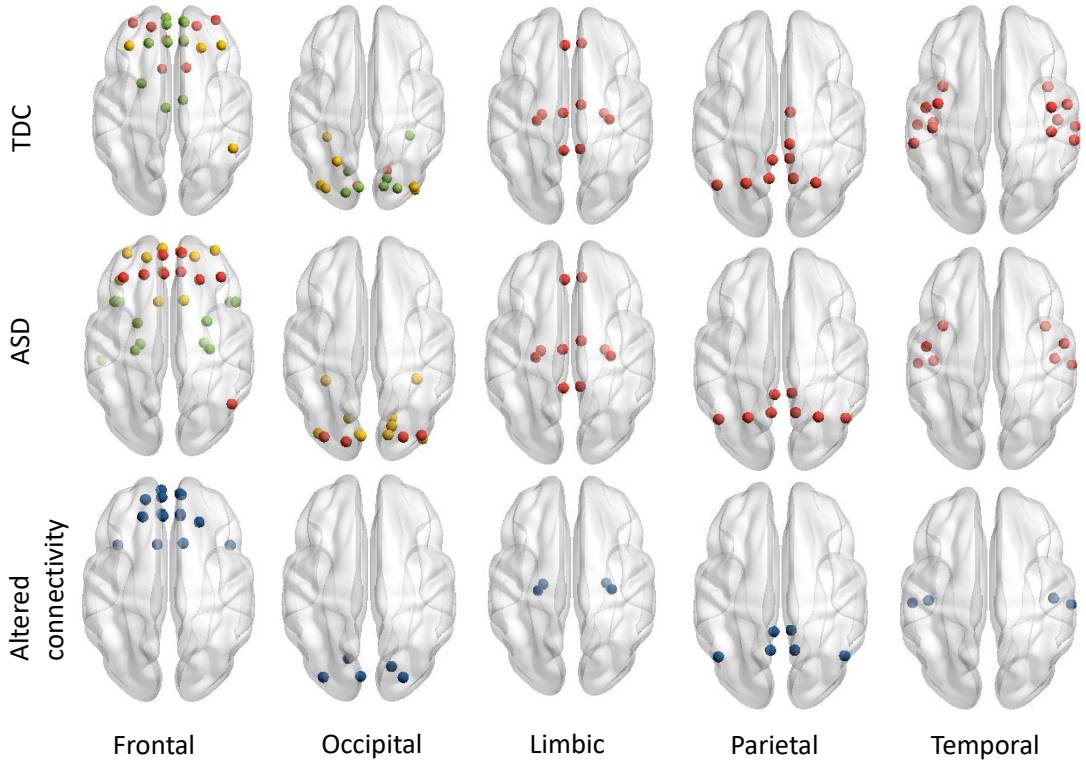


Figure 6.9: Dynamic communities obtained on KKI ABIDE dataset. First and second row of this figure show obtained communities, largely lying in the frontal, occipital, limbic, parietal and temporal parts of brain, in both the groups. Third row represents altered communites of both groups ($p<0.05$).

window.

To study brain networks at group-level, subject-summarization is carried out in each time window using Tucker decomposition based approach instead of the commonly used averaging technique. Generally, averaging about subjects is used to compare group-level networks. However, averaging may not fully capture the potential networks' organization and as a result subtle information may be lost. Subject-summarization in each time window results in three-mode spatio-temporal tensor (region x region x window).

We analyzed dynamic functional networks, both non-overlapping and overlapping, in ASD vis-à-vis TDC group. Non-overlapping dynamic networks are studied using modularity and alterations were noted in the cognitive control and default mode network. However, these networks could not be identified in all time win-

dows. Most of the windows showed random merging of ROIs belonging to different multiple brain networks. In addition, 'connector' and 'provincial' hubs were also studied based on the level of their intra-module connectivity and inter-module connectivity, i.e., their participation in multiple modules.

Overlapping networks were identified using spatio-temporal three mode tensor using PARAFAC decomposition on this tensor. Factorized component matrices, one and two, contained the loading vectors corresponding to ROIs, and component three contained loading vector corresponding to the temporal window, representing temporal evolution of dynamic networks at group-level. We applied thresholding on ROIs loading vectors resulting into networks of ROIs representing dFC states at the group-level. Thresholding allowed regions to be a part of multiple networks resulting in overlapping networks, while the existing method of modularity resulted in non-overlapping networks, where a ROI could be part of one network only. We observed networks to be differently activated across time in both the groups.

In this chapter, we studied overlapping dynamic networks alteration in healthy and autistic subjects and results show statistically significant differences in multiple functional brain networks. The differences in the functional brain networks of TDC and ASD groups is assessed using two-sample t-test. We considered differences to be significant when $p < 0.05$. We observed alterations in DMN, VN, SCN, BLN, CCN, and AN in ASD group compared to the healthy controls. We observed widespread altered networks in autistic subjects compared to TDC. Further investigation on time-varying abnormalities in these networks will improve the comprehensive understanding of aberrant dFC in ASD.

Consistency with the literature

VN related altered networks (4th network of TDC and 5th network of ASD) include regions such as right calcarine fissure (index no. 44 in AAL atlas), left cuneus (45), left lingual gyrus (47), and right superior occipital gyrus (50), left superior parietal gyrus (59) (10th network of both TDC and ASD groups). Calcarine fissure region forms the core of visual cortex. Abnormality in this region and in lingual gyrus region has been reported previously in ASD [130]. Cuneus region of VN is responsible for control of visual attention and refreshing information processing in

working memory [131]. Abnormality in superior parietal gyrus region in our work is consistent with the previous work [132]. In addition, we observed abnormality in fusiform gyrus (55, 56) region in network 9 of both the groups. This region is involved in a network having BLN regions. This region is known to be one of the abnormal connectivity key regions for face recognition and other social functions [115].

In our analyses, ASD subjects showed DMN activity in three overlapping networks (1, 8 and 10). On the other hand, TDC subjects showed DMN activity in four networks (5, 6, 8, and 10). From Figure-6.6, we observe that similar ROI based network 5 of TDC group and network 8 of ASD group are statistically different. This suggests alteration in left orbital superior frontal gyrus (5), medial orbital superior frontal gyrus (25, 26). Similarly, network 6 of TDC is statistically different from the similar network 1 of ASD, suggesting for alteration in dorsolateral superior frontal gyrus (3, 4), medial superior frontal gyrus (23, 24), anterior cingulate cortex (ACC) (31, 32). Moreover, network 10 of both TDC and ASD groups containing posterior cingulate cortex (PCC) (35, 36), angular gyrus (65, 66), and precuneus (67, 68) regions are statistically different.

Recently, a previous time-varying network study on ABIDE dataset has noted decreased connectivity for precuneus and PCC [115], while alteration in ACC connectivity is being reported in [133] as has been noted by us. In addition, we have observed reduced ranking of superior frontal gyrus and angular gyrus that is in consonance with global hypoconnectivity noted in these regions in autism subjects [134]. Further, it is noteworthy that our results show one additional DMN network (number 8) in TDC group compared to the ASD group. This network depicts activation in superior frontal gyrus dorsolateral, superior frontal gyrus orbital part, and middle frontal gyrus orbital part regions. These regions are a part of DMN and known to be active during resting-state. However, ASD group did not show activation in middle frontal gyrus orbital part region (9, 10) as observed from Figure-6.3. Our findings on aberrant activity in DMN regions confirm it's role in ASD.

ASD group showed activity of CCN in three overlapping networks (1, 3 and 8). TDC subjects also showed CCN activity in three networks (3, 8, and 10). From Figure-6.6, we observe that network 3 of TDC group is statistically different from

similar network 3 of the ASD group. This network involves insula region (29, 30) of CCN that generally shows aberrant activity in ASD due to restricted, repetitive behavior. Abnormal connectivity in insula (29,30) in our study is consistent with previous studies [115] and indicates loss in inhibitory control function. Similarly, some regions of CCN (right inferior frontal gyrus opercular part (12), inferior frontal gyrus triangular part (13, 14), and inferior parietal lobule (61, 62)) are found to be deactivated in the ASD group. Inferior frontal gyrus is related to deficit in social language processing and attentional mechanism and is commonly reported altered in ASD [135].

ASD subjects showed activity of AN in two overlapping networks (3 and 6). On the other hand, TDC subjects also showed AN activity in two networks (3 and 7). From Figure-6.6, we observe that network 3 of TDC group is statistically different from the similar network 3 of the ASD group, suggesting altered connectivity in rolandic operculum (17, 18), heschl gyrus (79, 80) and superior temporal gyrus (81, 82) regions. However, network 7 and 6 of TDC and ASD groups, respectively, are found to be statistically similar. In this network, rolandic operculum (17, 18) is associated with SMN and LN related regions. Our results show significantly abnormal connectivity linked to heschl gyrus and superior temporal gyrus in the ASD. These regions are related to speech comprehension and auditory processing in humans [136].

ASD group showed activity of SCN in one network (3). TDC group also showed SCN activity in one network (3). We observe from Figure-6.6 that this network is statistically different among the two groups suggesting for alteration in caudate nucleus (71, 72), putamen (73, 74), and pallidum (75, 76) regions. In addition, thalamus region (77, 78) that is a part of SCN is not observed to be activated in ASD. This indicates that probably there is less activity in the thalamus cortex of the autistic brain. This observation is also inline with the literature that autism subjects do not show activation in thalamus region [137].

ASD subjects showed activity of BLN in two overlapping networks (8 and 9). On the other hand, TDC subjects also showed activity in two networks (5 and 9). From Figure-6.6, we observe that network 5 of TDC group is statistically different from the similar community 8 of ASD group, suggesting for alteration in olfactory cortex (21, 22), gyrus rectus (27, 28) and temporal pole middle temporal

gyrus (87, 88). Similar statistically different network 9 of both the TDC and ASD groups indicates alteration in other BLN regions such as hippocampus (37, 38), parahippocampal gyrus (39, 40), amygdala (41, 42), temporal pole superior temporal gyrus (83, 84), and right inferior temporal gyrus (90). BLN regions such as amygdala (41,42) has been the focus of ASD related studies, given its important role in socio emotional processing [138]. Previous studies also noted hyperconnectivity in right parahippocampal gyrus [134] and hippocampus regions [139]. Inferior temporal gyrus is known to be the motion area within the visual cortex of the human brain [140] and abnormality in these regions suggests aberrant dFC in autistic subjects compared to age-matched healthy subjects.

ASD subjects showed activity of LN in three overlapping networks (6, 8 and 9). On the other hand, TDC subjects showed LN activity in one community (7). From Figure-6.6, we observe that network 7 of TDC group is statistically different from the similar network 6 of ASD group indicating alteration in supramarginal gyrus (63, 64) similar to results previously reported in [141].

6.4 Conclusions

We analyzed resting-state fMRI data to study changes in dynamic functional brain networks at the group level in ASD and TDC groups. We observed significant differences in overlapping dynamic functional brain networks that may prove helpful in finding biomarkers in ASD. In the proposed analysis pipeline, autistic subjects showed significant differences in dFC networks compared to healthy subjects. Group-level dynamic functional brain network analysis showed changes in DMN, VN, SCN, BLN, CCN, and AN in the ASD group compared to the control group. These findings are inline with the literature and suggest alterations of large scale brain networks in ASD.

This work is positioned towards contributing to the literature by proposing an advanced signal processing methods' based pipeline for extracting *dynamic overlapping* functional brain networks. So far, study of functional brain networks in ASD versus TDC is limited in three respects a) in terms of building adjacency matrices that are largely built by Pearson Correlation, while we have used an advanced method (MVRC method) to compute adjacency matrices, b) in terms of finding

networks that are largely computed using modularity yielding non-overlapping networks, while overlapping networks appear to be intuitively more correct and has been done here, and c) in terms of static versus dynamic networks. While building static networks is more common, dynamic networks are more realistic and are being actively researched now. The present study expands the current theory of extracting dynamic brain networks and advocates the use of dynamic analyses to better account for and understand FC differences compared to static FC analysis. Hence, this study contributes to the literature from the methods point of view.

We have also shown the validation of the proposed method on two publicly available dataset. In addition to this, we have also shown the validation of MVRC (an advanced method) based adjacency matrices compared to the Pearson correlation based adjacency matrices on the recent network state method by Yao et al. Results show that findings are more consistent with the literature on revealing differences between the ASD and TDC groups with the MVRC based adjacency matrices compared with those extracted with the Pearson correlation. This also provides credence to our work and establishes the need to use advanced signal processing methods. Since autism is a spectral disorder, we believe that our work on extracting dynamic overlapping functional brain networks using advanced methods is complementary to the work on finding network states. Joint analyses of such works can help in finding potential biomarkers to aid in the diagnosis of Autism disorder.

Chapter 7

Brain networks modeling via estimation of perceived stimulus

Till now, we have utilized region-averaged time series across all 90 AAL ROIs to construct fBNs. Recently some studies have attempted to compute fBNs via estimated intrinsic stimulus (or activity signal) at brain regions instead of commonly used region-averaged time series [128, 142, 143, 144, 145]. The idea behind these methods is that the functional connection between regions is largely due to intrinsic stimulus. Owing to this, we present novel intrinsic stimulus estimation method which overcomes certain limitations of existing methods and can be further utilized to build fBNs.

First, we begin with brief overview of existing work in the direction of estimating intrinsic stimulus. An fMRI BOLD signal $\mathbf{y}_i$ at particular brain's voxel V_i is comprised of hemodynamic response function (HRF) convolved with stimuli signal and is represented as [146]:

$$\mathbf{y}_i = \mathbf{s}_i \otimes \mathbf{h}_i + \xi_i, \tag{7.1}$$

where $\mathbf{s}_i$ represents intrinsic stimuli signal, $\mathbf{h}_i$ represents HRF at that voxel V_i, and ξ_i is the noise. In general, this noise is correlated in time and is generally modeled as autoregressive noise of order 1 (AR (1)) [147]. Therefore, $\xi_i \in N(0, \mathbf{\Gamma})$ is the vector of M-length representing colored AR(1) noise, where $\mathbf{\Gamma}$ is a symmetric positive definite covariance matrix of size MxM with its l^{th} element $\rho^{|l|}$. Equation

(1) has two unknown functions: the HRF $\mathbf{h}_i$ and the activity signal $\mathbf{s}_i$.

As aforementioned, in general, region representative time-series at a brain region is computed by averaging BOLD time-series $\mathbf{y}_i$ of all voxels belonging to that region. Recent methods have tried to build brain networks via estimated intrinsic stimulus or activity signal instead of utilizing region-averaged time-series [128, 142, 143, 144, 145]. In order to estimate stimuli signal, equ. (7.1) can be utilized. This equation in general has two unknowns: the HRF $\mathbf{h}_i$ and the stimuli signal $\mathbf{s}_i$, although most previous studies assume a pre-defined fixed shape of HRF despite the fact that the shape of HRF may be different across different brain regions as well as across subjects.

Recently in [128], fixed canonical HRF is assumed along with spatio-temporal priors on the underlying stimuli signal to estimate it. In [143, 144, 145], sparsity constraint is imposed on the underlying stimuli signal, but assuming apriori fixed canonical HRF shape. These methods assume apriori canonical shape of HRF and suffer with limitations of fixed HRF shape in their framework because HRF shape may vary across voxels or across different brain regions and across subjects. A fixed pre-defined canonical shape across all voxels may adversely impact the estimation of stimuli signal.

Motivated with the above literature review, we note that there is a need for developing a methodology for joint estimation of underlying stimuli signal and the HRF. In this chapter, we propose a dual iterative framework for the estimation of two unknowns in (7.1). We call the estimated stimuli signal as the *perceived stimulus*. This perceived stimulus may be different from the applied extrinsic stimulus for the case of task-based fMRI data because the estimated stimuli signal will reflect the total impact of the applied extrinsic stimulus and the intrinsic stimulus present, if any. Same has been observed by us as shown later in this chapter. The salient contribution of proposed work is as follows: we propose a dual iterative framework for the joint estimation of HRF and the *perceived stimulus* signal.

7.1 Proposed Formulation

In this section, we present our proposed dual iterative framework to estimate perceived stimuli from mean representative time-series in each ROI [148, 149]. In the proposed dual iterative framework, we first estimate the stimuli signal $\mathbf{s}_i$ followed by estimation of HRF $\mathbf{h}_i$.

After parcellation of brain into AAL ROIs, we averaged time-series of all the voxels within each ROI to generate region-representative time-series. Next, we carry out estimation of HRF and perceived stimulus on each region representative time-series using as a two-step iterative method. We estimate the stimulus signal $\mathbf{s}_i$ assuming a canonical shape of HRF as an initialization followed by estimation of HRF $\mathbf{h}_i$ and repeat this procedure iteratively.

Step-1: Estimation of stimulus signal

The model in (7.1) can be rewritten as

$$\mathbf{y}_i = \mathbf{H}_i\mathbf{s}_i + \xi_i, \tag{7.2}$$

where $\mathbf{H}_i$ is a Toeplitz convolution matrix of dimension MxM corresponding to M length stimuli signal $\mathbf{s}_i$ and L length HRF $\mathbf{h}_i$. The first column of $\mathbf{H}_i$ is filled with HRF padded with M-L zeros at the end.

In both the resting-state and the task-based fMRI data cases, we consider the perceived stimulus signal and its first difference to be sparse. Thus, we impose a fused least absolute shrinkage and selection operator (LASSO) penalty to the stimuli signal. Fused LASSO encourages sparsity on both the coefficients as well as on their successive first differences [150]. Thus, we formulate the problem of estimation of perceived stimulus signal as below:

$$\hat{\mathbf{s}}_i = \underset{\mathbf{s}_i}{argmin}\frac{1}{2}\left\|\mathbf{R}_v(\mathbf{y}_i - \mathbf{H}_i\mathbf{s}_i)\right\|_2^2 + \lambda_1\left\|\mathbf{s}_i\right\|_1 + \lambda_2\sum_{j=2}^{M}\mid s_{i,j} - s_{i,j-1}\mid, \tag{7.3}$$

where $\mathbf{R}_v$ is the decorrelation or the noise whitening matrix resulting from the Cholesky factorization of the inverse of noise covariance matrix $\mathbf{\Gamma}$ ($\mathbf{\Gamma}^{-1} = \mathbf{R}_v^T\mathbf{R}_v$) [151]. λ_1 and λ_2 are the regularization parameters. This problem is difficult to

solve because fused LASSO penalty is non-smooth and nonseparable. Thus, this estimation problem can be reformulated in matrix form as below:

$$\hat{\mathbf{s}}_i = \underset{\mathbf{s}_i}{argmin} \frac{1}{2} \left\| \mathbf{R}_v(\mathbf{y}_i - \mathbf{H}_i\mathbf{s}_i) \right\|_2^2 + \lambda_1 \left\| \mathbf{s}_i \right\|_1 + \lambda_2 \parallel \mathbf{T}\mathbf{s}_i \parallel_1, \tag{7.4}$$

where $\mathbf{T}$ is the first difference matrix operator as given below:

$$\mathbf{T} = \begin{bmatrix} -1 & 1 & 0 & . & . & 0 & 0 \\ 0 & -1 & 1 & 0 & . & . & 0 \\ 0 & 0 & -1 & 1 & 0 & . & . \\ 0 & 0 & . & . & . & . & . \\ . & . & . & . & . & . & 0 \\ 0 & . & . & . & 0 & -1 & 1 \\ 0 & 0 & . & 0 & 0 & 0 & -1 \end{bmatrix} \tag{7.5}$$

Since, in (7.4), $\mathbf{H}_i$ is unknown, we start with an initialization of HRF using the canonical shape of HRF. Next, we require to implement (7.4) that is a convex optimization problem. We utilize ADMM [71] that splits this problem into multiple subproblems that are easier to solve. Next, we explain ADMM based methodology to solve this problem. According to ADMM algorithm, we introduce auxiliary variable $\mathbf{A}$ and $\mathbf{B}$ for the l^1-norm terms in (7.4) [71]. It allows solving (7.4) via splitting it into subproblems. The new objective function is:

$$\underset{\mathbf{s}_i}{min} \frac{1}{2} \left\| \mathbf{y}_i - \mathbf{H}_i\mathbf{s}_i \right\|_2^2 + \lambda_1 \left\| \mathbf{A} \right\|_1 + \lambda_2 \parallel \mathbf{B} \parallel_1 \text{ s.t. } \mathbf{A} = \mathbf{s}_i, \ \mathbf{B} = \mathbf{T}\mathbf{s}_i. \tag{7.6}$$

For simplicity, we have removed term $\mathbf{R}_v$ from above equation. We will expalin more about this in the Results Section. Following [71], the augmented Lagrange function of (7.6) can be written as:

$$\begin{aligned} L(\mathbf{s}_i, \mathbf{A}, \mathbf{Y}_1, \mathbf{Y}_2, \mu, \beta) = \frac{1}{2} \left\| \mathbf{y}_i - \mathbf{H}_i\mathbf{s}_i \right\|_2^2 + \lambda_1 \left\| \mathbf{A} \right\|_1 + \lambda_2 \parallel \mathbf{B} \parallel_1 \\ + \frac{\mu}{2} \left\| \mathbf{A} - \mathbf{s}_i - \mathbf{Y}_1 \right\|_2^2 + \frac{\beta}{2} \left\| \mathbf{B} - \mathbf{T}\mathbf{s}_i - \mathbf{Y}_2 \right\|_2^2, \end{aligned} \tag{7.7}$$

where $\mu, \beta > 0$ are the penalty parameters and $\mathbf{Y}_1$, $\mathbf{Y}_2$ are the Lagrangian multipliers. The above equation consists of three unknown variables $\mathbf{s}_i$, $\mathbf{A}$ and $\mathbf{B}$. ADMM splits the above problem into three subproblems. Each subproblem may be treated as minimization over one variable while fixing other variables. The iterations of ADMM is described in *Algorithm* 5.

