Brain Connectivity Network Modeling using fMRI Signals

by

Aiping Liu

B.Sc., University of Science and Technology of China, 2009
M.A.Sc., The University of British Columbia, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL
STUDIES
(Electrical and Computer Engineering)

The University of British Columbia
(Vancouver)

May 2016

© Aiping Liu, 2016
Abstract

Functional magnetic resonance imaging (fMRI) is one of the most popular non-invasive neuroimaging technologies, which examines human brain at relatively good spatial resolution in both normal and disease states. In addition to the investigation of local neural activity in isolated brain regions, brain connectivity estimated from fMRI has provided a system-level view of brain functions. Despite recent progress on brain connectivity inference, there are still several challenges. Specifically, this thesis focuses on developing novel brain connectivity modeling approaches that can deal with particular challenges of real biomedical applications, including group pattern extraction from a population, false discovery rate control, incorporation of prior knowledge and time-varying brain connectivity network modeling.

First, we propose a multi-subject, exploratory brain connectivity modeling approach that allows incorporation of prior knowledge of connectivity and determination of the dominant brain connectivity patterns among a group of subjects. Furthermore, to integrate the genetic information at the population level, a framework for genetically-informed group brain connectivity modeling is developed.

We then focus on estimating the time-varying brain connectivity networks. The temporal dynamics of brain connectivity assess the brain in the additional temporal dimension and provide a new perspective to the understanding of brain functions. In this thesis, we develop a sticky weighted time-varying model to investigate the time-dependent brain connectivity networks. As the brain must strike a balance between stability and flexibility, purely assuming that brain connectivity is static or dynamic may be unrealistic. We therefore further propose making joint inference of time-invariant connections and time-varying coupling patterns by employing a
multitask learning model.

The above proposed methods have been applied to real fMRI data sets, and the disease induced changes on the brain connectivity networks have been observed. The brain connectivity study is able to provide deeper insights into neurological diseases, complementing the traditional symptom-based diagnostic methods. Results reported in this thesis suggest that brain connectivity patterns may serve as potential disease biomarkers in Parkinson’s Disease.
Preface

The research work in this thesis was jointly initiated by Dr. Z. Jane Wang, Dr. Martin J. McKeown and me. The thesis is based on a collection of manuscripts that have been accepted or submitted for publication in book chapter, international peer-reviewed journals and conferences.

Chapter 1 and Chapter 2 are based on the following manuscripts:


In the above manuscripts, the development of analysis techniques was jointly conducted by the author and Dr. Junning Li under the supervision of Dr. Z. Jane Wang and Dr. Martin J. McKeown. The experimental data were provided by Dr. Martin J. McKeown of PPRC (Clinical Research Ethics Board, H04-70177). I was responsible for technical literature survey, simulation, real data application and manuscripts
writing. The neurological interpretation of fMRI application was written with the guidance of Dr. Martin J. McKeown.

Chapter 3 is based on the following manuscript:


I was responsible for the development of algorithm, numerical simulation and real data application. Asymptotic properties of the algorithm was analyzed by Dr. Xi-aohui Chen. The genetic and fMRI data were collected, pre-processed and described by Dr. Qi Xu. I prepared the manuscript with suggestions and subsequent revisions from Dr. Silke Appel-Cresswell, Dr. Z. Jane Wang and Dr. Martin J. McKeown.

Chapter 4 is based on the following manuscripts:


In the above manuscripts, I was responsible for the model design, numerical simulation and real fMRI application. The experimental data were provided by Dr. Martin J. McKeown of PPRC (Clinical Research Ethics Board, H04-70177). I drafted the manuscript with subsequent editorial input from Dr. Xun Chen, Dr. Z. Jane Wang and Dr. Martin J. McKeown.

A version of Chapter 5 has been submitted for publication:

- Aiping Liu, Xun Chen, Xiaojuan Dan, Martin J. McKeown and Z. Jane Wang, “ A Combined Static and Dynamic Model for Resting State fMRI
Brain Connectivity Networks: Application to Parkinson’s Disease”, submitted, 2015.

I was responsible for the algorithm development, numerical simulation, real data application and paper writing. The data description was originally written by Dr. Martin J. McKeown (Clinical Research Ethics Board, H04-70177). Coauthors have provided editorial input for the manuscript.
# Table of Contents

Abstract ................................................................. ii

Preface ................................................................. iv

Table of Contents ...................................................... vii

List of Tables .......................................................... x

List of Figures .......................................................... xi

Glossary ................................................................. xiv

Acknowledgments ....................................................... xvi

1 Introduction and Overview ........................................ 1
  1.1 Functional Magnetic Resonance Imaging and Brain Activity Study 2
  1.2 Brain Connectivity Network Modeling ............................ 6
    1.2.1 Group Level Brain Connectivity Network Modeling ... 10
    1.2.2 Time Varying Brain Connectivity Network Modeling ... 12
    1.2.3 Brain Connectivity Network Interpretation .......... 14
  1.3 Research Objectives and Methodologies .................... 15
  1.4 Thesis Outline ................................................... 17

2 False Discovery Rate Controlled, Prior Knowledge Incorporated Group Brain Connectivity Modeling .......................... 19
  2.1 Introduction ....................................................... 19
2.2 Methods .................................................. 21
2.3 Real Application .............................. 25
2.4 Conclusion and Discussion ............... 27

3 A Genetically-informed, Group Brain Connectivity Modeling Framework .................. 29
3.1 Introduction ............................................ 29
3.2 Methods .............................................. 32
    3.2.1 Overlapped Group Fused Model .............. 33
    3.2.2 Model Selection and Degree of Freedom ...... 36
    3.2.3 Implementation of Brain Connectivity Modeling and Statistical Inference .......... 39
    3.2.4 Graph Theoretical Analysis .................... 40
3.3 Simulations .......................................... 41
3.4 Real Application in Schizophrenia ......... 44
    3.4.1 Subjects ......................................... 44
    3.4.2 Selection and Genotyping of SNPs ............. 44
    3.4.3 fMRI Data ....................................... 46
    3.4.4 Results .......................................... 46
3.5 Conclusion and Discussion .................... 51
3.6 Asymptotic Properties ......................... 53