Algorithm 5 Solving problem 7.4 by ADMM

Input: voxel time series $\mathbf{y}_i$, parameters λ_1 and λ_2.
Initialize: $\mu_0, \beta_0 = 0.1$, $\mu_{max}, \beta_{max} = 10^{10}$, $\rho_0 = 1.1$, $\mathbf{Y}_{1,0} = \mathbf{Y}_{2,0} = \mathbf{s}_{i,0} = \mathbf{0}$, $\mathbf{A}_0 = \mathbf{B}_0 = \text{randn}(N, N)$, k=0.
while not converge **do**

1: Fix the other variables and update $\mathbf{s}_i$ by

$$\mathbf{s}_{i,k+1} = \underset{\mathbf{s}_i}{argmin}\ \frac{1}{2}\left\|\mathbf{y}_i - \mathbf{H}_i\mathbf{s}_i\right\|_2^2 + \frac{\mu}{2}\left\|\mathbf{A}_k - \mathbf{s}_i - \mathbf{Y}_{1,k}\right\|_2^2 + \frac{\beta}{2}\left\|\mathbf{B}_k - \mathbf{T}\mathbf{s}_i - \mathbf{Y}_{2,k}\right\|_2^2.$$

where solution of above equation is provided in (7.8)

2: Fix the other variables and update $\mathbf{A}$ by

$$\mathbf{A}_{k+1} = \underset{\mathbf{A}}{argmin}\ \lambda_1\left\|\mathbf{A}\right\|_1 + \frac{\mu}{2}\left\|\mathbf{A} - (\mathbf{s}_{i,k+1} + \mathbf{Y}_{1,k})\right\|_2^2.$$

where solution of above equation is provided in (7.9)

3: Fix the other variables and update $\mathbf{B}$ by

$$\mathbf{B}_{k+1} = \underset{\mathbf{B}}{argmin}\ \lambda_2\left\|\mathbf{B}\right\|_1 + \frac{\beta}{2}\left\|\mathbf{B} - (\mathbf{T}\mathbf{s}_{i,k+1} + \mathbf{Y}_{2,k})\right\|_2^2.$$

where solution of above equation is provided in (7.10)

4: Update the multipliers by

$$\mathbf{Y}_{1,k+1} = \mathbf{Y}_{1,k} + \mathbf{s}_{i,k+1} - \mathbf{A}_{k+1}.$$
$$\mathbf{Y}_{2,k+1} = \mathbf{Y}_{2,k} + \mathbf{T}\mathbf{s}_{i,k+1} - \mathbf{B}_{k+1}.$$

5: Check the convergence condition

$$\left\|obj_k - obj_{k-1}\right\|_\infty < 10^{-5}.$$

6: Update $k : k \leftarrow k+1$

end while
Output: $\mathbf{s_i}$

It should be noted that Step 1 in *Algorithm* 5 has a closed form solution

$$\mathbf{s}_{i,k+1} = (\mathbf{H}_i^T\mathbf{H}_i + \mu\mathbf{I} + \beta\mathbf{T}^T\mathbf{T})^{-1}(\mathbf{H}_i^T\mathbf{y}_i + \mu\mathbf{A}_k - \mu\mathbf{Y}_{1,k} + \beta\mathbf{T}^T\mathbf{B}_k - \beta\mathbf{T}^T\mathbf{Y}_{2,k}). \tag{7.8}$$

Step 2 in *Algorithm* 5 can be solved using the soft thresholding as:

$$\mathbf{A}_{k+1} = S_{2\lambda_1/\mu}(\mathbf{s}_{i,k+1} + \mathbf{Y}_{1,k}), \tag{7.9}$$

where $S_\alpha(\mathbf{.})$ is the shrinkage thresholding operator defined in (3.9). Similarly, Step 3 also can be solved using soft thresholding as:

$$\mathbf{B}_{k+1} = S_{2\lambda_2/\beta}(\mathbf{T}\mathbf{s}_{i,k+1} + \mathbf{Y}_{2,k}), \tag{7.10}$$

Step-2: Estimation of HRF

Once the perceived stimulus signal $\mathbf{s}_i$ is estimated, we reformulate (7.1) as below:

$$\mathbf{y}_i = \mathbf{S}_i\mathbf{h}_i + \xi_i, \tag{7.11}$$

where $\mathbf{S}_i$ is a MxL Toeplitz convolution matrix consisting of lagged stimulus estimated covariates. Next, we impose sparsity constraint on $\mathbf{h}_i$ [148, 149]. In addition, HRF is assumed to be a smooth function over time [148, 152]. We incorporate this knowledge into our formulation and apply the Tikhonov regularisation technique for imposing a smoothness constraint on the HRF [152]. This smoothing constraint is imposed by the second difference matrix operator $\mathbf{D}$ defined as:

$$\mathbf{D} = \begin{bmatrix} 2 & -1 & 0 & . & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & . & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & . & . \\ 0 & 0 & . & . & . & . & . \\ . & . & . & . & 2 & -1 & 0 \\ 0 & . & . & . & -1 & 2 & -1 \\ 0 & 0 & . & 0 & 0 & -1 & 2 \end{bmatrix} \tag{7.12}$$

We minimize l^2 norm of $\mathbf{D}\mathbf{h}_i$ in order to impose smoothness. Using the above assumptions, we formulate the problem of estimation of HRF mathematically using Lagrangian multiplier method as below:

$$\hat{\mathbf{h}}_i = \underset{h_i}{argmin} \left\|\mathbf{y}_i - \mathbf{S}_i\mathbf{h}_i\right\|_2^2 + \lambda_3 \left\|\mathbf{h}_i\right\|_1 + \lambda_4 \left\|\mathbf{D}\mathbf{h}_i\right\|_2, \tag{7.13}$$

where λ_3 and λ_4 are the Lagrangian multipliers or the regularization parameters.

Next, we reformulate above equation as:

$$\hat{\mathbf{h}}_i = \underset{\mathbf{h}_i}{argmin} \left\| \begin{pmatrix} \mathbf{y}_i \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{S}_i \\ \sqrt{\lambda_4}\mathbf{D} \end{pmatrix} \mathbf{h}_i \right\|_2^2 + \lambda_3 \left\| \mathbf{h}_i \right\|_1, \tag{7.14}$$

where $\mathbf{0}$ is a zero vector. Above equation reduces to l^1 minimization problem also known as LASSO (least absolute shrinkage and selection operator) regression.

$$\hat{\mathbf{h}}_i = \underset{\mathbf{h}_i}{argmin} \left\| \mathbf{z}_i - \mathbf{Q}_i \mathbf{h}_i \right\|_2^2 + \lambda_3 \left\| \mathbf{W}\mathbf{h}_i \right\|_1, \tag{7.15}$$

Please note that (7.15) is more general form of l^1 minimization problem and can not be solved using soft thresholding. We utilize ADMM to solve (7.15) as it is found to be more efficient for such problems [71]. This method introduces auxiliary variable $\mathbf{C}$ for the second term in (7.15):

$$\underset{\mathbf{h}_i}{min} \left\| \mathbf{z}_i - \mathbf{Q}_i \mathbf{h}_i \right\|_2^2 + \lambda_3 \left\| \mathbf{C} \right\|_1 \text{ s.t. } \mathbf{C} = \mathbf{h}_i, \tag{7.16}$$

The augmented Lagrange function of the above problem is:

$$\begin{aligned} L(\mathbf{h}_i, \mathbf{C}, \mathbf{Y}_3, \gamma) = \left\| \mathbf{z}_i - \mathbf{Q}_i \mathbf{h}_i \right\|_2^2 + \lambda_3 \left\| \mathbf{C} \right\|_1 \\ + \frac{\gamma}{2} \left\| \mathbf{C} - \mathbf{h}_i - \mathbf{Y}_3 \right\|_2^2, \end{aligned} \tag{7.17}$$

where $\gamma > 0$ is a penalty parameter and $\mathbf{Y}_3$ is the Lagrangian multiplier. The above equation consists of two variables $\mathbf{h}_i$ and $\mathbf{C}$. ADMM splits the above problem into two subproblems. Each subproblem may be treated as minimization over one variable while fixing other variable. The iterations of ADMM are described in *Algorithm* 6.

It should be noted that Step 1 in *Algorithm* 6 has a closed form solution

$$\mathbf{h}_{i,k+1} = (\mathbf{Q}_i^T \mathbf{Q}_i + \gamma \mathbf{I})^{-1} (\mathbf{Q}_i^T \mathbf{z}_i + \gamma \mathbf{C}_k - \gamma \mathbf{Y}_{3,k}). \tag{7.18}$$

Step 2 can be solved using soft thresholding

$$\mathbf{C}_{k+1} = S_{2\lambda_3/\gamma}(\mathbf{h}_{i,k+1} + \mathbf{Y}_{3,k}), \tag{7.19}$$

Algorithm 6 Solving problem 7.17 by ADMM

Input: voxel time series $\mathbf{y}_i$, parameter λ_3.
Initialize: $\gamma = 10^{-4}$, $\gamma_{max} = 5$, $\mathbf{Y}_{3,0} = \mathbf{h}_{i,0} = \mathbf{0}$, $\mathbf{C}_0 = \text{randn}(N, N)$, k=0.
while not converge **do**

1: Fix the other variable and update $\mathbf{h}_i$ by

$$\mathbf{h}_{i,k+1} = \underset{\mathbf{h}_i}{argmin} \ \|\mathbf{z}_i - \mathbf{Q}_i\mathbf{h}_i\|_2^2 + \frac{\gamma}{2}\|\mathbf{C} - \mathbf{h}_i - \mathbf{Y}_3\|_2^2 .$$

where the solution of above equation is provided in equ. 7.18

2: Fix the other variables and update $\mathbf{C}$ by

$$\mathbf{C}_{k+1} = \underset{\mathbf{C}}{argmin} \ \lambda_3\|\mathbf{C}\|_1 + \frac{\gamma}{2}\|\mathbf{C} - \mathbf{h}_i - \mathbf{Y}_3\|_2^2 ,$$

where the solution of above equation is provided in (7.19)

3: Update the multipliers by

$$\mathbf{Y}_{3,k+1} = \mathbf{Y}_{3,k} + \mathbf{h}_{i,k+1} - \mathbf{C}_{k+1},$$
$$\gamma = min(\gamma_{max}, 1.1 * \gamma).$$

4: Check the convergence condition

$$\|obj_k - obj_{k-1}\|_\infty < 10^{-5}.$$

5: Update $k : k \leftarrow k+1$

end while
Output: $\mathbf{h_i}$

where $S_\alpha(\mathbf{.})$ is the shrinkage thresholding operator as defined in (3.9).

The above dual iterative framework is repeated iteratively until the shapes of $\hat{\mathbf{s}}_i$ and $\hat{\mathbf{h}}_i$ converge. This is to note that perceived stimuli estimated method is largely valid on the task-based dataset. It is not straight forward to estimate stimuli signal for the case of rest fMRI data due to noise involved. Perhaps, one needs to use different regularization terms altogether to account for noise level present in rest fMRI data. This requires further elaborate evaluation and discussion on the rest data.

7.2 Results

This is to note that the goal of this work is to estimate the underlying stimulus along with HRF. Thus, it is important that the ground truth is known for the purpose of the validation of the proposed method. In this work, we test the proposed method on the synthetic data constructed with a known HRF and known stimuli signal with varying duration and varying onset times. For the real fMRI data, block stimuli signal is used as the ground truth. This implies that we know ground truth of the stimuli signal in the real data too for the purpose of validation of the proposed method, although HRF is unknown in the real data.

Results on Synthetic fMRI Data

We generated synthetic fMRI time series by convolving stimuli signal with the canonical HRF. We designed canonical HRF of length $L = 32$ using the difference of two gamma functions [24]. The shape of this HRF is shown in Fig. 7.1 (a).

In order to assess the proposed framework, we test our algorithm on the above synthetic data where activity signal is generated with 5 ON periods of duration 6s, 5s, 10s, 3s, and 1s with onsets at 10s, 40s, 100s, 140s, 180s, respectively. We generated 200 time points of the synthetic BOLD fMRI signal as below:

$$\mathbf{y} \equiv y[n] = s[n] \otimes h[n] + \xi[n] \tag{7.20}$$

For the sake of simplicity, noise is assumed to be white. Additive white Gaussian noise is generated with variances 0.75, 0.5, 0.25, 0.1, and 0.05. For computing the mean square error (MSE), 500 Monte Carlo cycles have been performed over voxel time-series (i.e., considering 500 different realizations of noise time-series). MSE between the canonical and estimated HRF is calculated as below:

$$MSE = \frac{1}{500} \sum_{k=1}^{500} \left[\frac{1}{L} \sum_{n=0}^{L-1} (\hat{h}_k[n] - h_k[n])^2 \right] \tag{7.21}$$

The estimated HRF and the estimated stimuli signal for noise variance of $\sigma^2 = 0.1$ and the regularization parameters of $\lambda_1 = 0.001, \lambda_2 = 0.3, \lambda_3 = 1$, and $\lambda_4 = 0.7$ are shown in Fig. 7.1(a) and Fig. 7.1(b), respectively. These parameters are deter-

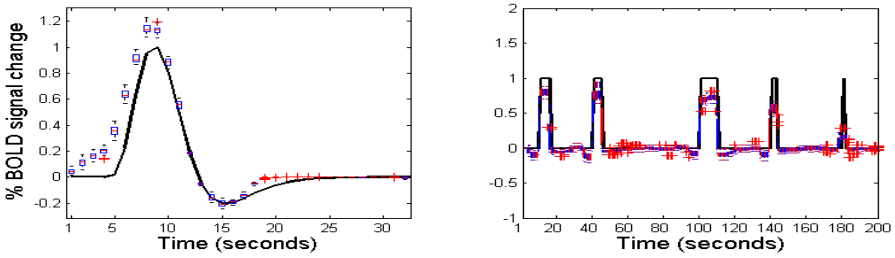


Figure 7.1: (a): Estimated HRF; (b) Estimated Activity Signal

Table-2: MSE calculated between estimated and the actual HRF (that is used in the synthetic data)

	Noise Variance σ^2				
	0.05	0.1	0.25	0.5	0.75
MSE using the Proposed method	0.0151	0.0159	0.0171	0.0183	0.0220

mined empirically where the MSE is minimum. Voxel time series is generated using canonical HRF and the algorithm was initialized using the shape of *Daubechies*-4 (or *db*4) scaling function.

From Fig. 7.1(b), we observe that the proposed method is able to extract stimuli signal with variable onset and duration. The MSE results on the estimated HRF using the proposed algorithm are tabulated in Table-2. Fig.7.2 shows the receiver operating characteristics (ROC) curve on the estimated activity signal. It is observed that the performance of the proposed method is satisfactory on the synthetic data.

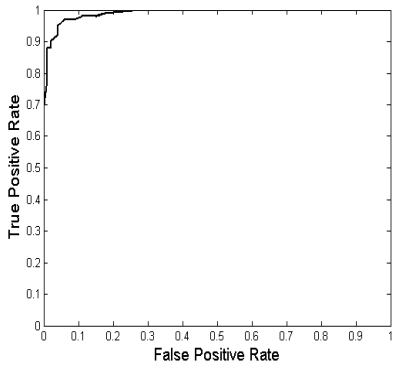


Figure 7.2: ROC curve for single voxel time series with σ^2=0.25

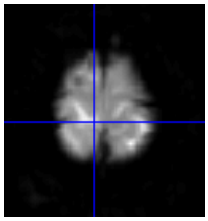


Figure 7.3: Seed voxel at (40, 43, 66)

Results on Real fMRI Data

In this section, we present results on real fMRI data acquired from one subject performing a right hand sense task in 3-T MR scanner. This dataset consists of an acquisition of 36 contiguous slices with $128 \times 128 \times 36$ voxels of voxel size $4 \times 4 \times 4$ mm^3. 100 brain volumes with repetition time of 3 s are acquired. The task involved is an example of block design paradigm starting with 10 volumes of rest followed by 10 volumes of activity, and so on. Data is preprocessed using SPM8 toolbox [25]. Pre-processing steps include realignment (with the first scan for removal of motion artifact), slice time correction (with the first slice of each volume), and normalization (with the MNI atlas). Resultant fMRI data had 100 scan points of $79 \times 95 \times 68$ voxels each. We did not use smoothing in preprocessing as our algorithm inherently enhances signal to noise ratio (SNR). We discarded first 12 dummy scans, resulting in 88 brain volumes. We also did detrending of real fMRI data prior to extraction of underlying stimuli signal. This helps in removing the trend from fMRI data and brings it to baseline.

First, we present result on the seed voxel which lies in somatosensory region of the brain and is supposed to be active in right hand sense task. In general, Brodmann regions 1, 2, 3 are found to be associated with somatosensory region [26]. Union of these Broadmann regions is extracted using the WFU Pickatlas Tool in Matlab [27]. Then, seed voxel was extracted in this region using our recent work [21]. The coordinates of the seed voxel were found to be (40, 43, 66). Fig. 7.3 depicts this seed voxel on the corresponding axial brain slice.

Next, we follow the procedure outlined in Table-1 and estimate both the HRF and the stimuli signal. We start our algorithm with the canonical HRF shape as shown in Fig. 7.1(a). The regularization parameters are set to the same values as that used in the synthetic fMRI experiment. Values of ρ in the range of 0 to 1

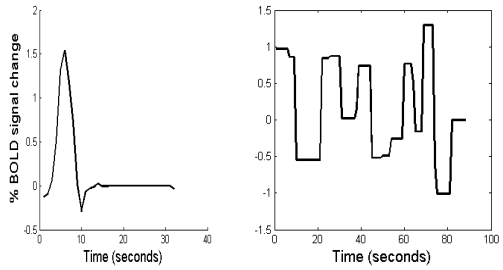


Figure 7.4: (a): Estimated HRF; (b) Estimated Activity Signal at voxel (40, 43, 66) using the proposed method

are tested for an initialization of $\mathbf{R}_v$. Empirically ρ =0.1 is selected for whitening of AR(1) noise on voxel time series. Fig. 7.4 shows the estimated HRF and the estimated activity signal using the proposed method. The estimated activated signal in Fig. 7.4(b) is having a block nature. Since first 12 scans were dummy, we observe estimated signal to start from activity (block) followed by the rest block. Although the applied stimulus is in uniform blocks of 10 rest and 10 activity, the estimated activity signal represents the perceived activity stimulus by brain and hence, has a slightly varied shape. This is to note that, in [21], we proposed a method for seed voxel detection using known activity signal. Thus, estimated activity signal at that voxel should closely resemble the applied stimuli which is indeed the case as observed from Fig. 7.4b. This experiment shows that our framework of joint estimation in this chapter is robust and reliable.

Next, for the sake of completeness, we test our algorithm on other active voxels. Voxel with coordinates [41, 45, 66] is extracted based on highest norm of voxel time-series lying in somatosensory region. Fig. 7.5 shows the estimated HRF and the estimated activity signal using the proposed method. We note that on this voxel, the perceived activity stimulus by brain differs from applied stimuli in a greater manner compared to the seed voxel. In future, we will extend the proposed method to resting-state data for the detection of intrinsic activity signal in order to build robust resting-state networks.

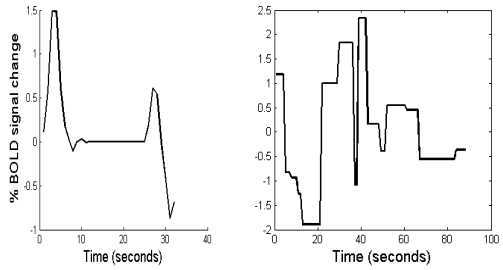


Figure 7.5: (a): Estimated HRF; (b) Estimated Activity Signal at voxel (41, 45, 66) using the proposed method

7.3 Conclusions

In this Chapter, we have introduced a joint iterative framework for the estimation of hemodynamic response function (HRF) and the underlying activity signal. This estimation is based on two-stage iterative method. The proposed framework estimates voxel-wise HRF via imposing constraints of sparsity in the wavelet-domain and smoothness in time-domain on HRF. The activity signal is estimated using fused LASSO penalty which imposes sparsity on coefficients as well as on first difference of the activity signal. The proposed method is observed to perform satisfactorily.

Chapter 8

Brain networks reproducibility from compressively sensed fMRI data

Despite the advancements in fMRI scanners, one of the biggest limitations of fMRI modality is slow imaging compared to the other medical imaging modalities [153] that leads to annoyance in patients resulting in low SNR due to subject's movement [154]. Hence, there is a need to capture images in the shortest possible time. Various remedies have been proposed to this end such as development of high magnetic field scanner [155, 156, 157, 158], coil sensitivity improvement inside fMRI scanner [159], advancements in pulse sequences [160, 161], usage of parallel imaging [162, 163], and compressed sensing (CS) based fMRI reconstruction from fewer k-space (spatial Fourier domain) measurements [164, 165, 166, 167, 168, 169, 170, 171, 172, 173].

In recent years, CS is gaining a lot of interest in the area of accelerated fMRI reconstruction [164, 174, 170, 175, 173]. It allows reconstruction of fMRI data using smaller number of k-space measurements that are picked up at a sampling rate below the required Nyquist sampling frequency [176]. This reconstruction using lesser samples leads to reduction in scanning time because scanning time is directly related to the number of sampling measurements.

In this chapter, we are interested in CS based fMRI reconstruction. The rest of this chapter is organized as follows. Section-8.1 discuss problem formulation of fMRI reconstruction and presents brief literature review thereafter. Section-8.2 presents the proposed reconstruction methods. In the results section 8.3, simu-

lation results using all proposed methods and some of the existing methods are presented on fMRI dataset. Finally, conclusions are presented in the last section.

8.1 Literature Review

In this section, we present the mathematical formulation of fMRI reconstruction problem. A fMRI data is a 4-dimensional data with 3D brain volume captured over number of time points. This data is represented as belonging to 4D space $\mathbb{R}^{n_x \times n_y \times n_z \times T}$ where n_z is the number of brain slices (or images) along z-axis with each slice of size $n_x \times n_y$. In slice-by-slice fMRI reconstruction, Casorati matrix [177] is formed corresponding to all time points of each slice such that the size of the matrix is $n \times T$, where $n = n_x \times n_y$ is the number of voxels in each brain slice and T is the number of brain volumes. Thus, Casorati matrix is formed by stacking one fMRI slice over each time point.

Let us consider one such matrix $\mathbf{X}$ corresponding to a single slice captured over T time points. In CS-based reconstruction, undersampled fMRI data is acquired in k-space with a sampling rate below the Nyquist rate. The fMRI reconstruction problem from undersampled k-space can be represented as follows:

$$\mathbf{Y} = \mathbf{\Phi F X} + \xi, \tag{8.1}$$

where $\mathbf{Y}$ denotes the compressively sensed $k - t$ space data, $\mathbf{F}$ denotes the 2-D Fourier transform operator applied on Casorati matrix $\mathbf{X}$ of one slice, $\mathbf{\Phi}$ is the sensing matrix that contains partial measurements of $k - t$ space, and $\xi \in \mathbb{R}^{n \times T}$ denotes the measurement noise. The aim of fMRI reconstruction problem is to recover Casorati data matrix $\mathbf{X}$, given partial Fourier measurements $\mathbf{Y}$ and the sensing matrix $\mathbf{\Phi}$. Reconstruction is done independently for all n_z brain slices.

Researchers have proposed various methods for efficient reconstruction from undersampled k-space measurements $\mathbf{Y}$ [170, 171, 175, 168, 173, 164, 174]. These methods largely rely on reconstructing data using an optimization framework under certain constraints on $\mathbf{X}$. The task of computing $\mathbf{X}$ from $\mathbf{Y}$ is an underdetermined inverse problem. It is not possible to solve (8.1) by computing the inverse because the sensing matrix $\mathbf{\Phi}$ in the forward equation is usually ill-conditioned due

to large undersampling. Moreover, solving a set of under-determined equations leads to infinitely many solutions. Therefore, in order to recover a unique solution, regularization terms are added. Hence, in general the problem needs to be regularized to find a solution. A relatively simple solution is a well known sparsity regularization [170].

Sparse recovery methods assume the desired signal to be sparse over some known *apriori* transform basis $\mathbf{\Psi}$ and hence, l^1 norm in the corresponding domain is used as regularization to recover the signal. l^1 norm is used as a surrogate for standard sparsity inducing l^0 norm because regularization over l^0 norm is a non-deterministic polynomial (NP) hard problem. Thus, using sparsity regularization, fMRI reconstruction problem can be formulated as

$$\hat{\mathbf{X}} = arg\ \underset{\mathbf{X}}{min} \left\|\mathbf{Y} - \mathbf{\Phi F X}\right\|_F^2 + \lambda_1 \left\|\mathbf{\Psi X}\right\|_1, \tag{8.2}$$

where $\left\|\cdot\right\|_F^2$ denotes the Frobenius norm that is defined as $\left\|\mathbf{Y} - \mathbf{\Phi F X}\right\|_F^2 = Tr[(\mathbf{Y} - \mathbf{\Phi F X})^T(\mathbf{Y} - \mathbf{\Phi F X})]$, $\mathbf{\Psi}$ denotes the sparsifying transform basis, $\left\|\cdot\right\|_1$ is the l^1 norm, and λ_1 is the regularization parameter that governs the sparsity on $\mathbf{X}$ over the $\mathbf{\Psi}$ basis. The first term in (8.2) is the data fidelity term that minimizes the variance of noise ξ, while the second term is the sparsity promoting term. Often, fMRI data $\mathbf{X}$ is assumed to be sparse in some transform domain. For example, in [170], fMRI data is reconstructed using CS by adding sparsity on the fMRI data $\mathbf{X}$ in the wavelet domain using Daubechies wavelet. Here, sparsity is being used in the temporal direction i.e. along columns of matrix $\mathbf{X}$. Recently, in [171], both temporal and spatial sparsity are exploited to recover fMRI data to gain high spatial resolution fMRI and method is named as High Spatial Resolution Compressed Sensing (HSPARSE) [171]. More details regarding these methods is being provided in section 8.3.

In general, fMRI data matrix $\mathbf{X}$, i.e., one brain slice data stacked over time, is observed to be low rank. Hence, low rank constraint can be imposed in the CS based optimization framework to recover fMRI data slice by slice. In [175], CS is utilized for fMRI reconstruction by adding rank deficiency as a constraint. It is assumed that the fMRI data is low rank having a small number of significant singular values. Accordingly, iterative hard thresholding algorithm is used to recover

low-rank matrix.

In another work [173], fMRI reconstruction is performed using low-rank plus sparse (LR+S) decomposition of the fMRI data. Here, fMRI data is represented as a linear combination of low rank and sparse components as below:

$$\mathbf{X} = \mathbf{L}(low\ rank) + \mathbf{S}(sparse). \tag{8.3}$$

Further, low rank and sparse components are reconstructed separately and iteratively using low rank and sparsity optimization constraints respectively.

In the nutshell, many offline reconstruction methods, such as compressed sensing with wavelet domain sparsity (CSWD) [170], compressed sensing with both temporal and spatial domain sparsity [171], compressed sensing with low rank constraint [172], and compressed sensing with low rank plus sparse decomposition [173], have been proposed in the fMRI literature. These methods largely differ in regularization constraints in CS fMRI reconstruction framework.

8.2 Proposed Formulations

Double Temporal Sparsity based Reconstruction (DTSR) method

In this section, we introduce a novel fMRI reconstruction method. In fMRI, same brain volumes are scanned repeatedly over time in order to study brain's function. Consider the Casorati matrix $\mathbf{X}$ for any brain slice, where each column $\mathbf{x}_t$ represents the vectorized brain slice captured at the t^{th} time point. Since a brain slice over adjacent time points may contain less amplitude changes, the difference of adjacent columns of this Casorati matrix exhibits strong sparsity. This fact is used as an advantage in the proposed method via total variation based regularization [178] because scanning of the same brain volume over time brings similarity in the temporal direction. This form of difference sparsity is also known as total variation in the case of 1-D signal recovery [178]. We call this sparsity in the context of fMRI as total variation temporal sparsity because it exploits sparsity in the temporal direction. In addition, we impose conventional temporal sparsity in

the proposed reconstruction framework. In other words, we impose two l^1-norm constraints. First l^1-norm sparsity is imposed on the voxel time series (temporal data) in the transformed domain (i.e. columns of $\mathbf{X}$) and second l^1 norm sparsity is imposed on the successive difference of the same temporal data. Hence, we name the proposed method as Double Temporal Sparsity based Reconstruction (DTSR) method [74].

In the total variation temporal sparsity, the difference matrix of $\mathbf{X}$ is assumed to be sparse. This difference matrix is formed by performing the first difference on the consecutive columns of $\mathbf{X}$. First, differencing is performed from 2^{nd} column onwards. Thus, the DTSR objective function can be formulated as

$$\hat{\mathbf{X}} = arg\ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 + \lambda_1 \|\mathbf{\Psi X}\|_1 + \lambda_2 \sum_{t=2}^{T} |\mathbf{x}_t - \mathbf{x}_{t-1}|\,, \tag{8.4}$$

where λ_1 and λ_2 are non-negative regularization parameters. Since second regularization term in (8.4) is non-differentiable, it is not easy to solve DTSR in this formulation. Thus, (8.4) is reformulated below with matrix version that provides efficient solution to this problem.

$$\hat{\mathbf{X}} = arg\ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 + \lambda_1 \|\mathbf{\Psi X}\|_1 + \lambda_2 \|\mathbf{XD}\|_1\,. \tag{8.5}$$

where $\mathbf{D}$ performs first order differencing on the successive columns of the given matrix $\mathbf{X}$ and is defined as:

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & . & . & 0 & 0 \\ 0 & -1 & 1 & 0 & . & . & 0 \\ 0 & 0 & -1 & 1 & 0 & . & . \\ 0 & 0 & . & . & . & . & . \\ . & . & . & . & . & . & 0 \\ 0 & . & . & . & 0 & -1 & 1 \\ 0 & 0 & . & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Implementation

In this subsection, an algorithm is designed to solve (8.5) using the ADMM [71]. Following [71], two auxiliary matrices $\mathbf{W} \in \mathbb{R}^{n\times T}$ and $\mathbf{Z} \in \mathbb{R}^{n\times T}$ are introduced in

(8.5) as

$$\hat{\mathbf{X}} = arg\ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 + \lambda_1 \|\mathbf{W}\|_1 + +\lambda_2 \|\mathbf{Z}\|_1 \quad s.t.\ \ \mathbf{W} = \mathbf{\Psi X},\ \mathbf{Z} = \mathbf{XD}. \tag{8.6}$$

In addition to this, constraints with equality are added for each of the auxiliary matrices. Thus, the new objective function is written as:

$$arg \min_{\mathbf{X},\mathbf{Z},\mathbf{W}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 + \lambda_1 \|\mathbf{W}\|_1 + \lambda_2 \|\mathbf{Z}\|_1 + \frac{\eta_1}{2} \|\mathbf{W} - \mathbf{\Psi X} - \mathbf{B_1}\|_F^2 + \frac{\eta_2}{2} \|\mathbf{Z} - \mathbf{XD} - \mathbf{B_2}\|_F^2, \tag{8.7}$$

where η_1 and η_2 are penalty parameters and $\mathbf{B}_1$ and $\mathbf{B}_2$ are the Lagrange multipliers used to enforce equality between the original and auxiliary matrices. ADMM updates variables $\mathbf{W}$, $\mathbf{Z}$, and $\mathbf{X}$ alternately in the above defined objective function. The minimization over one variable in an iteration assumes the other two variables to be fixed. Therefore, the above function can be alternately optimized over each variable separately by splitting it into different subproblems as stated in Algorithm 7.

W and Z Subproblems

The first two subproblems in Algorithm 7 are l^1 minimization problems. For any l^1 minimization problem such as

$$\min_{\mathbf{P}} \alpha \|\mathbf{P}\|_1 + \frac{\beta}{2} \|\mathbf{P} - \mathbf{Q}\|_F^2, \tag{8.8}$$

where $\mathbf{P}, \mathbf{Q} \in \mathbf{R}^{n \times T}$ and $\alpha, \beta > 0$, the solution is [73]

$$\mathbf{P} = Soft(\mathbf{Q}, 2\frac{\alpha}{\beta}\mathbf{A}), \tag{8.9}$$

where $\mathbf{A}$ is a matrix containing all ones and $\mathbf{Q}$ on the right hand side in the above equation is an initial estimate of $\mathbf{P}$. The definition of '*Soft*' is

$$Soft(\mathbf{Q}, \nu\mathbf{A}) = sgn(\mathbf{Q}) \otimes max\left\{0, |\mathbf{Q}| - \nu\mathbf{A}\right\}, \tag{8.10}$$

Algorithm 7 Pseudo code of the proposed DTSR method

1: Initialize λ_1, λ_2, η_1, η_2, $\mathbf{B}_1^0$, $\mathbf{B}_2^0$, $\mathbf{X}^0$, j=1
2: **while** convergence criteria not met **do**
3: $\mathbf{W}$-subproblem

$$\mathbf{W}^j = arg\ \min_{\mathbf{W}}\ \lambda_1 \|\mathbf{W}\|_1 + \frac{\eta_1}{2}\left\|\mathbf{W} - \mathbf{\Psi X}^{j-\mathbf{1}} - \mathbf{B}_{\mathbf{1}}^{j-\mathbf{1}}\right\|_F^2.$$

4: $\mathbf{Z}$-subproblem

$$\mathbf{Z}^j = arg\ \min_{\mathbf{Z}}\ \lambda_2 \|\mathbf{Z}\|_1 + \frac{\eta_2}{2}\left\|\mathbf{Z} - \mathbf{X}^{j-\mathbf{1}}\mathbf{D} - \mathbf{B}_{\mathbf{2}}^{j-\mathbf{1}}\right\|_F^2.$$

5: $\mathbf{X}$-subproblem

$$\mathbf{X}^j = arg\ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 + \frac{\eta_1}{2}\left\|\mathbf{W}^j - \mathbf{\Psi X} - \mathbf{B}_{\mathbf{1}}^{j-\mathbf{1}}\right\|_F^2 + \frac{\eta_2}{2}\left\|\mathbf{Z}^j - \mathbf{XD} - \mathbf{B}_{\mathbf{2}}^{j-\mathbf{1}}\right\|_F^2.$$

6: Lagrange multipliers update

$$\mathbf{B}_1^j = \mathbf{B}_1^{j-1} + \mathbf{\Psi X}^j - \mathbf{W}^j.$$
$$\mathbf{B}_2^j = \mathbf{B}_2^{j-1} + \mathbf{X}^j\mathbf{D} - \mathbf{Z}^j.$$

7: j=j+1
8: **end while**

where $\nu = \frac{\alpha}{\beta}$, $\otimes$ denotes the element-wise product, and $|\mathbf{Q}|$ denotes a matrix with absolute values of $\mathbf{Q}$. $\mathbf{A}$ in the above equation ensures soft thresholding on all elements of $\mathbf{Q}$. For the nonzero elements of $\mathbf{Q}$, $sgn(\mathbf{Q}) = \mathbf{Q}./|\mathbf{Q}|$, otherwise $\text{sgn}(\mathbf{Q}) = 0$. Hence, the closed form solution of $\mathbf{W}$ at the iteration number j is

$$\mathbf{W}^j = Soft((\mathbf{\Psi X}^{j-1} + \mathbf{B}_1^{j-1}), 2\frac{\lambda_1}{\eta_1}\mathbf{A}). \tag{8.11}$$

Once $\mathbf{W}$ at iteration j is estimated, the next step is to estimate $\mathbf{Z}$. Using $\mathbf{X}^{j-1}$ and $\mathbf{B}_2^{j-1}$ from the previous iteration, $\mathbf{Z}^j$ can be obtained via a closed form as

$$\mathbf{Z}^j = Soft((\mathbf{X}^{j-1}D + \mathbf{B}_2^{j-1}), 2\frac{\lambda_2}{\eta_2}\mathbf{A}). \tag{8.12}$$

X Subproblem

This subproblem is quadratic. It can be efficiently solved using the conjugate gradient algorithm. We used the line search conjugate gradient algorithm as used in [179]. This is an iterative method where, first, a descent direction is decided based on the minimization of objective function. Next, a step size is computed that determines how far the unknown variable of interest should move along the descent direction. Unlike a fixed step size in the normal gradient descent, this method performs a line search in order to find the best step size. In general, it has been observed that for convex quadratic objective functions, the line search conjugate gradient method achieves a finite convergence.

Update of Lagrange Multiplier Variables

Last step is the update of Lagrange multipliers that is explained in Algorithm 7. Lagrange multipliers help in achieving convergence in the subsequent iterations. In this algorithm, convergence is checked either by comparing convergence of the objective function in (8.4) with a threshold or with the maximum number of iterations reached.

Matrix Completion with Sparse Recovery (MCwSR) Reconstruction Method

In the proposed DTSR algorithm, we assume adjacent brain slices to be similar, although this assumption might not hold true for task based fMRI data. Thus, instead of utilizing total variation temporal sparsity, we apply combination of low rank and sparsity in CS based fMRI reconstruction. This method is applicable for reconstructing both resting-state and task-based fMRI data. We name the proposed method as Matrix completion with Sparse Recovery (MCwSR) and formulate the proposed MCwSR fMRI reconstruction problem as below:

$$\hat{\mathbf{X}} = arg\ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 + \mu_1 \|\mathbf{X}\|_* + \mu_2 \|\mathbf{\Psi X}\|_1 , \tag{8.13}$$

where $'*'$ denotes the nuclear norm which is the l^1 norm of singular values. μ_1 and μ_2 are the regularization parameters. The two regularization terms in equ. (8.13)

imply sparsity and low rank constraints on Casorati matrix $\mathbf{X}$.