4 A Sticky Weighted Time Varying Model for Resting State Brain Connectivity Estimation .......... 55
4.1 Introduction .......................................... 55
4.2 Methods .............................................. 58
    4.2.1 Sticky Weighted Time Varying Model .......... 58
    4.2.2 Model Selection .................................. 60
    4.2.3 Statistical Analysis ............................. 61
4.3 Simulations .......................................... 62
4.4 Real Application .............................. 67
    4.4.1 Subjects and fMRI Resting State Data Set .... 67
    4.4.2 Results .......................................... 69
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1</td>
<td>The name and abbreviation of 12 selected brain ROIs.</td>
<td>26</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Implementation of overlapped group fused model.</td>
<td>37</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>Implementation of model selection.</td>
<td>39</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>The index and name of 52 selected brain ROIs. 'L' represents the brain left side and 'R' represents the brain right side.</td>
<td>45</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Implementation of sticky weighted time varying model.</td>
<td>61</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>The index and name of 48 selected brain ROIs. 'L' represents the brain left side and 'R' represents the brain right side.</td>
<td>68</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Implementation of combined static and dynamic model for time varying coefficients estimation.</td>
<td>85</td>
</tr>
<tr>
<td>Table 5.2</td>
<td>The index and name of 54 selected brain ROIs. 'L' represents the brain left side and 'R' represents the brain right side.</td>
<td>91</td>
</tr>
</tbody>
</table>
# List of Figures

| Figure 1.1   | BOLD signal mechanism for fMRI.      | 2       |
| Figure 1.2   | Example of fMRI block design experiement. | 3       |
| Figure 1.3   | Example of modeling the functional brain connectivity network. | 5       |
| Figure 1.4   | Illustration of pairwise dependence and conditional dependence. | 8       |
| Figure 1.5   | The overview of challenges and objectives of this thesis work. | 16      |
| Figure 2.1   | Learned brain connectivity for normal and patient groups. | 28      |
| Figure 3.1   | The proposed framework for genetically-informed group brain connectivity modeling. | 32      |
| Figure 3.2   | The comparison of estimated degree of freedom and true degree of freedom. | 38      |
| Figure 3.3   | One example of the heterogeneous overlapped group structure used in the simulation. | 42      |
| Figure 3.4   | Simulation results: Simulation is designed to recover a heterogeneous group network based on measures such as Type I error rate, detection power, FDR and F1 score as the number of data points increase. | 43      |
| Figure 3.5   | The comparisons between normal control and Schizophrenia groups. | 47      |
| Figure 3.6   | The comparisons of node degree at Left Putamen region in different conditions. | 48      |
| Figure 3.7   | Examples of brain connectivity networks in the group modeling. | 49      |
Figure 3.8 The comparisons between two groups with rs2391191 AG and AA genotypes. ................................. 50

Figure 4.1 Results for the first simulation. ......................... 63
Figure 4.2 Results for the second simulation. .................... 64
Figure 4.3 Results for the third simulation. ....................... 65
Figure 4.4 The comparisons of F1 scores. ......................... 66
Figure 4.5 The comparisons of density and network variations with different tuning parameters. ....................... 70
Figure 4.6 Time varying brain connectivity networks learned with fixed parameters for one typical control subject at different time points. 74
Figure 4.7 Time varying brain connectivity networks learned with fixed parameters for one typical PD subject at different time points. 75
Figure 4.8 The comparison of network variation between the normal and PD groups. ................................. 76
Figure 4.9 The comparison of averaged time period between the normal and PD groups. ................................. 76
Figure 4.10 The comparison of switching ratio between the normal and PD group. .................................. 77
Figure 4.11 Connections that consistently appear in at least one time point in all subjects in (a) the control group and (b) the PD group. 78

Figure 5.1 Simulation example for CSDM. ....................... 88
Figure 5.2 The comparison of the L2-loss of the estimated coefficients for (a) simulation 1 and (b) simulation 2. ............... 89
Figure 5.3 Examples of static eigenconnectivities (fixed density). .... 92
Figure 5.4 Examples of dynamic eigenconnectivities (fixed density). ... 93
Figure 5.5 The comparisons of static eigenconnectivity networks and averaged contributions of control and PD groups to the static eigenconnectivity networks. ......................... 94
Figure 5.6 The comparisons of dynamic eigenconnectivity networks and averaged contributions of control and PD groups to the dynamic eigenconnectivity networks. ......................... 95
Figure 5.7 The comparison of network variations between control, PDon and PDoff groups.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIC</td>
<td>Bayesian information criterion</td>
</tr>
<tr>
<td>BOLD</td>
<td>Blood-oxygen-level dependence</td>
</tr>
<tr>
<td>CSDM</td>
<td>Combined static and dynamic model</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>DAOA</td>
<td>D-amino acid oxidase activator</td>
</tr>
<tr>
<td>DCM</td>
<td>Dynamic causal modeling</td>
</tr>
<tr>
<td>DF</td>
<td>Degree of freedom</td>
</tr>
<tr>
<td>DM</td>
<td>Dirty model</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalography</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>GC</td>
<td>Granger causality</td>
</tr>
<tr>
<td>GIMME</td>
<td>Group iterative multiple model estimation</td>
</tr>
<tr>
<td>GLM</td>
<td>General Linear Model</td>
</tr>
<tr>
<td>FMRI</td>
<td>Functional magnetic resonance imaging</td>
</tr>
<tr>
<td>ICA</td>
<td>Independent component analysis</td>
</tr>
<tr>
<td>LASSO</td>
<td>Least absolute selection and shrinkage operator</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>MAR</td>
<td>Multivariate autoregressive model</td>
</tr>
<tr>
<td>MEG</td>
<td>Magnetoencephalogram</td>
</tr>
<tr>
<td>OGFM</td>
<td>Overlapped group fused model</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s disease</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of interest</td>
</tr>
<tr>
<td>SEM</td>
<td>Structural equation modeling</td>
</tr>
<tr>
<td>SNPS</td>
<td>Single nucleotide polymorphisms</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular value decomposition</td>
</tr>
<tr>
<td>SWTVC</td>
<td>Sticky weighted time varying</td>
</tr>
<tr>
<td>WTV</td>
<td>Weighted time varying</td>
</tr>
</tbody>
</table>
Acknowledgments

It’s a happy journey walking towards my PhD degree, with support and encouragement from numerous people. Needless to say, I thank all of them. In particular, I want to take this opportunity to acknowledge some of them.