Implementation

We utilize ADMM algorithm to solve (8.13). We introduce two proxy variables $\mathbf{W}$ and $\mathbf{Z}$ in this equation as

$$\begin{aligned} arg\min_{\mathbf{X},\mathbf{Z},\mathbf{W}} \|\mathbf{Y}-\mathbf{\Phi F X}\|_F^2 + \mu_1 \|\mathbf{W}\|_* + \mu_2 \|\mathbf{\Psi Z}\|_1 + \\ \mathbf{s.t.}\ \mathbf{W}=\mathbf{X} \text{ and } \mathbf{Z}=\mathbf{X}. \end{aligned} \tag{8.14}$$

Here, $\mathbf{W}$ and $\mathbf{Z}$ are proxy variables to the variable of interest $\mathbf{X}$. Following [71], we substitute equality constraints for each of the proxy variables to split (8.14) and introduce Bregman variables $\mathbf{B}_1$ and $\mathbf{B}_2$ as below:

$$\begin{aligned} arg\min_{\mathbf{X},\mathbf{Z},\mathbf{W}} \|\mathbf{Y}-\mathbf{\Phi F X}\|_F^2 + \mu_1 \|\mathbf{W}\|_* + \mu_2 \|\mathbf{\Psi Z}\|_1 + \\ \frac{\eta_1}{2}\|\mathbf{W}-\mathbf{X}-\mathbf{B_1}\|_F^2 + \frac{\eta_2}{2}\|\mathbf{Z}-\mathbf{X}-\mathbf{B_2}\|_F^2, \end{aligned} \tag{8.15}$$

where η_1 and η_2 are regularization parameters and, $\mathbf{B}_1$ and $\mathbf{B}_2$ are the Lagrange multipliers used to enforce equality between original and proxy variables. The above equation consists of three variables $\mathbf{W}$, $\mathbf{Z}$, and $\mathbf{X}$ along with two more variables, $\mathbf{B}_1$ and $\mathbf{B}_2$. We split the above problem into three subproblems. Each subproblem may be treated as minimization over one variable while fixing other variables as shown in Algorithm 8. The solution of each subproblem in Algorithm 8 is explained in the following subsections.

W subproblem

The first subproblem is nuclear norm minimization of matrix $\mathbf{W}$ as shown in Algorithm 8. We solve this subproblem using soft thresholding [83]. Soft thresholding is found to be one of the best method among the many existing algorithms [83, 180] for low rank matrix recovery. The solution of this subproblem is summarized in Algorithm 9.

Algorithm 8 Pseudo code of proposed MCwSR method

1: Intialize μ_1, μ_2, η_1, η_2, $\mathbf{B}_1^0$, $\mathbf{B}_2^0$, $\mathbf{X}^0$, j=1
2: **while** convergence criteria not met **do**
3: **W**-subproblem

$$\mathbf{W}^j = arg\ \underset{\mathbf{W}}{min}\ \mu_1 \left\|\mathbf{W}\right\|_* + \frac{\eta_1}{2}\left\|\mathbf{W} - \mathbf{X}^{j\text{-}1} - \mathbf{B}_{\mathbf{1}}^{j\text{-}1}\right\|_F^2.$$

4: **Z**-subproblem

$$\mathbf{Z}^j = arg\ \underset{\mathbf{Z}}{min}\ \mu_2 \left\|\mathbf{\Psi Z}\right\|_1 + \frac{\eta_2}{2}\left\|\mathbf{Z} - \mathbf{X}^{j\text{-}1} - \mathbf{B}_{\mathbf{2}}^{j\text{-}1}\right\|_F^2$$

5: **X**-subproblem

$$\mathbf{X}^j = arg\ \underset{\mathbf{X}}{min}\ \left\|\mathbf{Y} - \mathbf{\Phi F X}\right\|_F^2 + \frac{\eta_1}{2}\left\|\mathbf{W}^j - \mathbf{X} - \mathbf{B}_{\mathbf{1}}^{j\text{-}1}\right\|_F^2 + \frac{\eta_2}{2}\left\|\mathbf{Z}^j - \mathbf{X} - \mathbf{B}_{\mathbf{2}}^{j\text{-}1}\right\|_F^2.$$

6: **Bregman variable update**

$$\mathbf{B}_1^j = \mathbf{B}_1^{j-1} + \mathbf{X}^j - \mathbf{W}^j.$$
$$\mathbf{B}_2^j = \mathbf{B}_2^{j-1} + \mathbf{X}^j - \mathbf{Z}^j.$$

7: *j=j+1*
8: **end while**

Algorithm 9 Pseudo code of **W**-subproblem in Algorithm 8

1: Intialize $\mathbf{W}^j = \mathbf{X}^{j-1} + \mathbf{B}_1^{j-1}$.
2: Compute singular value decomposition (SVD) of $\mathbf{W}^j = \mathbf{USV}^T$, where $\mathbf{U}$ and $\mathbf{V}$ are the matrices containing left and right singular vectors, respectively, and the matrix $\mathbf{S}$ contains the singular values.
3: Soft thresholding is applied on the singular values contained on diagonal of $\mathbf{S}$ as

$$\mathbf{\Gamma} = Soft(\mathbf{S}, \frac{\mu_1}{\eta_1}\mathbf{I}) = sgn(\mathbf{S}) \otimes max\left\{0, |\mathbf{S}| - \frac{\mu_1}{\eta_1}\mathbf{I}\right\},$$

where $\otimes$ denotes the element-wise product, $|\mathbf{S}|$ denotes absolute values of matrix $\mathbf{S}$ and $\mathbf{I}$ is identity matrix. $\mathbf{I}$ in the above equation ensures soft thresholding only on diagonal elements of $\mathbf{S}$. For the nonzero elements of $\mathbf{S}$, $sgn(\mathbf{S}) = \mathbf{S}./\left|\mathbf{S}\right|$, otherwise sgn($\mathbf{S}$) = 0.
4: Next, $\mathbf{W}^j$ is updated with updated singular values and older singular matrices

$$\mathbf{W}^j = \mathbf{U\Gamma V}^T,$$

SVD of the matrix $\mathbf{W}^j$ is required to be computed as in step 2 of Algorithm 9. Since direct SVD computation is time intensive due to large size of the observation matrix, a different strategy is adopted to minimize computational complexity. Instead of direct SVD computation, singular values and right singular vectors are determined using eigen decomposition of $(\mathbf{W}^j)^H \mathbf{W}^j$ as below:

$$\left(\mathbf{W}^j\right)^H \mathbf{W}^j = \mathbf{V}\mathbf{S}^2\mathbf{V}^T, \tag{8.16}$$

where H denotes matrix Hermitian transpose. Left singular vectors $\mathbf{U}$ are determined from $\mathbf{W}^j = \mathbf{U}\mathbf{S}\mathbf{V}^T$ where $\mathbf{W}^j$, singular values matrix $\mathbf{S}$, and the right singular vectors' matrix $\mathbf{V}$ are known. This completes the SVD computation of matrix $\mathbf{W}^j$.

Z subproblem

The second subproblem is analysis prior l^1 minimization problem and the closed form solution at iteration j is

$$\mathbf{Z}^j = \mathbf{\Psi}^H(Soft(\mathbf{\Psi}(\mathbf{X}^{j-1} + \mathbf{B}_2^{j-1}), \frac{\mu_2}{\eta_2}\mathbf{A})), \tag{8.17}$$

where $\mathbf{\Psi}^H$ denotes the Hermitian transpose of the sparsifying basis $\mathbf{\Psi}$ and the definition of '*Soft* is provided in (8.10).

X subproblem

With fixed $\mathbf{W}$ and $\mathbf{Z}$, this subproblem is quadratic as shown in Algorithm 8. It can be solved using conjugate gradient algorithm [181]. Last step in Algorithm 8 is the update of Lagrange multipliers that is explained in Algorithm 8.

OptShrink LR+S Method

In this section, we are interested in accelerated fMRI data reconstruction using low rank plus sparse decomposition. To this end, we first elucidate low rank plus sparse (LR+S) reconstruction problem. This method reconstructs fMRI data using low

rank and sparse matrix decomposition as:

$$\hat{\mathbf{L}}, \hat{\mathbf{S}} = arg \min_{\mathbf{L},\mathbf{S}} \|\mathbf{Y} - \mathbf{\Phi}\mathbf{F}(\mathbf{L}+\mathbf{S})\|_F^2 + \gamma_1 \|\mathbf{L}\|_* + \gamma_2 \|\mathbf{\Psi}\mathbf{S}\|_1 , \tag{8.18}$$

where γ_1 and γ_2 are regularization parameters. The fMRI data matrix $\mathbf{X}$ is reconstructed as:

$$\hat{\mathbf{X}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}. \tag{8.19}$$

Here, matrix $\mathbf{X}$ is estimated as the superposition of low rank $\mathbf{L}$ and sparse matrices $\mathbf{S}$. Hence, aim is to recover matrices $\mathbf{L}$ and $\mathbf{S}$, given a set of undersampled measurements $\mathbf{Y}$ and the corresponding measurement matrix $\mathbf{\Phi}$. In (8.18), $\| \, . \, \|_*$ is the nuclear norm which is used to solve low rank matrix recovery problem and l^1 norm on $\mathbf{S}$ is used to obtain sparse solution [176, 182].

Recently in [183], non convex approach of low rank matrix recovery is proposed which is observed to be perform better than the nuclear norm. They name their proposed approach as optimal singular value shrinkage (OptShrink), a data driven method, recently used for denoising of low rank matrix [184]. The OptShrink method requires noisy low rank matrix and its rank estimate as input and provides denoised low rank matrix estimate. This motivates us to explore a non convex surrogate of low rank matrix recovery in LR+S fMRI reconstruction problem. Hence, we call the proposed method as *Optshrink* LR+S method. Complete pseudo code to solve *Optshrink* LR+S is being provided in Algorithm 10, wherein solution of each subproblem is mentioned in next section.

Implementation

In this subsection, we explain the proposed *Optshrink* LR+S implementation details wherein, it is solved by breaking it into two subproblems of estimating $\mathbf{L}$ and $\mathbf{S}$ as described next.

S subproblem

Soft thresholding is used to solve l^1 norm penalty on $\mathbf{S}$ as:

$$\hat{\mathbf{S}} = Soft(\mathbf{\Psi}\mathbf{S}, \gamma_2), \tag{8.20}$$

Algorithm 10 Pseudo code of *Optshrink LR+S* method

1: Intialize γ_2, $\mathbf{X}^0$, tolerance=10^{-5}, $\mathbf{\Psi}$, j=0, $\mathbf{L}^0 = \mathbf{X}^0, \mathbf{S}^0 = \mathbf{0}$
2: Inputs $\mathbf{Y}$, $\mathbf{\Phi}$
3: Outputs $\hat{\mathbf{L}}, \hat{\mathbf{S}}, \hat{\mathbf{X}}$
4: **while** (obj(j)-obj(j-1)<tolerance) & (j>2) **do**
5: **S**-subproblem

$$\hat{\mathbf{S}}^j = arg\ \min_{\mathbf{S}} \quad \left\| \mathbf{Y} - \mathbf{\Phi F}\hat{\mathbf{X}}^{\mathbf{j-1}} \right\|_F^2 + \gamma_2 \left\| \mathbf{\Psi S} \right\|_1 .$$

6: **L**-subproblem

$$\hat{\mathbf{L}}^j = arg\ \min_{\mathbf{L}} \quad \left\| \mathbf{Y} - \mathbf{\Phi F}\hat{\mathbf{X}}^{\mathbf{j-1}} \right\|_F^2 + \left\| \mathbf{L} \right\|_{Optshrink} .$$

7: **X**-update

$$\hat{\mathbf{X}}^j = \hat{\mathbf{L}}^j + \hat{\mathbf{S}}^j - \mathbf{A}^T(\mathbf{A}(\hat{\mathbf{L}}^j + \hat{\mathbf{S}}^j) - \mathbf{Y}).$$

8: j=j+1
where, obj = [obj, $\left\| \mathbf{Y} - \mathbf{\Phi F}\hat{\mathbf{X}}^{\mathbf{j}} \right\|_F^2 + \lambda_1 \left\| \mathbf{\Psi}\hat{\mathbf{S}}^j \right\|_1 + \lambda_2 \left\| \hat{\mathbf{L}}^j \right\|_{Optshrink}$].
9: **end while**

Same approach of **S** recovery is being used previously in LR+S existing method [173].

L subproblem

In conventional nuclear norm minimization, low rank matrix is estimated at each iteration j via singular value thresholding (SVT) [83] as:

$$\hat{\mathbf{L}}^j = SVT(\hat{\mathbf{X}}^{j-1} - \hat{\mathbf{S}}^{j-1}; \gamma_1) = \sum_{i=1}^{q} Soft(\sigma_i, \gamma_1)\mathbf{u}_i\mathbf{v}_i^H, \tag{8.21}$$

where definition of $'Soft'$ is same as defined in (8.10) and $\sigma_i, \mathbf{u}_i$ and $\mathbf{v}_i$ are the singular values, left singular vectors, and right singular vectors, respectively. These are the values corresponding to estimated low rank matrix ($\hat{\mathbf{X}}^{j-1} - \hat{\mathbf{S}}^{j-1}$) at previous iteration. $q = min(n, T)$ denotes the rank of low rank matrix and $(.)^H$ denotes the conjugate transpose.

Recently in Optshrink method, instead of utilizing SVT low rank matrix at

iteration j is estimated as

$$\hat{\mathbf{L}}^{j} = \sum_{i=1}^{r} w_i \mathbf{u}_i \mathbf{v}_i^H, \tag{8.22}$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the left and right singular vectors as mentioned above and w_i are unknown singular values. Here, low rank matrix is assumed to have rank $r < q$ and the closed form solution of singular values in (8.22) for every $1 \leq i \leq r$ is expressed as [184]:

$$w_i = -2 \, \frac{D(\sigma_i; \mathbf{\Sigma})}{D'(\sigma_i; \mathbf{\Sigma})}, \tag{8.23}$$

where $\mathbf{\Sigma}$ is equal to $\text{diag}(\sigma_{r+1}, ...\sigma_q)$, and σ_i denotes the i^{th} singular value of previous iteration low rank matrix. $D(.)$ is the D-transform which is defined as

$$D(\sigma_i; \mathbf{\Sigma}) := \frac{1}{n}\text{Tr}(\sigma_i(\sigma_i^2\mathbf{I} - \mathbf{\Sigma}\mathbf{\Sigma}^H)^{-1}) \times \frac{1}{T}\text{Tr}(\sigma_i(\sigma_i^2\mathbf{I} - \mathbf{\Sigma}^H\mathbf{\Sigma})^{-1}), \tag{8.24}$$

and $D^{'}(.)$ is defined as

$$\begin{aligned} D^{'}(\sigma_i; \mathbf{\Sigma}) := \frac{1}{n}\text{Tr}(\sigma_i(\sigma_i^2\mathbf{I} - \mathbf{\Sigma}\mathbf{\Sigma}^H)^{-1}) \times \frac{1}{T}\text{Tr}(-2\sigma_i^2(\sigma_i^2\mathbf{I} - \mathbf{\Sigma}^H\mathbf{\Sigma})^{-2} + (\sigma_i^2\mathbf{I} - \mathbf{\Sigma}^H\mathbf{\Sigma})^{-1}) \\ + \frac{1}{T}\text{Tr}(\sigma_i(\sigma_i^2\mathbf{I} - \mathbf{\Sigma}^H\mathbf{\Sigma})^{-1} \times \frac{1}{n}\text{Tr}(-2\sigma_i^2(\sigma_i^2\mathbf{I} - \mathbf{\Sigma}\mathbf{\Sigma}^H)^{-2} + (\sigma_i^2\mathbf{I} - \mathbf{\Sigma}\mathbf{\Sigma}^H)^{-1})), \end{aligned} \tag{8.25}$$

where $\text{Tr}(.)$ is equal to the trace of a matrix, and $\mathbf{I}$ is an identity matrix.

Above formulation considers rank estimate ($=r$) on input, and provides low rank matrix at the output. In [184], it has been shown that the solution of Optshrink is quite robust to input rank specification and hence, a rough estimate of rank at the input is sufficient. Another advantage of Optshrink is that there is no need to specify shrinkage parameter as is required in SVT (refer to γ_1 in (8.21)). In SVT, we need to tune γ_1 for every dataset. It has been observed that Optshrink always outperforms SVT in estimation of low rank matrix.

Overall Solution of Optshrink LR+S

In nutshell, *Optshrink LR+S* is solved iteratively as below:

$$\begin{aligned} \hat{\mathbf{S}}^j &= Soft_{\gamma_2}\mathbf{\Psi}(\hat{\mathbf{X}}^{j-1} - \hat{\mathbf{L}}^{j-1}) \\ \hat{\mathbf{L}}^j &= Optshrink(\hat{\mathbf{X}}^{j-1} - \hat{\mathbf{S}}^{j-1}) \\ \hat{\mathbf{X}}^j &= \hat{\mathbf{L}}^j + \hat{\mathbf{S}}^j - \mathbf{A}^T(\mathbf{A}(\hat{\mathbf{L}}^j + \hat{\mathbf{S}}^j) - \mathbf{Y}), \end{aligned} \tag{8.26}$$

where $\mathbf{A} = \mathbf{\Phi F}$ in (8.26) and j denotes an iteration number.

8.3 Results

In this section, we present results of proposed methods in fMRI reconstruction and compare these results with existing state-of-the-art algorithms.

Dataset Description

To assess the performance of reconstruction methods, we have used two fMRI dataset: (*i*) Task-based fMRI dataset (OpenfMRI project publicly available dataset[1]) and (*ii*) Resting-state fMRI dataset (1000 Functional Connectomes Project publicly available dataset [2]).

Task-based Dataset

This dataset consists of acquisition of 36 axial interleaved brain slices with dimensions 72 × 72 at each time point with TE equal to 35ms and TR equal to 2s [185]. This data is collected over 179 time points, resulting in the matrix $\mathbf{X}$ of size 5184x179 for one brain slice. During the task in this dataset, the subject had to answer questions about stories that referred either to a person's false belief (mental trials) or to outdated physical representations such as an old photograph. For more details on this dataset, please refer to [185].

[1] `https://openfmri.org/dataset/`
[2] `http://www.nitrc.org/frs/?group_id=296`

Resting-state Dataset

This dataset is a freely available Beijing_Zang resting-state fMRI dataset. This dataset is a part of Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC) 1000 Functional Connectomes Project [186]. It consists of an acquisition of 33 axial interleaved ascending brain slices with a dimension of 64×64 at each time point with TR equal to 2s. The fMRI brain data is collected over 225 time points. This dataset is acquired while subjects' eyes were closed. For more details on this dataset, please refer to the website[3].

Sampling of *k*-space Data

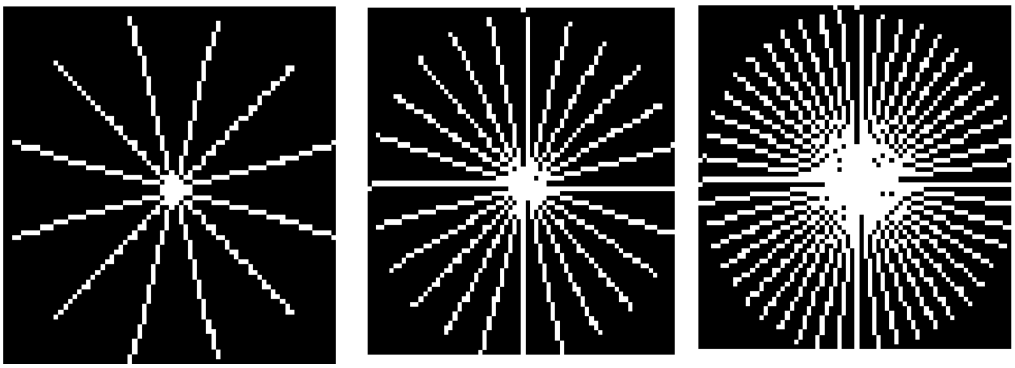
Since the raw k-t space data is not available from the dataset presented above, we simulated the undersampled measurement data $\mathbf{Y}$ by retrospective undersampling of the Fourier transform of the fully available data. Radial sampling patterns are used to undersample available k-space data as used in [187, 188, 73]. These patterns represent zeros at non-sampled locations and ones at sampled locations. Fig.8.1 shows three different radial sampling measurement patterns. As evident from Fig.8.1, these radial measurement patterns sample more data points in the low frequency region compared to the high frequency region. Sensing matrix $\mathbf{\Phi}$ in (8.1) is constructed by stacking sampling masks of all T time points. This is a standard methodology in any reconstruction related research work. We considered different sampling patterns for each time point data in order to maintain incoherency among the columns of matrix $\mathbf{X}$ [187].

Related Reconstruction Methods

We compare results of the proposed methods with other fMRI reconstruction methods including CSWD [170], k-t FASTER [172], LR+S [173], and HSPARSE [171]. Below we present brief overview of each of these existing methods.

k-t FASTER method [172]: This method reconstructs fMRI data assuming data matrix $\mathbf{X}$ to be low rank. This method is implemented by solving the below

[3] `http://fcon_1000.projects.nitrc.org/`



(a) (b) (c)

Figure 8.1: Radial sampling pattern on one slice: (a) 6 radial lines (12.856 acceleration factor); (b) 12 radial lines (6.065 acceleration factor); (c) 24 radial lines (3.495 acceleration factor)

optimization problem:

$$\hat{\mathbf{X}} = arg\ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 \quad s.t\ rank(\mathbf{X}) = r, \tag{8.27}$$

where r is pre-defined rank of $\mathbf{X}$. Here, hard thresholding is applied on the singular values of data matrix $\mathbf{X}$ as explained below. First, SVD of an initial crude estimate of matrix $\mathbf{X}$ is computed

$$\mathbf{X} = \mathbf{USV}^T. \tag{8.28}$$

Next, hard thresholding is applied on the singular values contained in $\mathbf{S}$ as

$$\hat{s_i} = \begin{cases} |s_i| - \mu & i \leq r \\ 0 & i > r \end{cases} \tag{8.29}$$

where μ is a constant, s_i is i^{th} singular value of $\mathbf{S}$, and $\hat{s_i}$ is updated singular value after hard thresholding. The value of constant μ is chosen to be 0.5 as used in [175]. Here, rank r is taken to be equal to the number of time points T.

Low rank plus sparse (LR+S) method [173]: This method reconstructs fMRI data using low rank and sparse matrix decomposition and hence, is solved using the equation as mentioned in (8.18).

CS with wavelet sparsity (CSWD) [170]: In this method, CS based reconstruction of fMRI data is carried out assuming the fMRI data to be sparse in the wavelet

domain. Hence, fMRI reconstruction is done by using the optimization framework mentioned in (8.2). We used Daubechies' orthogonal wavelet 'db4' (filter lengths 8) with 3-level decomposition as has been used in [170].

HSPARSE Method [171]: This method reconstructs fMRI data assuming data matrix $\mathbf{X}$ to be sparse in both the temporal and spatial domains. This method is implemented by solving the below optimization problem [171]:

$$\hat{\mathbf{X}} = arg\ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{\Phi F X}\|_F^2 + \lambda_3 \|\mathbf{\Psi_t X}\|_1 + \lambda_4 \|\mathbf{\Psi_s X}\|_1 , \tag{8.30}$$

where λ_3 and λ_4 are regularization parameters and, $\mathbf{\Psi_t}$ and $\mathbf{\Psi_s}$ are the temporal and spatial domain sparsifying basis, respectively. We chose discrete cosine transform (DCT) for both the temporal and the spatial sparsity as used in [171].

Reconstruction Performance measuring metric

We compare results of the proposed methods against the existing reconstruction methods using Normalized Mean Square Error (NMSE), and Peak Signal-to-Noise Ratio (PSNR). NMSE and PSNR are two well known reconstruction quality assessment metrics. Given a reference brain slice $\mathbf{x}_t$ at time point t and it's reconstructed estimate $\hat{\mathbf{x}}_t$ (t^{th} column of $\hat{\mathbf{X}}$), NMSE is calculated as

$$\text{NMSE} = \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2 \,/\, \|\mathbf{x}_t\|_2 , \tag{8.31}$$

where $\|.\|_2$ denotes l^2 norm. Similarly, PSNR is calculated as

$$\text{PSNR} = 20\log_{10} \frac{255}{\frac{1}{n_x n_y} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2}, \tag{8.32}$$

where $n_x \times n_y$ denotes the size of brain slices. In this work, one slice is being reconstructed simultaneously over all time points. Hence, NMSE and PSNR are calculated using (8.31) and (8.32) for T number of time points for a given slice and are subsequently time-averaged. In the following text, NMSE and PSNR signify average NMSE and average PSNR respectively.

Parameter values in each method

In the LR+S method (refer to (8.18)), we empirically selected $\gamma_1 = 200$ and $\gamma_2 = 2$ that provided us minimum NMSE. This is to note that we used same values of γ_2 in the proposed Optshrink LR+S method. In the CSWD method (refer to (8.2)), we used λ_1=0.1 as specified in [170] and solved this equation using non-linear conjugate gradient method [189]. In the HSPARSE method (refer to (8.30)), we empirically selected λ_3=0.1 and λ_4=0.1 and used the non-linear conjugate gradient method to solve HSPARSE [189].

The proposed DTSR (refer to (8.5)) and MCwSR (refer to (8.13)) methods require λ_1, λ_2, μ_1, μ_2 parameters to be initialized. We decided these values based on the minimization of NMSE compared to the ground truth (fully available dataset). We empirically arrived at the following values: λ_1=λ_2=0.5 and μ_1=10^2, μ_2=10^2. In addition, ADMM algorithm requires η_1, η_2, $\mathbf{B}_1^0$, $\mathbf{B}_2^0$ and $\mathbf{X}^0$ to be initialized. η_1 and η_2 are initialized as 10^{-2}. Lagrange multipliers $\mathbf{B}_1^0$ and $\mathbf{B}_2^0$ are initialized to matrices containing all one's. The fMRI data matrix $\mathbf{X}^0$ is initialized using the crude initial estimate obtained via direct IFT.

For all of the methods explained above, we stop iterations either by comparing objective function value with the predefined tolerance value or stopping with fixed number of iterations. We set the maximum number of iterations (required in optimization) to be equal to 500 and set the following convergence criteria for all methods: $optization\ function\ value(end) - optimization\ function\ value(end-1) < 10^{-5}$. All simulation was run on a computer with an Intel Core i7 CPU at 2.4 GHz, and 16 GB of RAM utilizing MATLAB (R2014a: The Mathworks, Natick, MA, USA).

Sparsifying transform $\mathbf{\Psi}$ in proposed methods

In general, sparsity is imposed in the transform domain $\mathbf{\Psi}$. We observed that fMRI data is more sparse in the temporal Fourier domain, i.e., in the Fourier domain of every voxel's time series. Hence, the matrix resulting by computing the Fourier transform of $\mathbf{X}$ along every row leads to a temporal Fourier transformed matrix that is sparse. In order to demonstrate this, we plot the sorted transformed coefficients of matrix $\mathbf{X}$ corresponding to one subject of Beijing_Zang data (Refer

to Fig.8.2).

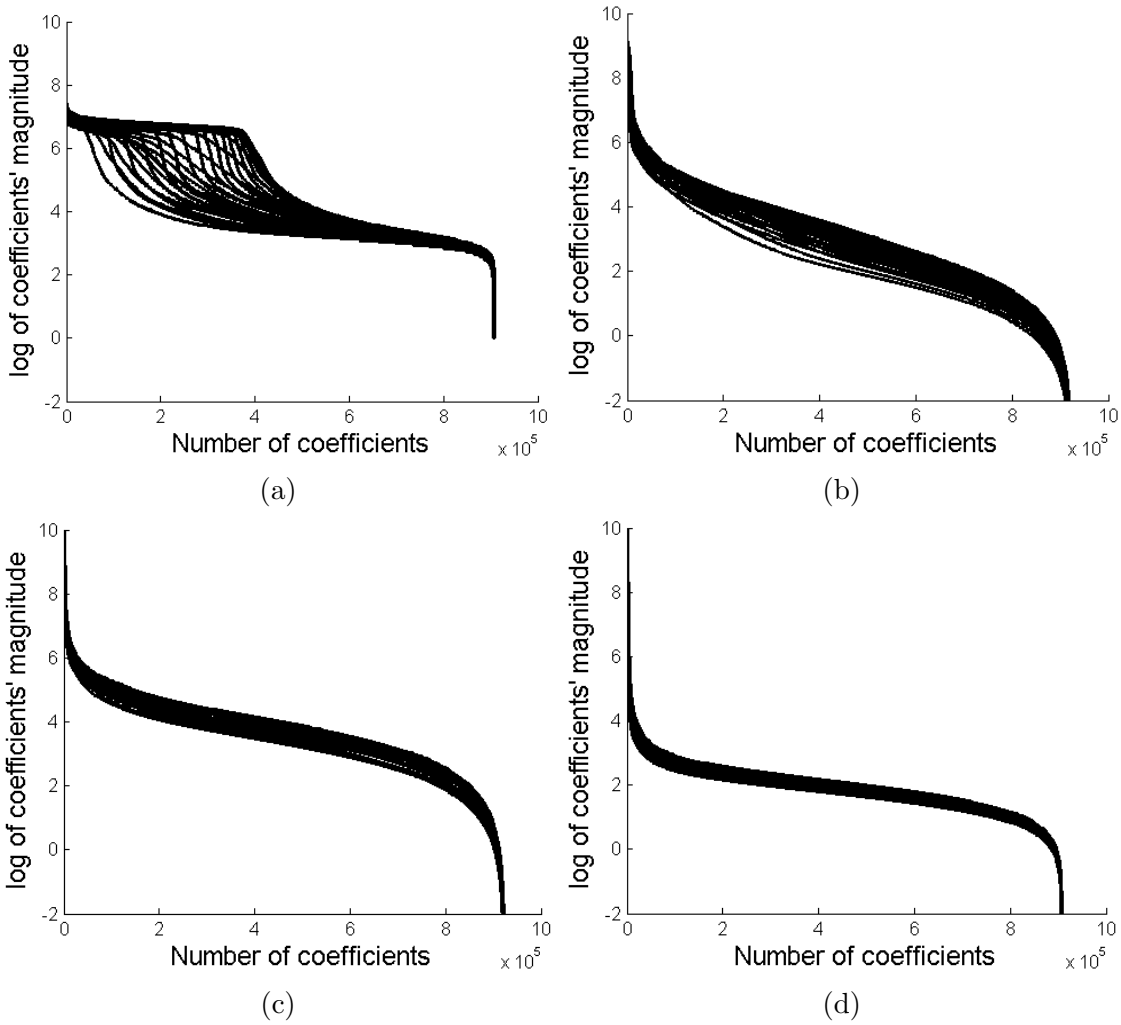


Figure 8.2: Illustration of increased sparsity of Beijing_Zang fMRI data in the temporal Fourier Transform domain. One subject's sorted log magnitude values of coefficients of matrix $\mathbf{X}$ obtained using: (a) no transform; (b) 3-level dB4 wavelet in spatial direction; (c) 2-D Discrete Cosine Transform in spatial directions; (d) 1-D Discrete Fourier Transform along the rows of $\mathbf{X}$.

From this figure, we observe that fMRI data is sparser in the temporal frequency domain. Hence, we consider $\mathbf{\Psi}$ as the corresponding temporal Fourier domain sparsifying matrix, i.e., it computes the Fourier transform along every row of $\mathbf{X}$.

Reconstruction Results

Quantitative analysis

We present reconstruction results, for varying number of radial sampling lines, in terms of average NMSE and PSNR over all slices for subject 1 of both dataset in Table 8.1 and 8.2. Results are tabulated at 6, 12, and 24 number of radial sampling lines. We present reconstruction results averaged over all volumes and over all slices.

Table 8.1: **Reconstruction results with different methods on task fMRI dataset**[a]

		NMSE			**PSNR**	
Method	**6 lines**	**12 lines**	**24 lines**	**6 lines**	**12 lines**	**24 lines**
CSWD [170]	0.2936	0.2114	0.1187	7.201	10.08	15.11
HSPARSE [171]	0.1978	0.1303	0.1062	10.968	14.289	17.745
k-t FASTER [172]	0.2776	0.2419	0.1633	6.458	8.93	12.37
LR+S [173]	0.2115	0.1175	0.0680	10.07	15.09	19.99
Proposed MCwSR	**0.0554**	**0.0519**	**0.0443**	**22.27**	**22.86**	**23.15**
Proposed OptShrink LR+S (r=1)	**0.0541**	**0.0501**	**0.0422**	**19.78**	**20.69**	**21.45**

[a] Dataset- False belief task fMRI data, results are averaged over all slices, over all time points for subject 1.

From these Tables, we observe that the proposed reconstruction methods consistently perform better than the existing reconstruction methods on both the dataset. We obtain similar consistent results across other subjects of both dataset. Though, it is ambiguous to say which proposed method is performing best among all the three proposed methods. In order to compare performance among them, we further see reproduciblity of brain networks using reconstructed data based on all the three proposed methods.

Statistical analysis for activation maps of task fMRI dataset

In this section, we explore the quality of activation maps for the task fMRI data reconstructed using the OptShrink LR+S and MCwSR methods. To this end,

Table 8.2: **Reconstruction results with different methods on rest fMRI dataset**[a]

		NMSE			**PSNR**	
Method	**6 lines**	**12 lines**	**24 lines**	**6 lines**	**12 lines**	**24 lines**
CSWD [170]	0.2478	0.1689	0.1247	5.78	8.45	11.59
HSPARSE [171]	0.1752	0.0785	0.0551	7.989	13.895	18.014
k-t FASTER [172]	0.2476	0.1697	0.1178	4.878	8.412	11.758
LR+S [173]	0.1857	0.0783	0.0604	9.012	13.846	19.781
Proposed DTSR	**0.0541**	**0.0351**	**0.031**	**19.14**	**21.76**	**22.01**
Proposed MCwSR	**0.0575**	**0.0416**	**0.0402**	**22.52**	**23.68**	**23.79**
Proposed OptShrink LR+S (r=1)	**0.0569**	**0.0443**	**0.0404**	**20.88**	**22.93**	**23.45**

[a] Dataset- Beijing_Zang resting-state fMRI data, results are averaged over all slices, over all time points for subject 1.

first pre-processing of fMRI data is done using SPM[4]. Reconstructed and original ground truth fMRI data is motion corrected followed by smoothing with 6mm FWHM as a standard pre-processing step. Next, we compute activation maps on the subject 1 of task fMRI dataset.

After preprocessing, fMRI data is fitted to four input block conditions as mentioned in [185]. These conditions are false belief story, false belief question, false belief photo story, and false belief photo question. We computed task-related activation on reconstructed fMRI data using general linear model (GLM). GLM is a univariate method used for activation detection on task fMRI data [190]. In this method, a linear model of applied stimuli is fitted to each voxel time series resulting in a set of voxel specific parameters. These parameters can be used to form contrast maps or activation maps [191]. We generated these activation maps using SPM that is a standard fMRI statistical analysis toolbox. The resulting maps are thresholded using p-value ($p < 0.05$) wherein clusters with atleast 12 voxels are considered active.

Reconstruction is performed on undersampled fMRI data on radial sampling pattern of 12 radial line. Figures 8.3 show the corresponding statistical maps obtained using (a) original fully sampled *k-t* space data (b) reconstructed data

[4]http://www.fill.ion.ucl.ac.uk/spm/

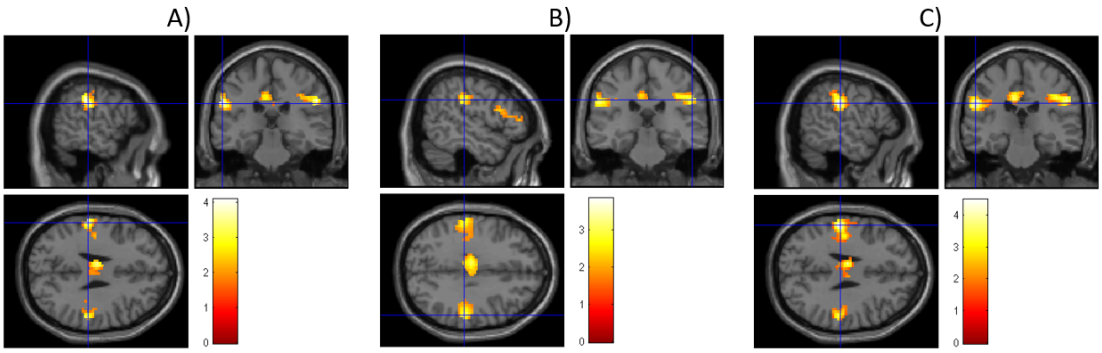


Figure 8.3: False belief fMRI data shown on saggital, coronal, and axial planes: (A) fully sampled fMRI data; (B) reconstructed fMRI data using MCwSR (12 radial lines); (C) reconstructed fMRI data using proposed *Optshrink LR+S* method (rank=1) (12 radial lines)

using MCwSR method, (c) reconstructed data using the proposed *Optshrink* LR+S method. We present results on representative slices having peak voxel of activation, whereas Montreal Neurological Institute (MNI) position of this most active voxel is being listed in Table 8.3. We also report cluster sizes and maximum z-scores values in this table. Activation maps are thresholded using t-test at the cluster level with uncorrected p-value = 0.05. Clusters with less than 12 voxels in count are rejected.

As evident from Figures 8.3 and Table 8.3, we notice that the MNI position of the most active voxel on the reconstructed data using the proposed MCwSR method is different as that obtained with the original data (with x coordinate sign reversed). In addition, on visually comparing these activation maps, we observe false activation blob with MCwSR reconstructed data activation maps as shown in 8.3.

Reproducibility of resting-state networks

In this section, we test the efficacy of all the three proposed reconstruction methods on resting-state fMRI dataset. We compare and evaluate the reproducibility of brain RSNs constructed using the proposed methods based reconstruction fMRI data and using the fully available whole brain resting fMRI dataset. Spatial Independent Component Analysis (ICA) of the reconstructed fMRI data and the

Table 8.3: **Statistical analysis results for uncorrected p=0.05**

			6 lines			**12 lines**	
	Method	**Cluster size**	***Z* score**	**MNI position**	**Cluster size**	***Z* score**	**MNI position**
1	Proposed MCwSR	96	2.83	54 -27 31	122	3.26	54 -27 31
2	Proposed Optshrink LR+S (r=1)	218	4.93	-54 -24 24	204	4.34	-54 -27 28

Task-based data - false belief task fMRI data, subject no. 1
Smoothed fully sampled fMRI data with FWHM = 6mm - **cluster size = 112, *Z* score = 3.99, MNI position (in mm) = -57 -27 28**
Please note that the coordinates of most active voxel are reported via Z score. Cluster size denotes the number of active voxels surrounding this most active voxel.

original fully available fMRI dataset is performed via GIFT toolbox [5]. ICA is a data driven method that has been widely used in resting-state fMRI to recover the set of spatially independent brain RSNs [192, 193, 194, 56]. Before applying ICA, data is preprocessed similar to as we did with task fMRI data.

For the sake of completeness, ICA model as used in fMRI is briefly discussed below. Consider matrix $\mathbf{S} \in \mathbb{R}^{T \times V}$, where T is the number of time points and V is the total number of voxels. After ICA, $\mathbf{S}$ can be expressed as:

$$\mathbf{S} = \mathbf{MN}, \tag{8.33}$$

where $\mathbf{M}$ is the $T \times C$ mixing matrix and $\mathbf{N}$ is the $C \times V$ source matrix. C is the total number of spatially independent component. Each row of source matrix $\mathbf{N}$ represents one spatially independent component and the corresponding column of the mixing matrix $\mathbf{M}$ represents time course of that independent component. The goal of spatial ICA is to model fMRI data as a mixture of maximally independent

[5] `https://www.nitrc.org/projects/gift`

spatial components.