First and foremost, I would like to express my great appreciation to my supervisors, Dr. Z. Jane Wang and Dr. Martin J. McKeown, who are the best supervisors I could ever expect. Thanks for their constant academic and personal support over the years. Without their enlightening guidance, this thesis would never have been written. Many thanks go to Dr. Junning Li and Dr. Xiaohui Chen for providing me a lot of valuable discussions and assistance in my research.

I want to thank my thesis examination committee members for their time and efforts. I would also like to acknowledge Dr. Silke Appel-Cresswell, Dr. Nazanin Baradaran, Suejin Lin and Sun Nee Tan from Pacific Parkinson’s Research Centre for providing me with the data and helpful discussions regarding the Parkinson’s Disease studies.

I am thankful to all my labmates, colleges and coauthors, Dr. Joyce Chiang, Dr. Xudong Lv, Dr. Chen He, Dr. Zhenyu Guo, Liang Zou, Huan Qi, Jiannan Zheng and Yiming Zhang for creating supportive research environment and for the fun we have had together. I also would like to thank my friends, Haolu Zhang, Shumin Wang, Dr. Di Xu, Dr. Qiang Tang, Dr. Haoming Li, Dr. Erkai Liu, Dr. Mengzhe Shen, Yifeng Zhou, and many others from the UBC, for the friendship and moral support they provided.

I am deeply indebted to my parents for their unconditional love and constant encouragement. A special thank goes to my husband, Dr. Xun Chen for his understanding, support and love. Dr. Xun Chen is my friend, collaborator and great
partner. Ryan Y. Chen, my son, thank you for your sweet smile and making me a multitask learner.

This work was supported by the Four Year Doctoral Fellowship (4YF) program of UBC, Natural Sciences and Engineering Research Council (NSERC) of Canada and Canadian Institutes of Health Research (CIHR) Grants.
Chapter 1

Introduction and Overview

Human brain is considered as one of the most complex systems and attracts many efforts in studying its structures and functions, where neuroimaging technologies are shown to be the powerful tools. Neuroimaging technologies such as Electroencephalography (EEG), Magnetoencephalogram (MEG), Computed tomography (CT), Positron emission tomography (PET) and Functional magnetic resonance imaging (fMRI), are becoming prevalent in recent neuroscience studies as they are capable to non-invasively examine brain activities in vivo. In particular, fMRI, which measures brain functions at better spatial resolution than other functional neuroimaging modalities, has been widely employed in biomedical applications, achieving remarkable progress in understanding our brain.

The advances in neuroimaging technologies have also encouraged the development of suitable data modeling approaches. The conventional methods focus on functional segregation studies which try to identify the functional specialization of particular brain regions. While most cognitive states involve the coherent activation of several functionally specialized regions, investigating brain activity in isolated brain areas may be insufficient. As a result, an alternative organizational principal, functional integration, has been introduced to characterize brain activity. This has led to the emergence of the concept of brain connectivity which estimates the interaction patterns of a set of distant brain regions.

Brain connectivity has provided a system-level view of how the brain works and facilitated the exploration of brain functions in normal states. In addition, it
has expanded our insights into related brain diseases. For instance, neurological disorders such as Parkinson’s disease (PD) have enormous impact on the whole population and the neuroimaging studies benefit the detection of these diseases in their early stage [40, 58]. The discovery of altered connectivity patterns is promising in assisting the disease diagnosis, severity detection and medication evaluation. In this thesis, we are thus particularly interested in the estimation of brain connectivity. A set of novel network modeling approaches are developed with the ultimate goal to investigate the disease induced effects on brain connectivity patterns.

In the remainder of this chapter, we first introduce the background on fMRI and brain activity study in Section 1.1. Section 1.2 reviews a list of popular brain connectivity network modeling approaches, and discusses on the network interpretations briefly. Section 1.3 introduces the research objectives and methodologies. The thesis outline is finally presented in Section 1.4.

1.1 Functional Magnetic Resonance Imaging and Brain Activity Study

fMRI is one of the most widely used neuroimaging technologies which measures relative changes in deoxygenated haemoglobin in the form of Blood-oxygen-level dependence (BOLD) signals as a result of ongoing brain activity [108]. When a brain region becomes active, more oxygen contained in increasing blood flow is delivered to the neurons by haemoglobin which could exhibit different magnetic properties during oxygenation to deoxygenation. The mechanism of BOLD signal is shown in Fig. 1.1. It is an indirect marker of neural activity, as it is based on focal blood flow modulated by local brain metabolism.

![Figure 1.1: BOLD signal mechanism for fMRI.](image)

The majority of fMRI analyses to date have examined changes in BOLD amplitude as a result of external stimulus, which identify the functional specialization
of a particular brain region (functional segregation).

**Figure 1.2:** Example of fMRI block design experiment. A subject performs a certain task, e.g. bulb squeezing, when BOLD signals are collected. Alternating between task and rest generates the images required for inferring brain activity. The voxel time courses are then extracted for subsequent analysis such as activation detection. The figure is adapted from [105].

The exact amplitude of the BOLD signal is not directly comparable across subjects since BOLD fMRI is a contrast imaging technique (i.e. unitless), but not a quantitative imaging technique. It is standard to determine the relative differences in BOLD signal amplitude across two tasks (e.g. bulb squeezing vs rest). Traditionally, this has been done in a block design, where the subjects may, for example,
squeeze a bulb for 20-30s followed by 20-30s of rest, and the cycle is repeated (Fig. 1.2). This alternating of experimental and control tasks will tend to make analysis methods more robust against any erroneous interpretations based on non-neuronal slow drifts in the signal and/or fatigue effects that would bias interpretations if the experimental task was done only at the beginning or end of the run. Block-related designs generally assume that any hemodynamic response to neuronal activity is saturated by rapidly and repeatedly performing the same task within a block.