The InfomaxICA algorithm [195] is used to obtain ICA components and the corresponding time courses. It is one of the most popular ICA algorithms that is used in fMRI data analysis [196]. The number of spatially independent components C is predefined to 100 in accordance with the previous studies [193, 194, 56, 197]. High number of components facilitate good segregation of cortical and subcortical brain functional networks [197].

To identify brain RSNs among 100 spatially independent components, spatial distribution of each component can be identified by spatial overlapping with the available template images of brain RSNs. We identified number of ICs, in all three reconstruction methods, from the mean maps of all 20 fully available fMRI subjects after removing the artifact components. These ICs can be broadly categorized into 10 RSNs: 1. Visual Network (VN), 2. SomatoMotor Network (SMN), 3. Limbic Network (LN), 4. Dorsal Attention Network (DAN), 5. Ventral Attention Network (VAN), 6. Default Mode Network (DMN), 7. Frontoparietal Network (FPN), 8. Temporal + Frontal Network (TFN), 9. Subcortical Network (SCN), and 10. Cerebellar Network (CN).

We manually arranged various ICs into various RSNs stated above. We observed RSNs to be similar across all three proposed methods. The spatial maps of some RSNs discovered by the fully available Beijing_Zang data and the DTSR reconstructed data are shown in Fig.8.4 to Fig.8.5. Left part of each figure represents networks identified using the fully available data and the right part represents networks identified using the DTSR reconstructed data. It is clear that spatial activation maps of RSNs obtained from the reconstructed data overlap significantly with the RSNs of the fully available fMRI data. Similar results were obtained with the other two proposed methods.

Further, we evaluate the performance of proposed methods in terms of identifying functional brain networks using our proposed approach mentioned in chapter 3. To this end, we first computed connectivity matrices across all subjects using MVRC and later, utilized modularity to discover communities from the averaged connectivity matrices. We mapped the identified communities using proposed DTSR method onto the human brain, as shown in Fig 8.6. Fig 8.6 displays the communities derived the proposed method refer to default mode, subcortical,

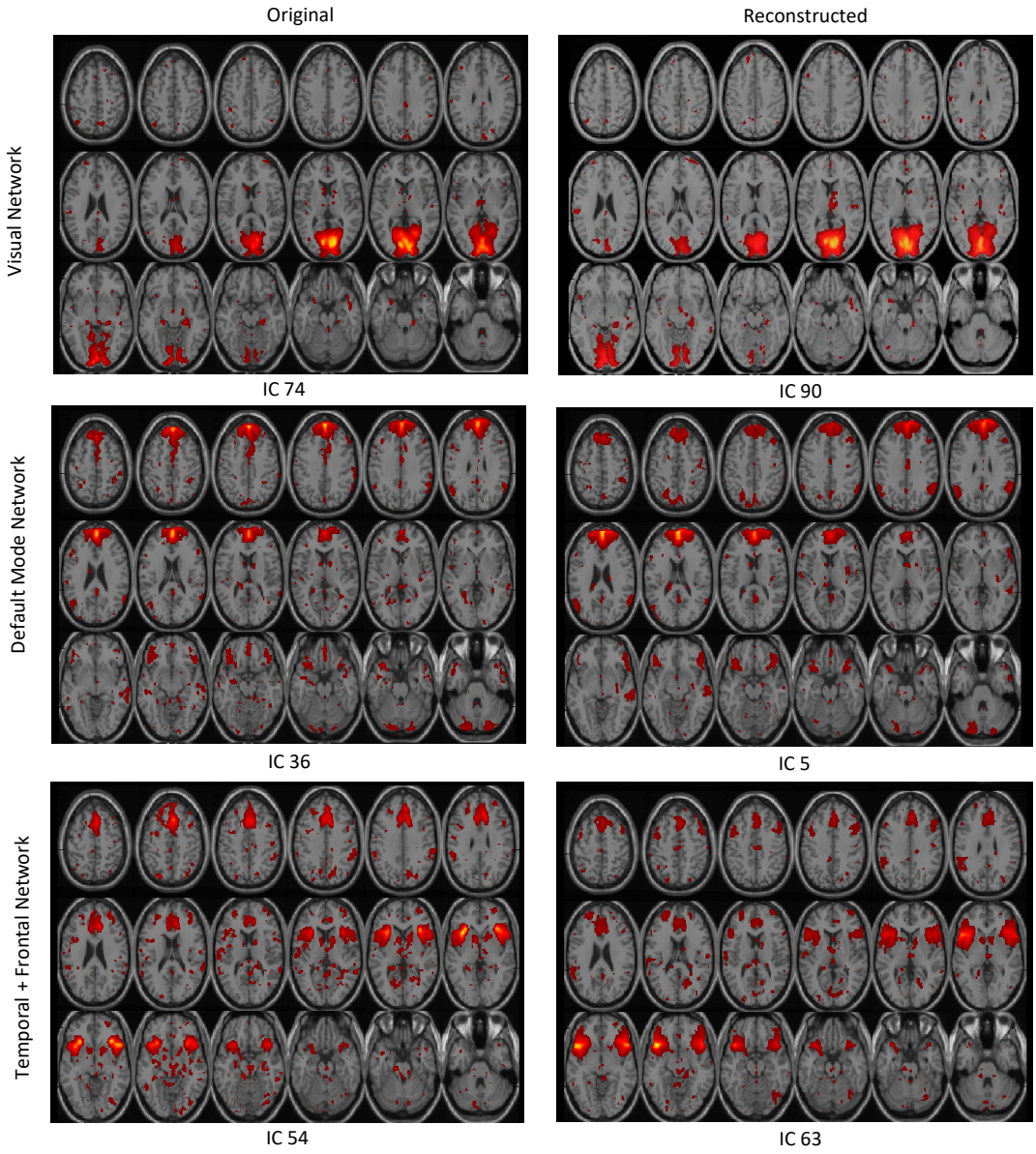


Figure 8.4: Axial view of spatial maps of various RSNs where the left part of each figure represents networks identified using the fully available (original) Beijing_Zang data and the right part represents networks identified using the DTSR reconstructed data. Each row corresponds to results on one RSN. Number in brackets below each image represents independent component (IC) number obtained after group ICA.

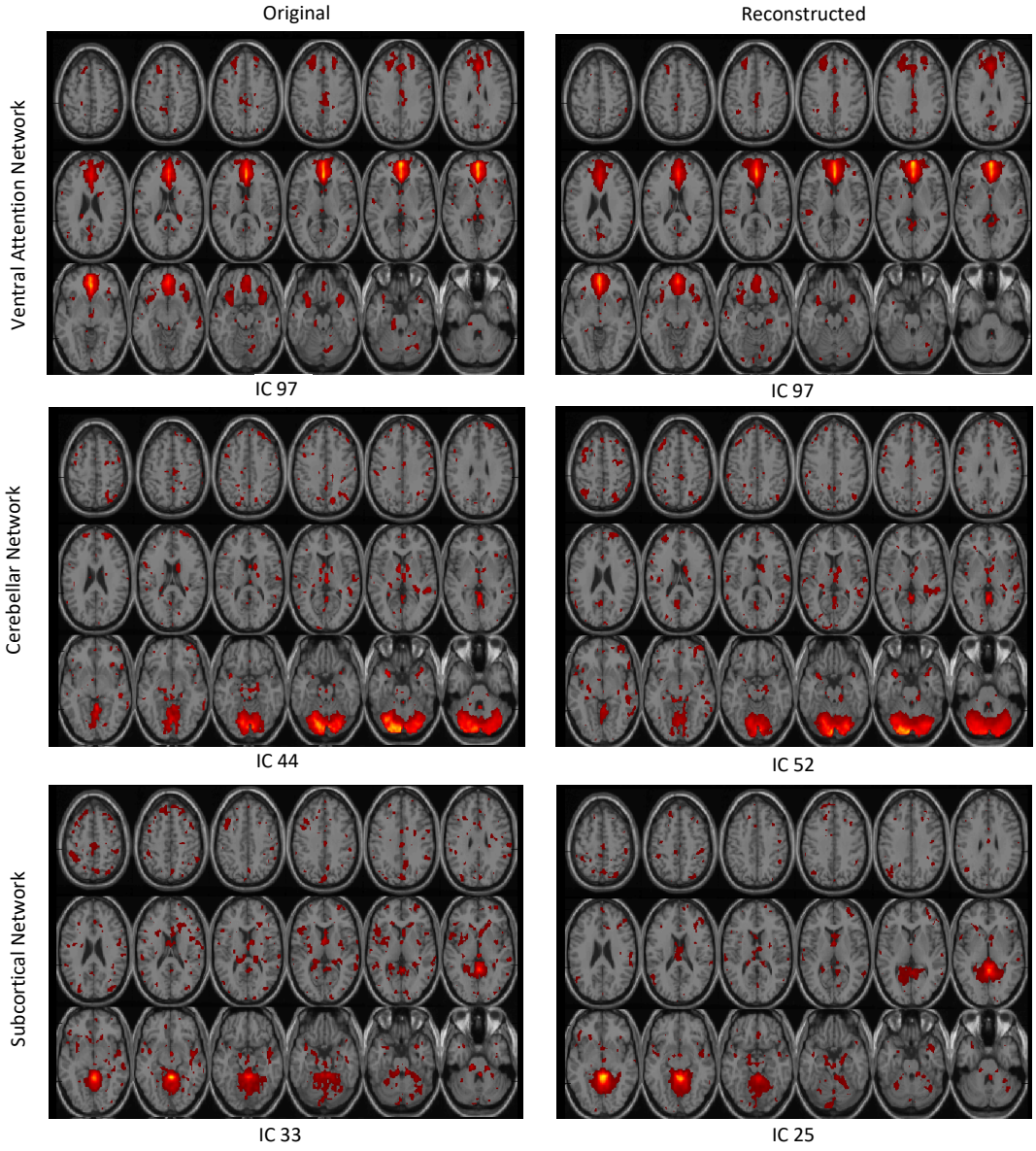


Figure 8.5: Axial view of spatial maps of various RSNs where the left part of each figure represents networks identified using the fully available (original) Beijing_Zang data and the right part represents networks identified using the DTSR reconstructed data. Number in brackets below each image represents independent component (IC) number obtained after group ICA.

visual, bilateral limbic, motor and auditory networks.These networks are highly consistent with resting-state fBNs discovered by the fully available ground truth data. In essence, we observed similar results with other proposed reconstruction methods also.

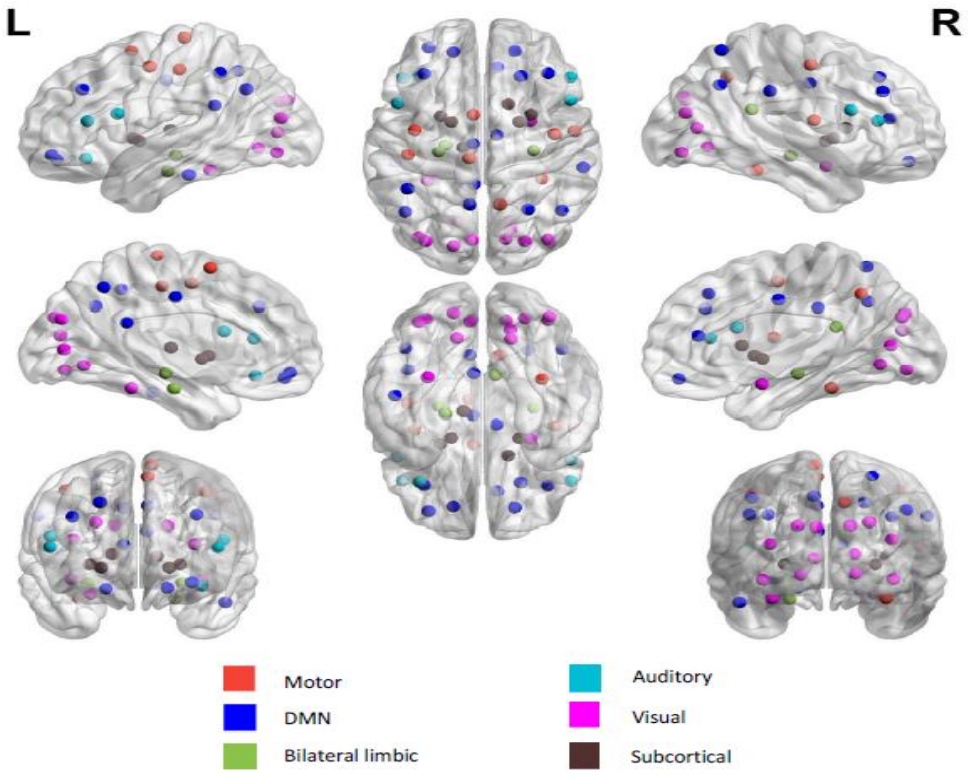


Figure 8.6: Communities derived from DTSR reconstructed data.

Regarding communities detection evaluation of reconstruction methods, we compare the community partitions with the ground-truth communities obtained with fully-sampled data. In order to assess community partition robustness of reconstructed data, NMI can be used as a powerful tool for assessing these algorithms in terms of the agreement of detected communities with the "fully sampled" data. For the comparison of the communities, it is examined whether the reconstructed data identified communities show higher NMI value with the fully sampled data community structures. We observed the NMI value to be 0.76. It seems that the agreement with fully sampled community structures is high, suggesting the robust

performance of the proposed method.

8.4 Conclusions

In this chapter, we have proposed novel fMRI reconstruction methods that exploits combination of both sparsity and low-rank to improve fMRI reconstruction quality. We compared the performance of proposed methods with the existing methods including k-t FASTER, LR+S, CSWD, and HSPARSE on two real fMRI dataset. Our results show that the proposed methods yield the most faithful reconstruction. Further, their performance is consistently good at very low sampling ratios.

In addition, proposed methods is able to preserve the voxel activation maps of fMRI data that is not observed with any other existing method. Rather, most of the existing methods provide false activation that can yield misleading findings on fMRI data. This establishes the significance of the proposed method. In addition, we provide a mechanism to validate the reconstruction quality of an fMRI reconstruction method via building brain activation maps.

Chapter 9

Conclusions and Future Research Directions

In this chapter, we provide the conclusions by summarizing our main contributions and also indicate future research directions.

9.1 Conclusions

Human brain is a complex network system that comprises of multiple functionally coupled brain regions. Understanding of brain is not limited to identification of active brain regions. It further requires an identification of functionally coupled brain regions. This functional coupling of brain regions in the literature is known as functional brain networks (fBNs). In this dissertation, we attempted to advance the theory, in particular, of identifying fBNs by proposing methods for non-invasive neuro-imaging fMRI modality. We have developed analytical techniques to extract fBNs. The primary goal of this thesis is to develop new analytic techniques and apply them to model whole-brain functional connectivity (FC) based on resting-state fMRI data. In addition, we have also presented methods for accelerated reconstruction of brain fMRI data using undersampled measurement data for faster data acquisition.

We began with the discussion of nodes and edges for the case of fBNs. Within the context of FC framework, human brain is characterized by the set of nodes (brain regions) and pairwise relationships between them (edges). This is repre-

sented by a matrix with all possible pairwise connections between brain regions. Pairwise connections between nodes is assessed by calculating the statistical dependency between the activity of two brain regions, such as Pearson correlation, and then fBNs can be identified by considering all pairwise functional connections.

This discussion is followed by presenting proposed multivariate regression method to build FC matrix. Proposed multivariate regression method considers weighted combination of different brain regions while computing functional connections between them and captures sparse and dense fBNs simultaneously. We name the proposed method as Multivariate Vector Regression-based Connectivity (MVRC). In our next work, we extended proposed MVRC method to identify low-rank and sparse fBNs by imposing low-rank and sparsity constraints simultaneously. This work also proposed method to identify overlapping fBNs that are often ignored in the literature. By overlapping, we mean that one brain region may be a part of multiple fBNs. This sounds plausible because one triggering stimulus, say auditory, may stimulate memory and other fBNs apart from the auditory network. This indicates a need of identifying overlapping fBNs compared to the commonly identified disjoint fBNs.

In our next work, we proposed a method that automates the process of identifying fBNs at the group-level comprising of multiple subjects using group-fused optimization constraint. Here, we considered a group of few normal subjects and presented a framework to identify group-fused FC matrices across all subjects. This work also proposed method to identify overlapping fBNs at the group-level.

In our later work, we generalized proposed MVRC method to identify dynamic fBNs. Recent methods of computing dynamic fBNs end up computing edges in some fixed time window duration and repeat it until entire scan duration is covered by moving time windows. Thus, in order to identify dynamic fBNs by considering state-of-the-art time window approach, proposed method utilized multivariate regression method to compute FC in each time window and further, utilized tensor factorization approach to identify dynamic overlapping fBNs. The proposed approach is applied to real fMRI data derived from subjects with autism disorder and healthy controls, and we find robust group evidence of disease-related changes in brain networks. We use the statics to compare/investigate statistical difference between networks across healthy and diseased (autism) subjects, showing marked

difference between the two groups.

Our next work involved compressed sensing based accelerated reconstruction of fMRI data. Lower temporal resolution in neuroimaging fMRI data results in artifacts due to long scanning time. This also causes annoyance in patients inside scanner, leading to images blurred with motion and other physiological artifacts. Therefore, it is necessary to capture images in shortest possible time. One possible solution is to use either parallel imaging or use compressed sensing (CS) approaches. In the end of this dissertation, we presented our works carried out on developing new methods for CS fMRI for faster data acquisition, based on sparsity and matrix recovery techniques. We aim to build methods that are able to better preserve fBNs compared to the existing methods. This is essential to show that the proposed method is accurate as well as is able to preserve fBNs, which is one of the crucial motive for studying fMRI data.

In the nutshell, we have addressed the problem of fBNs identification by providing novel multivariate regression technique, followed by methods to identify overlapping fBNs at the individual and group-level. In the subsequent chapter, we presented proposed method to identify dynamic fBNs and to extract intrinsic perceived stimuli that is responsible for the functional coupling of brain regions. At last, we presented multiple compressed sensing techniques to recover undersampled fMRI data for faster acquisition.

9.2 Future Research Directions

As a future work, it is worthwhile to investigate following directions:

- The methods developed in this dissertation allow us to extract undirected brain networks. However, directional brain networks should also be extracted and studied for complete understanding of the diseased brain's networks. Directed adjacency matrix weights can be estimated using the time-shifted version of region time-series, as opposed to static functional connectivity. Functional connectivity do not provide information about the directed casual interactions among brain regions.

- A sliding-window approach is commonly used on fMRI time-series data to elucidate dynamic aspects of functional brain networks. Besides its simplicity, this technique carries some obvious limitations. Foremost, the choice of the window length has long been a matter of debate. Too short window lengths increase the risks of introducing spurious fluctuations in the observed dFC. On the other hand, too long windows would impede the detection of the temporal variations of FC. Therefore, a trade-off must be considered to keep satisfactory ranges of both specificity (window length long enough to detect reliable dFC fluctuations) and sensitivity (window length short enough not to miss genuine dFC variations). The commonly used rule of thumb (1/fmin, fmin being the cut-off frequency of the highpass filter applied to the fMRI time-courses), eventually leads to the choice of a fixed window length. A different family of approaches detecting connectivity at each time point might help to effectively escape from this constraint.

- Most existing approaches examine the temporal changes of functional connectivity in the resting-state, although connectivity may also be altered in the task performance. Therefore, it would be interesting to investigate the brain connectivity patterns under different task categories and also during the period of transition from the task-state to the resting-state. It might act as a potential way to investigate a variety of neurological disorders.

- Building and comparing brain networks with estimated perceived stimulus is worth exploring in the future. In this direction, we can also study the multivariate approach for the extraction of FC.

- Finally, the functional brain imaging data alone may not be sufficient to gain a comprehensive understanding of the brain's functional organization. A multimodal data combination and analysis of human brain may provide better knowledge about the underlying networks' organization and guide us to a deeper understanding of the human brain.

Bibliography

[1] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich, "Network modelling methods for FMRI," *NeuroImage*, vol. 54, no. 2, pp. 875 – 891, 2011.