An alternate approach is to assess the BOLD response to a single stimulus, such as a simple motor movement. This has the advantage of comparing stimuli that might have the same loci of activation but different amplitudes of neuronal (and subsequent hemodynamic) response, but has the disadvantage of significantly prolonging scanning time, as the hemodynamic response must decay sufficiently before the next stimulus can be presented.

To infer task-related activation or specialization, hypothesis-driven methods have been widely adopted to examine changes in BOLD amplitude, like the standard General Linear Model (GLM) [48].

The spatial patterns of the BOLD signals could also be altered by task-related activity. For instance, spatial or “3D texture descriptors” such as 3D Moment Invariants (3DMI) that are invariant to the exact coordinate system, can be used to examine the effects of task-related changes in fMRI [106].

While most cognitive states involve the coherent activation of several functionally specialized regions, the concept of brain connectivity has emerged which estimates the interaction patterns between discrete brain regions [13]. Brain connectivity is a promising way to investigate the functional coordination and has provided a system-level view of how the brain works.

A more recent type of paradigm, the so-called “resting-state” study, whereby individuals are instructed to simply rest quietly with their eyes closed and remain awake, is well suited for the brain connectivity studies. In this condition, spatially widespread, unprompted activity not attributable to specific external stimuli can be observed. Statistical analyses on the spontaneous fluctuations in the BOLD signal can then be performed to detect the intrinsic activities of human brain [46]. Resting-state fMRI has emerged as a powerful tool for discovery science which is capable of generating detailed maps of complex neural systems [14].
Figure 1.3: Example of modeling the functional brain connectivity network at the ROI level. The functional connectivity network can be represented as a graph with each node representing the brain ROI and each edge representing the relationship between two ROIs. (A). A collection of brain ROIs are firstly selected. (B). The time course of each ROI is extracted for pairwise or multivariate analysis. (C). The brain connectivity network is estimated. The figure is adapted from [146].

To reveal the whole brain organizations, brain connectivity studies, which we are interested in, play significant roles. A typical example of brain connectivity network modeling at the Region of interest (ROI) level is shown in Fig. 1.3. The interactions between brain regions have been increasingly recognized as being important for understanding normal brain functions and the pathophysiology of many
related diseases. It has been suggested that some disorders such as Schizophrenia and PD are related with the dysfunctions of connectivity networks [135]. Compared with traditional methods for functional specialization analysis, brain connectivity network modeling allows the exploration of the cooperation between multiple brain regions and the extraction of more informative features of the neural systems.

1.2 Brain Connectivity Network Modeling

The study of brain connectivity has enhanced our understanding of the underlying brain function in both normal and disease states, in addition to the traditional approaches focusing on regional activity detection.

Brain connectivity patterns have been inferred based on bivariate analysis such as correlation threshold [19], frequency-based coherence analysis [125], mutual information [128], Granger causality (GC) derived from bivariate autoregressive models [54] and so on. Multivariate models, including Multivariate autoregressive model (MAR) [62], Structural equation modeling (SEM) [97] and Dynamic causal modeling (DCM) [47] have also been proposed to assess brain connectivity. Other commonly-used approaches include linear decomposition methods such as Independent component analysis (ICA) [99, 100], sparse induced modeling [24, 149] and Bayesian network models [82, 117]. There are different (but not mutually exclusive) ways in which all these proposed brain connectivity modeling approaches can be categorized: exploratory vs confirmatory, linear vs nonlinear, directional connectivity vs bidirectional and voxel level vs ROI level modeling.

The most straightforward approach, the correlation threshold method [19] estimates how strongly two brain regions interact with each other by calculating the correlation coefficient between their activities. If the correlation coefficient is sufficiently high that it is not possible only coming from the randomness, the two regions are considered associated with each other. Correlation threshold is statistically rigorous. However, standard pairwise correlation cannot distinguish between direct and indirect interactions (whether two components interact directly or indirectly through a third component).

Instead of simply making inference of the co-varied brain regions, partial correlation can be employed to estimate if one brain region has direct influence over
another \cite{96}, as it measures the normalized correlation with the effect of all other variables being removed. The application of partial correlation in inferring the relationship between two variables is based on the conditional independence test. The definition of conditional independence is as follows: \(X\) and \(Y\) are conditional independent given \(Z\) if and only if \(P(\text{XY}|Z) = P(X|Z)P(Y|Z)\). It is similar to the pair-wise independence definition \(P(\text{XY}) = P(X)P(Y)\), but conditional on a third variable \(Z\). Note that pairwise independence does not imply conditional independence, and \textit{vice versa} (Fig. 1.4). For instance, the activities of two brain regions \(A\) and \(B\) are commonly driven by that of a third region \(C\), then the activities of \(A\) and \(B\) maybe correlated in pairwise fashion. But if the influence from \(C\) is removed, their activities will become independent, as shown in Fig. 1.4 (b). On the other hand, if the activities of two brain regions are correlated even after all possible influence from other regions are removed (as shown in Fig. 1.4 (c)), then very likely there is a direct connection between them (i.e., the two regions are conditionally dependent). Therefore, the conditional dependence implies that two brain regions are directly connected. It is a key concept in multivariate analyses such as graphical modeling, where two nodes are connected if and only if the corresponding variables are not conditionally independent.

Different from the pairwise analysis between two variables discussed before, linear decomposition methods, such as Principal component analysis (PCA) or ICA \cite{100} can be used to assess which voxels have a tendency to coactive. They are the data driven approaches which are suitable in the applications when the models of brain activity are not available. For instance, ICA decomposes BOLD patterns into spatially independent maps and their associated time courses. It was a significant shift from the traditional hypothesis-based approach for fMRI analysis when first proposed \cite{100}. Because no time course of activation needs to be specified a priori, it is ideally suited to assess resting-state fMRI data \cite{38, 43} or in situations where the anticipated activation patterns may deviate from normal. Thus ICA analysis of fMRI has been widely used to study clinical populations, e.g. Alzheimer’s disease \cite{56}, depression \cite{57}, schizophrenia \cite{69}, mild cognitive impairments \cite{116} and non-communicative brain damaged patients \cite{147}.

Similar to the linear decomposition methods, clustering techniques, such as the self-organizing map (SOM) \cite{45, 113}, k-means clustering \cite{9}, hierarchical clus-
Figure 1.4: Illustration of pairwise dependence and conditional dependence.

tering [34] and graph clustering [64], are also data-driven approaches to explore unknown interactions between brain regions. They are based on the assumption that if the time courses of voxels and/or ROIs tend to cluster together, they likely have interactions between them. Clustering is usually implemented with fast and heuristic algorithms and thus is suitable for large-scale problems where it is difficult to perform rigorous statistical analysis. However, the data-driven feature also brings disadvantages, as certain algorithms may either fall in local optimal solutions or their convergence cannot be proved. Statistical criteria such as specificity and sensitivity generally cannot be theoretically analyzed for clustering methods.

It is noticed that aforementioned approaches can not determine the direction of the connections. To estimate the direction, one popular approach employed is Granger Causality (GC) [54]. It is based on the statistical hypothesis testing for determining whether one time series can be used to at least partially predict another. Another way is using Patel’s conditional dependence measure which estimates the connectivity between two variables by measuring the imbalance of condi-
tional probability between them [112]. These two methods are usually considered as confirmatory methods due to the requirement of prior knowledge on the network structures.

Multivariate autoregression (MAR) model [62], structural equation model (SEM) [97] and dynamic causal model (DCM) [47] are popular multivariate regression models proposed to estimate brain connectivity. MAR model focuses on the lagged interactions and incorporates the covariance as well as temporal information across the samples. It represents one sample of a time series as the weighted sum of its previous samples,

\[ Y_t = \sum_{p=1}^{P} A_p Y_{t-p} + e_p \]  \hspace{1cm} (1.1)

where \( Y_t \) is the K-dimensional vector denoting BOLD signal values of K ROIs at time \( t \), \( A_p \) is the MAR coefficient matrix at time lag \( p \), and \( e_t \) is the noise term. Different from the MAR model that infers the lagged information, SEM estimates the simultaneous interactions between brain ROIs,

\[ Y_t = M Y_t + e_t \]  \hspace{1cm} (1.2)

where \( M \) represents the connection strength matrix. DCM is different from them by accommodating the nonlinear and dynamic activities between brain regions. It models neural activities as hidden variables. Multivariate regression models are considered as statistically rigorous, flexible, and supported by many well-developed algorithms. However, a major drawback of these models is that the computational cost grows exponentially with the number of ROIs. This typically restricts their use to confirmatory studies examining a few ROIs.

fMRI has relatively few time points and the number of ROIs may be large. Modeling the brain connectivity using fMRI signals is a difficult statistical inference problem. The sparsity assumption thus has been made on the connectivity networks which favors both the computational efficiency and biological interpretations. Least absolute selection and shrinkage operator (LASSO) based approaches combine computational efficiency with the ability to deal with high dimensionality, and hence such methods including the robust LASSO and the sparse inverse covari-
ance estimation have been proposed [24, 66]. Sparse dictionary learning techniques have also been developed to assess the functional connectivity [148, 150]. For instance, by fitting a model to all the variables, the graphical LASSO estimates a sparse network, whereby ROIs are represented as vertices and variable-wise relationships are represented as edges [24].

Graphical models are suitable for modeling the brain connectivity as their graphical nature assists in the biological interpretation of the connectivity patterns. The Linear, Non-Gaussian, Acyclic model (LiNGAM) is a causal graphical model that assumes the variables have non-Gaussian distributions of non-zero variances, and identify the brain connectivity structure as a directed acyclic graph [130]. The Bayesian network models which encode the conditional dependence/independence into the graph, are the most popular graphical models proposed for studying the interactions between brain regions [82, 118]. They are capable to handle relatively large number of brain regions and provide more rigorous model selection procedure [104].

It’s worth noting that a lot of novel approaches have been proposed, aiming to enhance the network modeling with different assumptions made. Each method has its own advantages and limitations. Based on different scenarios, specific approaches may be designed to meet the demands in their applications [131]. However, due to the lack of underlying truth of brain interaction mechanisms, revealing brain connectivity patterns is still on the exploratory stage with the goal to uncover the large scale brain connectivity maps.

1.2.1 Group Level Brain Connectivity Network Modeling

Rather than focusing on an individual subject, the biomedical researches typically involve a group of subjects in order to make inference about a population. In addition, as the number of fMRI data sets increasing, it offers the opportunities and also becomes necessary to investigate the brain connectivity networks at the group level.

However, group level approaches for modeling brain connectivity need to handle not only the variances and correlations across subjects, but also the fact that the exact structures of brain connectivity may vary across individuals. The large
inter-subject variability poses challenges on the fMRI data processing and the underlying differences existing may lead to draw the false connections.

Several methods have been developed to infer group connectivity networks in neuroimaging. As one of the most popular method, group ICA extends the single subject ICA algorithm to the multi-subject setting by estimating a set of group components [16]. Based on multi-subject ICA, dual regression has been proposed to identify between-subject differences in resting state brain connectivity networks [10].

Bayesian model selection [136] handles inter-subject variability and error rate control. However, its current proposed implementation does not scale well, making it more suitable for confirmatory, rather than an exploratory research. Varoquaux et al. [149] propose a data driven method to estimate large-scale brain connectivity using Gaussian modeling and deals with the variability between subjects by using optimal regularization schemes. Ramsey et al. [119] describe and evaluate a combination of a multi-subject search algorithm and the orientation algorithm for group level brain connectivity inference. Extending the structural equation modeling approach, Group iterative multiple model estimation (GIMME) adopts a forward and then backward search algorithm to eliminate the insignificant connections, and thus estimates the brain connectivity networks [50].