[2] S. B. Laughlin and T. J. Sejnowski, "Communication in neuronal networks," *Science*, vol. 301, no. 5641, pp. 1870–1874, 2003.

[3] O. Sporns, *Networks of the Brain*. MIT press, 2010.

[4] M. Newman, *Networks: an introduction*. Oxford university press, 2010.

[5] Q. K. Telesford, S. L. Simpson, J. H. Burdette, S. Hayasaka, and P. J. Laurienti, "The brain as a complex system: Using network science as a tool for understanding the brain," *Brain Connectivity*, vol. 1, no. 4, pp. 295–308, Oct 2011.

[6] A. Fornito, A. Zalesky, and E. Bullmore, *Fundamentals of brain network analysis*. Academic Press, 2016.

[7] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, no. 4, pp. 175 – 308, 2006.

[8] H.-J. Park and K. Friston, "Structural and functional brain networks: From connections to cognition," *Science*, vol. 342, no. 6158, 2013.

[9] M. P. Van Den Heuvel and O. Sporns, "Network hubs in the human brain," *Trends in Cognitive Sciences*, vol. 17, no. 12, pp. 683–696, 2013.

[10] L. Pessoa, "Understanding brain networks and brain organization," *Physics of Life Reviews*, vol. 11, no. 3, pp. 400 – 435, 2014.

[11] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *J. Neurosci.*, vol. 26, no. 1, pp. 63–72, 2006.

[12] D. S. Bassett and E. Bullmore, "Small-world brain networks," *The Neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.

[13] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nat. Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, 2009.

[14] O. Sporns, *Discovering the human connectome*. MIT press, 2012.

[15] C. T. Butts, "Revisiting the foundations of network analysis," *Science*, vol. 325, no. 5939, pp. 414–416, 2009.

[16] J. W. Lichtman and W. Denk, "The big and the small: Challenges of imaging the brain's circuits," *Science*, vol. 334, no. 6056, pp. 618–623, 2011.

[17] O. Sporns, G. Tononi, and R. Ktter, "The human connectome: A structural description of the human brain," *PLOS Computational Biology*, vol. 1, no. 4, 2005.

[18] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," *Magn. Reson. Med.*, vol. 34, no. 4, pp. 537–541, 1995.

[19] D. Cordes, V. M. Haughton, K. Arfanakis, G. J. Wendt, P. A. Turski, C. H. Moritz, M. A. Quigley, and M. E. Meyerand, "Mapping functionally related regions of brain with functional connectivity MR imaging," *American Journal of Neuroradiology*, vol. 21, no. 9, pp. 1636–1644, 2000.

[20] M. De La Iglesia-vaya, A. S. Kanaan, J. Molina-Mateo, L. Martí-Bonmati, and M. J. Escarti-Fabra, *Brain Connections-Resting State fMRI Functional Connectivity*. Intech Open Access Publisher, 2013.

[21] K. J. Friston, "Functional and effective connectivity: A review," *Brain Connectivity*, vol. 1, no. 1, pp. 13–36, 2011.

[22] N. K. Logothetis, "What we can do and what we cannot do with fMRI," *Nature*, vol. 453, no. 7197, pp. 869–878, 2008.

[23] M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore, "Detection of functional connectivity using temporal correlations in MR images," *Human Brain Mapping*, vol. 15, no. 4, pp. 247–262, 2002.

[24] T. Chen, S. Ryali, S. Qin, and V. Menon, "Estimation of resting-state functional connectivity using random subspace based partial correlation: A novel method for reducing global artifacts," *NeuroImage*, vol. 82, pp. 87–100, 2013.

[25] G. Marrelec, A. Krainik, H. Duffau, M. Plgrini-Issac, S. Lehricy, J. Doyon, and H. Benali, "Partial correlation for functional brain interactivity investigation in functional MRI," *NeuroImage*, vol. 32, no. 1, pp. 228 – 237, 2006.

[26] S. Ryali, T. Chen, K. Supekar, and V. Menon, "Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty," *Neuroimage*, vol. 59, no. 4, pp. 3852–3861, 2012.

[27] G. Varoquaux and R. C. Craddock, "Learning and comparing functional connectomes across subjects," *NeuroImage*, vol. 80, pp. 405–415, 2013.

[28] B. Cassidy, C. Rae, and V. Solo, "Brain activity: Connectivity, sparsity, and mutual information," *IEEE Transactions on Medical Imaging*, vol. 34, no. 4, pp. 846–860, 2015.

[29] K. Li, L. Guo, J. Nie, G. Li, and T. Liu, "Review of methods for functional brain connectivity detection using fMRI," *Computerized Medical Imaging and Graphics*, vol. 33, no. 2, pp. 131–139, 2009.

[30] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. D. Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, "Dynamic functional connectivity: Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360 – 378, 2013.

[31] A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen, "Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses," *Science*, vol. 273, no. 5283, pp. 1868–1871, 1996.

[32] S. Makeig, A. Delorme, M. Westerfield, T.-P. Jung, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Electroencephalographic brain dynamics following manually responded visual targets," *PLOS Biology*, vol. 2, no. 6, 2004.

[33] C. Chang and G. H. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI," *NeuroImage*, vol. 50, no. 1, pp. 81–98, 2010.

[34] A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, and M. Breakspear, "Time-resolved resting-state brain networks," *Proc. Natl. Acad. Sci. USA*, vol. 111, no. 28, pp. 10 341–10 346, 2014.

[35] R. Hindriks, M. Adhikari, Y. Murayama, M. Ganzetti, D. Mantini, N. Logothetis, and G. Deco, "Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI?" *NeuroImage*, vol. 127, pp. 242–256, 2016.

[36] M. Najafi, B. W. McMenamin, J. Z. Simon, and L. Pessoa, "Overlapping communities reveal rich structure in large-scale brain networks during rest and task conditions," *NeuroImage*, vol. 135, pp. 92–106, 2016.

[37] M. D. Luca, C. Beckmann, N. D. Stefano, P. Matthews, and S. Smith, "fMRI resting state networks define distinct modes of long-distance interactions in the human brain," *NeuroImage*, vol. 29, no. 4, pp. 1359 – 1367, 2006.

[38] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional Magnetic Resonance Imaging," *Nat. Rev. Neurosci.*, vol. 8, no. 9, pp. 700–711, Sep 2007.

[39] M. P. van den Heuvel and H. E. H. Pol, "Exploring the brain network: A review on resting-state fMRI functional connectivity," *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519 – 534, 2010.

[40] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation." *Proc Natl Acad Sci USA*, vol. 87, no. 24, pp. 9868–9872, 1990.

[41] R. M. Birn, J. B. Diamond, M. A. Smith, and P. A. Bandettini, "Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI," *NeuroImage*, vol. 31, no. 4, pp. 1536 – 1548, 2006.

[42] Y. Behzadi, K. Restom, J. Liau, and T. T. Liu, "A component based noise correction method (compcor) for bold and perfusion based fMRI," *NeuroImage*, vol. 37, no. 1, pp. 90 – 101, 2007.

[43] M. D. Fox, D. Zhang, A. Z. Snyder, and M. E. Raichle, "The global signal and observed anticorrelated resting state brain networks," *Journal of neurophysiology*, vol. 101, no. 6, pp. 3270–3283, 2009.

[44] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion," *NeuroImage*, vol. 59, no. 3, pp. 2142 – 2154, 2012.

[45] P. Kundu, N. D. Brenowitz, V. Voon, Y. Worbe, P. E. Vrtes, S. J. Inati, Z. S. Saad, P. A. Bandettini, and E. T. Bullmore, "Integrated strategy for improving functional connectivity mapping using multiecho fMRI," *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 40, pp. 16 187–16 192, 2013.

[46] T. E. Lund, K. H. Madsen, K. Sidaros, W.-L. Luo, and T. E. Nichols, "Non-white noise in fMRI: Does modelling have an impact?" *NeuroImage*, vol. 29, no. 1, pp. 54 – 66, 2006.

[47] K. R. A. Van Dijk, T. Hedden, A. Venkataraman, K. C. Evans, S. W. Lazar, and R. L. Buckner, "Intrinsic functional connectivity as a tool for human connectomics: Theory, properties, and optimization," *Journal of Neurophysiology*, vol. 103, no. 1, pp. 297–321, Nov 2009.

[48] V. M. Eguíluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian, "Scale-free brain functional networks," *Phys. Rev. Lett.*, vol. 94, p. 018102, 2005.

[49] M. van den Heuvel, C. Stam, M. Boersma, and H. H. Pol, "Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain," *NeuroImage*, vol. 43, no. 3, pp. 528 – 539, 2008.

[50] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, 2002.

[51] J. L. Lancaster, M. G. Woldorff, L. M. Parsons, M. Liotti, C. S. Freitas, L. Rainey, P. V. Kochunov, D. Nickerson, S. A. Mikiten, and P. T. Fox, "Automated talairach atlas labels for functional brain mapping," *Human brain mapping*, vol. 10, no. 3, pp. 120–131, 2000.

[52] R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, and E. Bullmore, "Neurophysiological architecture of functional magnetic resonance images of human brain," *Cerebral Cortex*, vol. 15, no. 9, pp. 1332–1342, 2005.

[53] J. Richiardi, S. Achard, H. Bunke, and D. V. D. Ville, "Machine learning with brain graphs: Predictive modeling approaches for functional imaging in systems neuroscience," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 58–70, 2013.

[54] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9673–9678, 2005.

[55] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith, "Investigations into resting-state connectivity using independent component analysis,"

Philosophical Transactions of the Royal Society of London B: Biological Sciences, vol. 360, no. 1457, pp. 1001–1013, 2005.

[56] S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird *et al.*, "Correspondence of the brain's functional architecture during activation and rest," *Proceedings of the National Academy of Sciences*, vol. 106, no. 31, pp. 13 040–13 045, 2009.

[57] J. Power, A. Cohen, S. Nelson, G. Wig, K. Barnes, J. Church, A. Vogel, T. Laumann, F. Miezin, B. Schlaggar, and S. Petersen, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.

[58] B. T. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *J. Neurophysiol.*, vol. 106, no. 3, pp. 1125–1165, 2011.

[59] S. Ostojic, N. Brunel, and V. Hakim, "How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains," *Journal of Neuroscience*, vol. 29, no. 33, pp. 10 234–10 253, 2009.

[60] M. Schlvinck, K. Friston, and G. Rees, "The influence of spontaneous activity on stimulus processing in primary visual cortex," *NeuroImage*, vol. 59, no. 3, pp. 2700 – 2708, 2012.

[61] M. W. Cole, G. J. Yang, J. D. Murray, G. Repovš, and A. Anticevic, "Functional connectivity change as shared signal dynamics," *Journal of neuroscience methods*, vol. 259, pp. 22–39, 2016.

[62] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[63] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Natl. Acad. Sci. USA*, vol. 103, no. 23, pp. 8577–8582, 2006.

[64] M. Newman, "Equivalence between modularity optimization and maximum likelihood methods for community detection," *Phys. Rev. E*, vol. 94, p. 052315, 2016.

[65] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, no. 2, p. 026113, 2004.

[66] G. Marrelec, B. Horwitz, J. Kim, M. Plgrini-Issac, H. Benali, and J. Doyon, "Using partial correlation to enhance structural equation modeling of functional mri data," *Magnetic Resonance Imaging*, vol. 25, no. 8, pp. 1181–1189, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0730725X07001415

[67] R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, and E. Bullmore, "Neurophysiological architecture of functional magnetic resonance images of human brain," *Cerebral Cortex*, vol. 15, no. 9, pp. 1332–1342, 2005. [Online]. Available: http://dx.doi.org/10.1093/cercor/bhi016

[68] P. Aggarwal, A. Gupta, and A. Garg, "Multivariate brain network graph identification in functional MRI," *Medical Image Analysis*, vol. 42, pp. 228–240, 2017.

[69] X. Li and H. Wang, "Identification of functional networks in resting state fmri data using adaptive sparse representation and affinity propagation clustering," *Frontiers in Neuroscience*, vol. 9, p. 383, 2015.

[70] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.

[71] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[72] C. Hu, S. Reeves, D. C. Peters, and D. Twieg, "An efficient reconstruction algorithm based on the alternating direction method of multipliers for joint estimation of R_2^* and off-resonance in fMRI," *IEEE Transactions on Medical Imaging*, vol. 36, no. 6, pp. 1326–1336, 2017.

[73] P. Aggarwal and A. Gupta, "Accelerated fMRI reconstruction using matrix completion with sparse recovery via split bregman," *Neurocomputing*, vol. 216, pp. 319 – 330, 2016.

[74] P. Aggarwal and A.Gupta, "Double temporal sparsity based accelerated reconstruction of compressively sensed resting-state fMRI," *Computers in Biology and Medicine*, vol. 91, pp. 255 – 266, 2017.

[75] I. W. Selesnick, "Sparse signal restoration," 2010.

[76] O. Sporns and R. F. Betzel, "Modular brain networks," *Annual review of psychology*, vol. 67, pp. 613–640, 2016.

[77] K. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273 – 1302, 2003.

[78] C. Lu, J. Feng, Z. Lin, and S. Yan, "Correlation adaptive subspace segmentation by trace lasso," in *IEEE International Conference on Computer Vision*, 2013, pp. 1345–1352.

[79] S. Simpson, R. Lyday, S. Hayasaka, A. Marsh, and P. Laurienti, "A permutation testing framework to compare groups of brain networks," *Frontiers in Computational Neuroscience*, vol. 7, p. 171, 2013.

[80] J.-i. Hirayama, A. Hyvärinen, and S. Ishii, "Sparse and low-rank matrix regularization for learning time-varying markov networks," *Machine Learning*, vol. 105, no. 3, pp. 335–366, 2016.

[81] D. Precup and P. Bachman, "Improved estimation in time varying models," in *Proceedings of the 29th International Conference on Machine Learning*. ACM, 2012, pp. 1459–1466.

[82] P. Aggarwal and A. Gupta, "Low rank and sparsity constrained method for identifying overlapping functional brain networks," *PLOS ONE*, vol. 13, no. 11, pp. 1–19, 2018.

[83] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[84] I. Psorakis, S. Roberts, M. Ebden, and B. Sheldon, "Overlapping community detection using bayesian non-negative matrix factorization," *Physical Review E*, vol. 83, no. 6, p. 066114, 2011.

[85] J. Yang and J. Leskovec, "Overlapping community detection at scale: A nonnegative matrix factorization approach," in *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*, 2013, pp. 587–596.

[86] Z.-Y. Zhang, Y. Wang, and Y.-Y. Ahn, "Overlapping community detection in complex networks using symmetric binary matrix factorization," *Phys. Rev. E*, vol. 87, p. 062803, Jun 2013.

[87] S. Sra and I. S. Dhillon, *Nonnegative matrix approximation: Algorithms and applications*. Computer Science Department, University of Texas at Austin, 2006.

[88] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. John Wiley & Sons, Ltd, 2009.

[89] Z. Zheng, J. Yang, and Y. Zhu, "Initialization enhancer for non-negative matrix factorization," *Engineering Applications of Artificial Intelligence*, vol. 20, no. 1, pp. 101–110, 2007.

[90] M. N. Schmidt, O. Winther, and L. K. Hansen, "Bayesian non-negative matrix factorization," in *International Conference on Independent Component Analysis and Signal Separation*. Springer, 2009, pp. 540–547.

[91] M. Rubinov and O. Sporns, "Weight-conserving characterization of complex functional brain networks," *NeuroImage*, vol. 56, no. 4, pp. 2068–2079, 2011.

[92] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis," *Human brain mapping*, vol. 14, no. 3, pp. 140–151, 2001.

[93] S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson, M. F. Glasser, D. C. Van Essen, D. A. Feinberg, E. S. Yacoub, and K. Ugurbil, "Temporally independent functional modes of spontaneous brain activity," *Proceedings of the National Academy of Sciences*, vol. 109, no. 8, pp. 3131–3136, 2012.

[94] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[95] F. Barkhof, S. Haller, and S. A. R. B. Rombouts, "Resting-state functional mr imaging: A new window to the brain," *Radiology*, vol. 272, no. 1, pp. 29–49, 2014.

[96] X.-N. Zuo, C. Kelly, J. S. Adelstein, D. F. Klein, F. X. Castellanos, and M. P. Milham, "Reliable intrinsic connectivity networks: Test retest evaluation using ica and dual regression approach," *NeuroImage*, vol. 49, no. 3, pp. 2163–2177, 2010.

[97] M. Mennes, B. B. Biswal, F. X. Castellanos, and M. P. Milham, "Making data sharing work: The FCP/INDI experience," *NeuroImage*, vol. 82, no. Supplement C, pp. 683 – 691, 2013.

[98] D. C. V. Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, and K. Ugurbil, "The WU-minn human connectome project: An overview," *NeuroImage*, vol. 80, no. Supplement C, pp. 62 – 79, 2013.

[99] R. A. Poldrack, T. O. Laumann, O. Koyejo, B. Gregory, A. Hover, M.-Y. Chen, K. J. Gorgolewski, J. Luci, S. J. Joo, R. L. Boyd *et al.*, "Long-term neural and physiological phenotyping of a single human," *Nature communications*, vol. 6, p. 8885, 2015.

[100] T. D. Satterthwaite and C. Davatzikos, "Towards an individualized delineation of functional neuroanatomy," *Neuron*, vol. 87, no. 3, pp. 471 – 473, 2015.

[101] H. Li, T. D. Satterthwaite, and Y. Fan, "Large-scale sparse functional networks from resting state fMRI," *NeuroImage*, vol. 156, no. Supplement C, pp. 1 – 13, 2017.

[102] D. Meunier, S. Achard, A. Morcom, and E. Bullmore, "Age-related changes in modular organization of human brain functional networks," *NeuroImage*, vol. 44, no. 3, pp. 715 – 723, 2009.

[103] R. C. Craddock, G. James, P. E. Holtzheimer, X. P. Hu, and H. S. Mayberg, "A whole brain fMRI atlas generated via spatially constrained spectral clustering," *Human Brain Mapping*, vol. 33, no. 8, pp. 1914–1928, 2012.

[104] S. Mueller, D. Wang, M. D. Fox, B. T. Yeo, J. Sepulcre, M. R. Sabuncu, R. Shafee, J. Lu, and H. Liu, "Individual variability in functional connectivity architecture of the human brain," *Neuron*, vol. 77, no. 3, pp. 586 – 595, 2013.

[105] N. D. Cahill, H. Singh, C. Zhang, D. A. Corcoran, A. M. Prengaman, P. S. Wenger, J. F. Hamilton, P. Bajorski, and A. M. Michael, "Multiple-view spectral clustering for group-wise functional community detection," *arXiv preprint arXiv:1611.06981*, 2016.

[106] L. Dodero, A. Gozzi, A. Liska, V. Murino, and D. Sona, *Group-wise functional community detection through joint laplacian diagonalization.* Cham: Springer International Publishing, 2014, pp. 708–715.

[107] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[108] M. L. Overton and R. S. Womersley, "Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices," *Mathematical Programming*, vol. 62, no. 1, pp. 321–357, Feb 1993.

[109] X. Liu, S. Ji, W. Glnzel, and B. D. Moor, "Multiview partitioning via tensor methods," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 5, pp. 1056–1069, 2013.

[110] Y.-Y. Chuang, "Affinity aggregation for spectral clustering," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 773–780.

[111] P. K. Gopalan and D. M. Blei, "Efficient discovery of overlapping communities in massive networks," *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 36, pp. 14 534–14 539, 2013.

[112] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3, pp. 75 – 174, 2010.

[113] A. Alexander-Bloch, R. Lambiotte, B. Roberts, J. Giedd, N. Gogtay, and E. Bullmore, "The discovery of population differences in network community structure: New methods and applications to brain functional networks in schizophrenia," *NeuroImage*, vol. 59, no. 4, pp. 3889 – 3900, 2012.

[114] A. Strehl and J. Ghosh, "Cluster ensembles — a knowledge reuse framework for combining multiple partitions," *J. Mach. Learn. Res.*, vol. 3, pp. 583–617, Mar. 2003.