Group studies are often closely linked to exploratory analysis. In contrast to confirmatory studies that usually involve verification of just a few pre-selected models, exploratory studies must search through a huge number of possible models to find one or a few models that are best supported by data. Thus, efficiency of the search strategy becomes important, especially since the number of possible network structures increases super-exponentially to the number of brain ROIs involved. For example, with just seven ROIs, there are more than a billion possible network structures.

Accuracy is another important criterion for an exploratory method. In biological scenarios, the goal is not just to adequately model the overall multivariate time series derived from multiple ROIs, but also that the structure of the model, from which biological interpretations are made, is accurately depicted. Colloquially, accuracy can be inferred from answering the questions, “How many of the connections in the model inferred from data are actually true”, “how many true
connections can be detected by the model?” and “how many non-existing connections in the model are falsely reported?” Therefore, error control is a crucial point in the design of reliable methods for discovering group level connectivity.

A theoretically elegant and feasible method for the group-level exploratory analysis of brain connectivity must include both efficiency and accuracy of the learned networks.

Moreover, current group analyses attempt to summarize information into one group by treating all the subjects equally. However, in the neuroimaging studies with the subjects in disease states, simply learning the common group structure may neglect the heterogeneity in the population. Taken the similarity as well as variability into consideration, the subjects could be categorized into a hierarchical structure according to their clinical information and the group connectivity modeling should be able to characterize the diversity among subjects in a hierarchical manner.

Incorporating a priori knowledge into brain connectivity modeling allows for more flexibility in estimation the connectivities at the group level, benefits the biological interpretations and potentially leads to greater sensitivity in accurately discovering the true brain connectivity. Since the neuroscientific interpretation is largely based on the pattern of inferred connections, compared with barely creating a model that fits the overall data well, imposing the prior knowledge into the connectivity models benefit the final biological interpretations. Incorporating a priori knowledge into connectivity modeling also provides a way for the combinational modality studies.

1.2.2 Time Varying Brain Connectivity Network Modeling

The dynamics of brain connectivity networks are particularly important as they are associated with a variety of neurodisorders such as Schizophrenia [124], multiple sclerosis [80], PD [90] and post-traumatic stress disorder [85]. For instance, altered contributions of brain connectivity dynamic patterns have been reported in subjects with multiple sclerosis [80]. The network variations of subjects with Parkinson’s Disease is decreasing compared with that of control subjects, an observation that may be related with the cognitive rigidity, i.e. difficulty between switching tasks,
that is frequently observed in PD [90].

A few studies have investigated brain connectivity dynamics [7, 42, 60, 77, 124] and have demonstrated that connectivity can be mediated by learning and/or task performance [7, 42, 124]. In addition to task-related connectivity changes, various groups have assessed dynamic changes during resting state fMRI.

Frequently, the assumption is made that the brain networks change slowly and smoothly with time, and a sliding window based approach is used. By specifying a fixed window length and shifting the window by a given number of sample points, different network learning methods, such as correlation [61, 67, 140], covariance [1, 49] and ICA [72, 123], have been applied to estimate time-dependent interactions at each time window.

A key problem with sliding window approaches is that it is critical and difficult to determine the appropriate window length. Too long a window will reduce temporal variability and miss possibly important brain state changes, while too short a window may suffer from large fluctuations as the small number of samples results in unstable estimates [79]. Some features, such as patterns of co-varying connections, may be largely invariant to different window lengths [80]. To circumvent the issue associated with window length, time-frequency based approaches, such as wavelet transform based coherence analysis, have been proposed to estimate dynamic brain connectivity in the time-frequency plane [21].

An alternative to assuming smoothly changing connectivity is that the brain states are relatively stable between a few critical time points [92]. These critical time points thus can be used to segment the entire brain signals into quasistationary sections for the purposes of brain connectivity estimation [37, 86, 92]. Nevertheless, change point detection models can be particularly sensitive to artifacts contained in the data.

Lagged interaction based approaches, such as multivariate auto-regressive models [54] have also been employed to study the brain dynamics which examine brain interactions simultaneously and over adjacent time steps. However, since the lagged interactions themselves are assumed to be time-invariant, they technically can be categorized as stationary models. State space model based approaches, by combining lagged interaction and filtering theory, estimate non-stationary brain connectivity at each time point [74].
Although more and more studies have been conducted to investigate time-
dependent brain networks, there are still some challenges. For instance, there is
no consensus on the underlying brain connectivity patterns. Whether the brain
coupling changes smoothly or suddenly, is still unclear. What is the time scale of
the brain connectivity networks, is still in debate. Therefore, further investigation
of dynamic brain connectivity is still in demand. They may ultimately provide us
deeper insights into the flexibility and adaption of underlying brain functions.

1.2.3 Brain Connectivity Network Interpretation

Based on the inferred brain connectivity networks, the extraction of the useful in-
formation for evaluating the network properties are of great importance. However,
the interpretations of the estimated brain connectivity networks are challenging due
to their spatial complexity.

Several approaches have been proposed to summarize and represent the in-
ferred networks such as the graph theoretical measures which was originated from
graph theory [15]. With solid theoretical bases, it serves as the feature extraction
tool for better describing our brain at the network level, and interesting results
have been found by adopting graph theoretical analysis. Small-worldness topology
which describes brain networks with high efficiency but low wiring cost has been
identified in the normal brains [6]. It demonstrates that brain is an efficient infor-
mation processing system. Some other graph measures such as motif, clustering
and modularity have also been investigated, and provide significant insights into
the brain organizations [15]. In addition, graph theoretical analysis also provides
insightful information when investigating brain connectivity related impairments
and disorders such as Parkinsons Disease, Schizophrenia and etc. [59].

Time varying connectivity patterns are particularly difficult in interpretation
due to its additional temporal complexity. More measures may be required to rep-
resent the full potential of temporal and spatial information. To facilitate such
interpretations, the eigenconnectivity networks have been proposed based on the
concatenated dynamic connectivity networks using the decomposition approach
[80]. They serve as the representative spatial connectivity patterns, which can then
be compared across groups. The clustering approach as described in [1] is an alter-
native popular method to identify the temporal functional connectivity states based on the estimated dynamic connectivity networks.