[115] Z. Yao, B. Hu, Y. Xie, F. Zheng, G. Liu, X. Chen, and W. Zheng, "Resting-state time-varying analysis reveals aberrant variations of functional connectivity in autism," *Front. Hum. Neurosci.*, vol. 10, p. 463, 2016.

[116] N. Leonardi and D. V. D. Ville, "On spurious and real fluctuations of dynamic functional connectivity during rest," *NeuroImage*, vol. 104, pp. 430 – 436, 2015.

[117] W. H. Thompson and P. Fransson, "The frequency dimension of fMRI dynamic connectivity: Network connectivity, functional hubs and integration in the resting brain," *NeuroImage*, vol. 121, pp. 227–242, 2015.

[118] L. F. Robinson, L. Y. Atlas, and T. D. Wager, "Dynamic functional connectivity using state-based dynamic community structure: Method and application to opioid analgesia," *NeuroImage*, vol. 108, pp. 274–291, 2015.

[119] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[120] A. G. Mahyari, D. M. Zoltowski, E. M. Bernat, and S. Aviyente, "A tensor decomposition-based approach for detecting dynamic network states from EEG," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 225–237, 2017.

[121] Q. Zhang, Y. Wang, M. D. Levine, X. Yuan, and L. Wang, "Multisensor video fusion based on higher order singular value decomposition," *Information Fusion*, vol. 24, pp. 54–71, 2015.

[122] R. Guimera and L. A. Nunes Amaral, "Functional cartography of complex metabolic networks," *Nature*, vol. 433, no. 7028, pp. 895–900, 2005.

[123] A. Ponce-Alvarez, G. Deco, P. Hagmann, G. L. Romani, D. Mantini, and M. Corbetta, "Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity," *PLoS Comput. Biol.*, vol. 11, no. 2, pp. 1–23, 2015.

[124] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[125] L. Gauvin, A. Panisson, and C. Cattuto, "Detecting the community structure and activity patterns of temporal networks: A non-negative tensor factorization approach," *PLoS One*, vol. 9, no. 1, pp. 1–13, 2014.

[126] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the royal statistical society*, vol. 57, pp. 289–300, 1995.

[127] J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Methods to detect, characterize, and remove motion artifact in resting state fMRI," *NeuroImage*, vol. 84, pp. 320–341, 2014.

[128] F. I. Karahanolu, C. Caballero-Gaudes, F. Lazeyras, and D. V. D. Ville, "Total activation: fMRI deconvolution through spatio-temporal regularization," *NeuroImage*, vol. 73, pp. 121–134, 2013.

[129] T. Itahashi, T. Yamada, H. Watanabe, M. Nakamura, D. Jimbo, S. Shioda, K. Toriizuka, N. Kato, and R. Hashimoto, "Altered network topologies and hub organization in adults with autism: A resting-state fMRI study," *PLOS ONE*, vol. 9, no. 4, pp. 1–15, 04 2014.

[130] C. P. Chen, C. L. Keown, A. Jahedi, A. Nair, M. E. Pflieger, B. A. Bailey, and R.-A. Mller, "Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism," *NeuroImage: Clinical*, vol. 8, pp. 238 – 245, 2015.

[131] Y. Makino, K. Yokosawa, Y. Takeda, and T. Kumada, "Visual search and memory search engage extensive overlapping cerebral cortices: an fMRI study," *NeuroImage*, vol. 23, no. 2, pp. 525 – 533, 2004.

[132] T. B. Jones, P. A. Bandettini, L. Kenworthy, L. K. Case, S. C. Milleville, A. Martin, and R. M. Birn, "Sources of group differences in functional connectivity: An investigation applied to autism spectrum disorder," *NeuroImage*, vol. 49, no. 1, pp. 401 – 414, 2010.

[133] M. Assaf, K. Jagannathan, V. D. Calhoun, L. Miller, M. C. Stevens, R. Sahl, J. G. O'Boyle, R. T. Schultz, and G. D. Pearlson, "Abnormal functional

connectivity of default mode sub-networks in autism spectrum disorder patients," *NeuroImage*, vol. 53, no. 1, pp. 247–256, 2010.

[134] C. S. Monk, S. J. Peltier, J. L. Wiggins, S.-J. Weng, M. Carrasco, S. Risi, and C. Lord, "Abnormalities of intrinsic functional connectivity in autism spectrum disorders," *NeuroImage*, vol. 47, no. 2, pp. 764–772, 2009.

[135] S. Ha, I.-J. Sohn, N. Kim, H. J. Sim, and K.-A. Cheon, "Characteristics of brains in autism spectrum disorder: Structure, function and connectivity across the lifespan," *Experimental neurobiology*, vol. 24, no. 4, pp. 273–284, 2015.

[136] R. K. Kana, M. A. Patriquin, B. S. Black, M. M. Channell, and B. Wicker, "Altered medial frontal and superior temporal response to implicit processing of emotions in autism," *Autism Research*, vol. 9, no. 1, pp. 55–66, 2016.

[137] A. Nair, R. A. Carper, A. E. Abbott, C. P. Chen, S. Solders, S. Nakutin, M. C. Datko, I. Fishman, and R.-A. Mller, "Regional specificity of aberrant thalamocortical connectivity in autism," *Human Brain Mapping*, vol. 36, no. 11, pp. 4497–4511, 2015.

[138] S. Baron-Cohen, H. Ring, E. Bullmore, S. Wheelwright, C. Ashwin, and S. Williams, "The amygdala theory of autism," *Neuroscience & Biobehavioral Reviews*, vol. 24, no. 3, pp. 355 – 364, 2000.

[139] R. A. Cooper, F. R. Richter, P. M. Bays, K. C. Plaisted-Grant, S. Baron-Cohen, and J. S. Simons, "Reduced hippocampal functional connectivity during episodic memory retrieval in autism," *Cerebral Cortex*, vol. 27, no. 2, pp. 888–902, 2017.

[140] G. A. Orban, D. V. Essen, and W. Vanduffel, "Comparative mapping of higher visual areas in monkeys and humans," *Trends in Cognitive Sciences*, vol. 8, no. 7, pp. 315 – 324, 2004.

[141] J. Salmi, U. Roine, E. Glerean, J. Lahnakoski, T. N. von Wendt, P. Tani, S. Leppmki, L. Nummenmaa, I. Jskelinen, S. Carlson, P. Rintahaka, and M. Sams, "The brains of high functioning autistic individuals do not synchronize with those of others," *NeuroImage: Clinical*, vol. 3, pp. 489 – 497, 2013.

[142] Z. Dogan, T. Blu, and D. V. D. Ville, "Detecting spontaneous brain activity in functional magnetic resonance imaging using finite rate of innovation," in *IEEE 11th International Symposium on Biomedical Imaging*, 2014, pp. 1047–1050.

[143] C. Caballero Gaudes, N. Petridou, S. T. Francis, I. L. Dryden, and P. A. Gowland, "Paradigm free mapping with sparse regression automatically detects single-trial functional magnetic resonance imaging blood oxygenation level dependent responses," *Human Brain Mapping*, vol. 34, no. 3, pp. 501–518, 2013.

[144] C. C. Gaudes, F. I. Karahanolu, F. Lazeyras, and D. V. D. Ville, "Structured sparse deconvolution for paradigm free mapping of functional MRI data," in *9th IEEE International Symposium on Biomedical Imaging*, 2012, pp. 322–325.

[145] I. Khalidov, D. Van De Ville, J. Fadili, and M. Unser, "Activelets and sparsity: a new way to detect brain activation from fMRI data," in *Optical Engineering+ Applications*. International Society for Optics and Photonics, 2007, p. 67010Y.

[146] F. D. Bowman, "Brain imaging analysis," *Annu Rev Stat Appl*, vol. 1, pp. 61–85, 2014.

[147] M. W. Woolrich, B. D. Ripley, M. Brady, and S. M. Smith, "Temporal autocorrelation in univariate linear modeling of fMRI data," *NeuroImage*, vol. 14, no. 6, pp. 1370 – 1386, 2001.

[148] P. Aggarwal, A. Gupta, and A. Garg, *Joint Estimation of Hemodynamic Response Function and Voxel Activation in Functional MRI Data*. Cham: Springer International Publishing, 2015, pp. 142–149.

[149] P. Aggarwal, A. Gupta, and A.Garg, "Joint estimation of activity signal and hrf in fmri using fused lasso," in *IEEE Global Conference on Signal and Information Processing*, 2015, pp. 829–833.

[150] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, "Sparsity and smoothness via the fused lasso," *Journal of the Royal Statistical Society*, vol. 67, no. 1, pp. 91–108, 2005.

[151] F. Ahmad, M. Maqbool, and N. Lee, "Regularization of voxelwise autoregressive model for analysis of functional magnetic resonance imaging data," *Concepts in Magnetic Resonance Part A*, vol. 38A, no. 5, pp. 187–196, 2011.

[152] N. Bazargani and A. Nosratinia, "Joint maximum likelihood estimation of activation and hemodynamic response function for fMRI," *Medical Image Analysis*, vol. 18, no. 5, pp. 711 – 724, 2014.

[153] L. R. Frank, R. B. Buxton, and E. C. Wong, "Estimation of respiration-induced noise fluctuations from undersampled multislice fMRI data," *Magnetic Resonance in Medicine*, vol. 45, no. 4, pp. 635–644, 2001.

[154] D. A. Feinberg, S. Moeller, S. M. Smith, E. Auerbach, S. Ramanna, M. F. Glasser, K. L. Miller, K. Ugurbil, and E. Yacoub, "Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging," *PLoS ONE*, vol. 5, no. 12, p. e15710, 12 2010.

[155] T. Q. Duong, E. Yacoub, G. Adriany, X. Hu, K. Ugurbil, J. T. Vaughan, H. Merkle, and S.-G. Kim, "High-resolution, spin-echo BOLD, and cbf fMRI at 4 and 7 T," *Magnetic Resonance in Medicine*, vol. 48, no. 4, pp. 589–593, 2002.

[156] E. Yacoub, T. Q. Duong, P.-F. Van De Moortele, M. Lindquist, G. Adriany, S.-G. Kim, K. U?urbil, and X. Hu, "Spin-echo fMRI in humans using high spatial resolutions and high magnetic fields," *Magnetic Resonance in Medicine*, vol. 49, no. 4, pp. 655–664, 2003.

[157] N. Harel, "Ultra high resolution fMRI at ultra-high field," *NeuroImage*, vol. 62, no. 2, pp. 1024 – 1028, 2012.

[158] E. Yacoub, N. Harel, and K. Uurbil, "High-field fMRI unveils orientation columns in humans," *Proc Natl Acad Sci USA*, vol. 105, no. 30, pp. 10 607–10 612, 2008.

[159] N. K. Logothetis, H. Merkle, M. Augath, T. Trinath, and K. Ugurbil, "Ultra high-resolution fMRI in monkeys with implanted RF coils," *Neuron*, vol. 35, no. 2, pp. 227–242, 2002.

[160] J. Pfeuffer, P.-F. van de Moortele, E. Yacoub, A. Shmuel, G. Adriany, P. Andersen, H. Merkle, M. Garwood, K. Ugurbil, and X. Hu, "Zoomed functional imaging in the human brain at 7 tesla with simultaneous high spatial and high temporal resolution," *NeuroImage*, vol. 17, no. 1, pp. 272–286, 2002.

[161] P.-H. Wu, P.-H. Tsai, M.-L. Wu, T.-C. Chuang, Y.-Y. Shih, H.-W. Chung, and T.-Y. Huang, "High spatial resolution brain functional MRI using submillimeter balanced steady-state free precession acquisitiona)," *Medical Physics*, vol. 40, no. 12, 2013.

[162] T. Hugger, B. Zahneisen, P. LeVan, K. J. Lee, H.-L. Lee, M. Zaitsev, and J. Hennig, "Fast undersampled functional magnetic resonance imaging using nonlinear regularized parallel image reconstruction," *PLoS ONE*, vol. 6, no. 12, p. e28822, 2011.

[163] L. Chaari, P. Ciuciu, S. Mériaux, and J.-C. Pesquet, "Spatio-temporal wavelet regularization for parallel MRI reconstruction: application to functional MRI," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 27, no. 6, pp. 509–529, 2014.

[164] W. Lu, T. Li, I. Atkinson, and N. Vaswani, "Modified-cs-residual for recursive reconstruction of highly undersampled functional MRI sequences," in *Proc Int Conf Image Process*, Sept 2011, pp. 2689–2692.

[165] S. Yan, L. Nie, C. Wu, and Y. Guo, "Linear dynamic sparse modelling for functional MR imaging," *Brain Inform*, vol. 1, no. 1, pp. 11–18, 2014.

[166] P. K. Han, S.-H. Park, S.-G. Kim, and J. C. Ye, "Compressed sensing for fMRI : feasibility study on the acceleration of non-EPI fMRI at 9.4 T," *Biomed Res Int*, vol. 2015, no. 131926, pp. 1–24, 2015.

[167] H. Jung and J. C. Ye, "Performance evaluation of accelerated functional MRI acquisition using compressed sensing," in *Proc IEEE Int Symp Biomed Imaging*, 2009, pp. 702–705.

[168] X. Zong, J. Lee, A. J. Poplawsky, S.-G. Kim, and J. C. Ye, "Compressed sensing fMRI using gradient-recalled echo and EPI sequences," *NeuroImage*, vol. 92, pp. 312 – 321, 2014.

[169] C. Chavarrías, J. Abascal, P. Montesinos, and M. Desco, "Exploitation of temporal redundancy in compressed sensing reconstruction of fMRI studies with a prior-based algorithm," *Medical physics*, vol. 42, no. 7, pp. 3814–3821, 2015.

[170] D. J. Holland, C. Liu, X. Song, E. L. Mazerolle, M. T. Stevens, A. J. Sederman, L. F. Gladden, R. C. N. D'Arcy, C. V. Bowen, and S. D. Beyea, "Compressed sensing reconstruction improves sensitivity of variable density spiral fMRI," *Magnetic Resonance in Medicine*, vol. 70, no. 6, pp. 1634–1643, 2013.

[171] Z. Fang, N. Van Le, M. Choy, and J. H. Lee, "High spatial resolution compressed sensing functional MRI," *Magn Reson Med*, vol. 76, no. 2, pp. 440–455, 2016.

[172] M. Chiew, N. N. Graedel, J. A. McNab, S. M. Smith, and K. L. Miller, "Accelerating functional MRI using fixed-rank approximations and radial-cartesian sampling," *Magnetic Resonance in Medicine*, 2016.

[173] V. Singh, A. Tewfik, and D. Ress, "Under-sampled functional MRI using low-rank plus sparse matrix decomposition," in *Proc IEEE Int Conf Acoust Speech Signal Process*, 2015, pp. 897–901.

[174] S. Yan, L. Nie, C. Wu, and Y. Guo, "Linear dynamic sparse modelling for functional MR imaging," *Brain Informatics*, vol. 1, no. 1-4, pp. 11–18, 2014.

[175] M. Chiew, S. M. Smith, P. J. Koopmans, N. N. Graedel, T. Blumensath, and K. L. Miller, "k-t FASTER: Acceleration of functional fMRI data acquisition using low rank constraints," *Magnetic Resonance in Medicine*, vol. 74, no. 2, pp. 353–364, 2015.

[176] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[177] Z.-P. Liang, "Spatio-temporal imaging with partially separable functions," in *4th IEEE International Symposium Biomedical Imaging*, 2007, pp. 988–991.

[178] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[179] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: kt SLR," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1042–1054, 2011.

[180] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and bregman iterative methods for matrix rank minimization," *Mathematical Programming*, vol. 128, no. 1-2, pp. 321–353, 2011.

[181] H. Zhu, Y. Xiao, and S.-Y. Wu, "Large sparse signal recovery by conjugate gradient algorithm based on smoothing technique," *Computers & Mathematics with Applications*, vol. 66, no. 1, pp. 24 – 32, 2013.

[182] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[183] B. Moore, R. Nadakuditi, and J. Fessler, "Improved robust pca using low-rank denoising with optimal singular value shrinkage," in *IEEE Workshop on Statistical Signal Processing*, 2014, pp. 13–16.

[184] R. Nadakuditi, "Optshrink: An algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 3002–3018, 2014.

[185] J. M. Moran, E. Jolly, and J. P. Mitchell, "Social-cognitive deficits in normal aging," *The Journal of Neuroscience*, vol. 32, no. 16, pp. 5553–5561, 2012.

[186] B. B. Biswal, M. Mennes, X.-N. Zuo, S. Gohel, C. Kelly, S. M. Smith, C. F. Beckmann, J. S. Adelstein, R. L. Buckner, S. Colcombe, A.-M. Dogonowski, M. Ernst, D. Fair, M. Hampson, M. J. Hoptman, J. S. Hyde, V. J. Kiviniemi, R. Ktter, S.-J. Li, C.-P. Lin, M. J. Lowe, C. Mackay, D. J. Madden, K. H.

Madsen, D. S. Margulies, H. S. Mayberg, K. McMahon, C. S. Monk, S. H. Mostofsky, B. J. Nagel, J. J. Pekar, S. J. Peltier, S. E. Petersen, V. Riedl, S. A. R. B. Rombouts, B. Rypma, B. L. Schlaggar, S. Schmidt, R. D. Seidler, G. J. Siegle, C. Sorg, G.-J. Teng, J. Veijola, A. Villringer, M. Walter, L. Wang, X.-C. Weng, S. Whitfield-Gabrieli, P. Williamson, C. Windischberger, Y.-F. Zang, H.-Y. Zhang, F. X. Castellanos, and M. P. Milham, "Toward discovery science of human brain function," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4734–4739, 2010.

[187] S. Zhang, K. T. Block, and J. Frahm, "Magnetic resonance imaging in real time: Advances using radial flash," *Journal of Magnetic Resonance Imaging*, vol. 31, no. 1, pp. 101–109, 2010.

[188] P. Aggarwal, P. Shrivastava, T. Kabra, and A. Gupta, "Optshrink lr+s: accelerated fmri reconstruction using non-convex optimal singular value shrinkage," *Brain Inform*, vol. 4, no. 1, pp. 65–83, 2017.

[189] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid mr imaging," *Magn Reson Med*, vol. 58, no. 6, pp. 1182–1195, 2007.

[190] M. A. Lindquist, "The statistical analysis of fMRI data," *Statist. Sci.*, vol. 23, no. 4, pp. 439–464, 2008.

[191] K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. E. Penny, *Statistical Parametric Mapping: The analysis of functional brain images*. Academic Press, 2006.

[192] V. D. Calhoun, V. K. Potluru, R. Phlypo, R. F. Silva, B. A. Pearlmutter, A. Caprihan, S. M. Plis, and T. Adalı, "Independent component analysis for brain fMRI does indeed select for maximal independence," *PloS one*, vol. 8, no. 8, p. e73309, 2013.

[193] V. Kiviniemi, T. Starck, J. Remes, X. Long, J. Nikkinen, M. Haapea, J. Veijola, I. Moilanen, M. Isohanni, Y.-F. Zang *et al.*, "Functional segmentation of the brain cortex using high model order group PICA," *Hum Brain Mapp*, vol. 30, no. 12, pp. 3865–3886, 2009.

[194] A. Abou-Elseoud, T. Starck, J. Remes, J. Nikkinen, O. Tervonen, and V. Kiviniemi, "The effect of model order selection in group PICA," *Human brain mapping*, vol. 31, no. 8, pp. 1207–1216, 2010.

[195] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[196] F. Esposito, E. Formisano, E. Seifritz, R. Goebel, R. Morrone, G. Tedeschi, and F. Di Salle, "Spatial independent component analysis of functional MRI time-series: To what extent do results depend on the algorithm used?" *Hum Brain Mapp*, vol. 16, no. 3, pp. 146–157, 2002.

[197] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking whole-brain connectivity dynamics in the resting state," *Cereb Cortex*, vol. 24, no. 3, pp. 663–676, Mar 2014.