To summarize, the brain connectivity network interpretation approaches play vital roles in facilitating the understanding of estimated brain networks. Comprehensively extraction of the useful information from the inference results have great significance of the real biomedical studies. Further exploration of network analysis approaches should be emphasized in order to promote the understanding of inferred brain connectivity.

1.3 Research Objectives and Methodologies

The goal of this thesis is to develop a set of novel brain connectivity modeling approaches which are able to cope with several challenges present in real brain connectivity studies, and hence further investigate the disease induced changes on brain interaction patterns.

Brain connectivity has not only provided great potential in understanding the brain functioning in normal state, but also extended the insights into related neurological disorders. However, due to the lack of underlying truth of brain interaction mechanisms and complexity of brain itself, modeling the brain connectivity is a difficult topic with some challenges present including the group level inference, hard biological interpretation, poor signal to noise ratio, temporal dynamics, accuracy control, prior knowledge incorporation, multimodality, efficiency and so on.

Motivated by our particular applications, we are interested in addressing the following concerns and challenges. First, group level inference is one of the most important issues as biomedical studies typically involve a group of subjects in order to make inference about a population, rather than focusing on a single subject. However, it’s challenging to deal with the large inter-subject variability while making the inference with sufficient estimation accuracy. Error rate control thus is another important criterion for an exploratory group level analysis method.

To facilitate the accuracy, efficiency and results interpretation of group network modeling, the prior knowledge of the brain connectivity could be taken into consideration. In addition, the prior knowledge coming from other modalities may further assist the group level brain connectivity estimation which is the data fusion
approach for the multimodality studies.

Another challenge in brain network modeling is the time-dependent brain connectivity estimation. As our brain is inherently non-stationary, the dynamics of brain connectivity may thus provide us insights into adaption and reorganization of brain connections. However, most existing approaches using fMRI signals are based on the stationary assumption which may neglect the dynamics of brain interactions. The time-dependent brain connectivity estimation approaches are thus required to fill the gap.

**Figure 1.5:** The overview of challenges and objectives of this thesis work.

According to the aforementioned discussions, the main technical challenges could be summarized as generality, accuracy, efficiency, prior knowledge incorporation, multimodality and temporal dynamics modeling. The objectives of this thesis will focus on developing a set of brain connectivity modeling approaches that are able to deal with aforementioned challenges, including group pattern extraction from a population, false discovery rate control, incorporation of prior knowledge, and dynamic brain connectivity estimation. Specifically, the main research contributions of this thesis are:

- Propose a group level, error rate controlled, prior knowledge incorporated
network modeling approach by extending the original PC<sub>fdr</sub> algorithm.

- Propose a framework for genetically-informed group brain connectivity modeling.

- Develop a dynamic connectivity estimation approach which is able to recover the smoothly changing coefficients as well as suddenly changing connectivity structures.

- Propose a combined temporal network modeling approach for simultaneously static and dynamic brain connectivity estimation.

Fig. 1.5 illustrates the challenges and objectives of this thesis work.

1.4 Thesis Outline

The outline of the remainder thesis is as follows:

In Chapter 2, we extend the original PC<sub>fdr</sub> algorithm and propose a multi-subject, exploratory brain connectivity modeling approach that allows incorporation of prior knowledge of connectivity and determination of the dominant brain connectivity patterns among a group of subjects. The proposed approach is applied to real fMRI data derived from subjects with Parkinson’s Disease on and off L-dopa medication and normal controls performing a motor task, and we find robust group evidence of disease-related changes, compensatory changes and the normalizing effect of L-dopa medication.

In Chapter 3, we propose a framework for genetically-informed group brain connectivity modeling. The proposed method is able to model the diversity of brain connectivity in a population. Subjects are first stratified according to their genotypes, and then a group regularized regression model is employed for brain connectivity modeling. The simulations have been performed to test the performance. The proposed method has then applied to resting state fMRI data from Schizophrenia and normal control subjects, and interesting results have been found. It represents a multi-modal analysis approach for incorporating genotypic variability into brain connectivity analysis directly.

Chapter 4 presents a time varying model to investigate the temporal dynamics of brain connectivity networks. The proposed method allows for estimation of
abrupt changes in network structure via a fused LASSO scheme, as well as recovery of time-varying networks with smoothly changing coefficients via a weighted regression technique. Simulations demonstrate that the proposed method yields improved accuracy on estimating time-dependent connectivity patterns. The proposed method is then applied to real resting state fMRI data sets from Parkinson’s Disease and control subjects. Significantly different temporal and spatial patterns are found to be associated with PD.

In Chapter 5, we further leverage assumptions on the brain dynamics, where joint inference of time-invariant connections as well as time-varying coupling patterns is proposed. We employ a multitask learning model followed by a least square approach to precisely estimate the connectivity coefficients. The proposed method is applied to resting state fMRI data from PD and control subjects, and the eigenconnectivity networks are estimated to obtain the representative patterns of both static and dynamic brain connectivity features.

Finally, chapter 6 summarizes the contributions of this thesis, and discusses on the future research directions.
Chapter 2

False Discovery Rate Controlled, Prior Knowledge Incorporated Group Brain Connectivity Modeling

2.1 Introduction

Graphical models, when applied to functional neuroimaging data, represent brain Regions of Interest (ROIs) as nodes, and the stochastic interactions between ROIs as edges. However, in most non-brain imaging graphical model applications, the primary goal is to create a model that fits the overall multivariate data well, not necessarily accurately reflect the particular connections between nodes. Yet in the applications of graphical models to brain connectivity, the neuro-scientific interpretation is largely based on the pattern of connections inferred by the model. This places a premium on accurately determining the “inner workings” of the model such as accounting for the error rate of the edges in the model.

False discovery rate (FDR) [11, 137], defined as the expected ratio of spurious connections to all learned connections, has been suggested as a suitable error-rate control criterion when inferring brain connectivity. Compared with traditional type
I and type II statistical error rates, the FDR is more informative in bioinformatics and neuroimaging, since it is directly related with the uncertainty of the reported positive results. When selecting candidate genes for genetic research, for example, researchers may want 70% of selected genes to be truly associated with the disease, that is, an FDR of 30%.

Naively controlling traditional type I and type II error rates at specified levels may not necessarily result in reasonable FDR rates, especially in the case of large, sparse networks. For example, consider an undirected network with 40 nodes, with each node interacting, on average, with 3 other nodes, i.e. there are 60 edges in the network. An algorithm with the realized type I error rate of 5% and the realized power of 90% (i.e. the realized type II error rate = 10%) will recover a network with $60 \times 90\% = 54$ correct connections and $[40 \times (40 - 1)/2 - 60] \times 5\% = 36$ false connections, which means that $36/(36 + 54) = 40\%$ of the claimed connections actually would not exist in the true network! This example, while relatively trivial, demonstrates that the FDR may not be kept suitably low by simply controlling traditional type I and type II error rates.

Recent work in the machine learning field has started to investigate controlling the FDR in network structures using a generic Bayesian approach and classical FDR assessment [87]. This work was subsequently extended to look specifically at graphical models where the FDR was assessed locally at each node [143].

Li and Wang proposed a network learning method that allows asymptotically control of the FDR globally. They based their approach on the PC algorithm (named after Peter Spirtes and Clark Glymour), a computationally efficient and asymptotically reliable Bayesian network-learning algorithm. The PC algorithm assesses the (non)existence of an edge in a graph by determining the conditional dependence/independence relationships between nodes [134]. However, different from the original PC algorithm, which controls the type I error rate individually for each edge during conditional independence testing, the Li and Wang algorithm, referred as the $\text{PC}_{\text{fdr}}$ algorithm, is capable of asymptotically controlling the FDR under pre-specified levels [82]. The $\text{PC}_{\text{fdr}}$ algorithm does this by interpreting the learning of a network as testing the existence of edges, and thus the FDR control of edges becomes a multiple-testing problem, which has a strong theoretical basis and has been extensively studied by statisticians [82].

20
In this chapter, we introduce the extension to the original PC_{fdr} algorithm and propose a multi-subject brain connectivity modeling approach, allowing it to incorporate prior knowledge and extending it to robustly assess the dominant brain connectivity in a group of subjects. The major distinguishing feature of the proposed approach compared to these aforementioned approaches is that the current data-driven approach aims at controlling the FDR directly at the group level network. When applied the proposed approach to fMRI data derived from ten subjects with Parkinson’s disease on and off L-dopa medication and ten normal controls performing a motor task, we found robust group evidence of disease-related changes, compensatory changes and the normalizing effect of L-dopa medication.

2.2 Methods

The initial version of Li’s method, called the PC_{fdr} algorithm, was proved to be capable of asymptotically controlling the FDR. It does this by interpreting the learning of a network as testing the existence of edges, and thus the FDR control of edges becomes a multiple-testing problem, which has a strong theoretical basis and has been extensively studied by statisticians [82].

Here we present an extension of the PC_{fdr} algorithm that can incorporate a priori knowledge, which was not specified in the original PC_{fdr} algorithm. We name the extension as the PC_{fdr}^{+} algorithm where the superscript “+” indicates that it is an extension. It handles prior knowledge with two inputs: $E_{\text{must}}$, the set of edges assumed to appear in the true graph, and $E_{\text{test}}$, the set of edges to be tested from the data. The original PC_{fdr} algorithm can thus be regarded as a special case of the extended algorithm, by setting $E_{\text{must}} = \emptyset$ and $E_{\text{test}} = \{\text{all possible edges}\}$. The asymptotic performance of PC_{fdr}^{+} algorithm can refer to [88].

Another extension is the group-level inference. Assessing group level activity is done by considering a mixed-effect model (Step 7 of Algorithm 1). When also incorporating a priori knowledge, the resulting algorithm is named the gPC_{fdr}^{+} algorithm.

Suppose we have $m$ subjects within a group. Then for subject $i$, the conditional independence between the activities of two brain regions $a$ and $b$ given other regions $C$ can be measured by the partial correlation coefficient between $X_a(i)$ and
\( X_b(i) \) given \( X_C(i) \), denoted as \( r_{ab|C}(i) \). Here \( X_\bullet \) denotes variables associated with a vertex or a vertex set, and index \( i \) indicates that these variables are for subject \( i \). By definition, the partial correlation coefficient \( r_{ab|C}(i) \) is the correlation coefficient between the residuals of projecting \( X_a(i) \) and \( X_b(i) \) onto \( X_C(i) \), and can be estimated by the sample correlation coefficient as

\[
\hat{r}_{ab|C}(i) = \frac{\text{Cov}[Y_{a|C}(i), Y_{b|C}(i)]}{\sqrt{\text{Var}[Y_{a|C}(i)]\text{Var}[Y_{b|C}(i)]}}, \tag{2.1}
\]

where

\[
\beta_{a|C}(i) = \arg\min_\beta |X_a(i) - X_C(i)\beta|^2, \tag{2.2}
\]

\[
\beta_{b|C}(i) = \arg\min_\beta |X_b(i) - X_C(i)\beta|^2, \tag{2.3}
\]

\[
Y_{a|C}(i) = X_a(i) - X_C(i)\beta_{a|C}(i), \tag{2.4}
\]

\[
Y_{b|C}(i) = X_b(i) - X_C(i)\beta_{b|C}(i). \tag{2.5}
\]

For clarity, in the following discussion we omit the subscript “\( ab|C \)”, and simply use index “\( i \)” to emphasize that a variable is associated with subject \( i \).

To study the group-level conditional-independence relationships, a group-level model should be introduced for \( r_i \). Since partial correlation coefficients are bounded and their sample distributions are not Gaussian, we apply Fisher’s z-transformation to convert (estimated) partial correlation coefficients \( r \) to a Gaussian-like distributed z-statistic \( z = Z(r) \).

The group model we employ is

\[
Z_i = z_g + e_i, \tag{2.6}
\]

where \( e_i \) follows a Gaussian distribution \( N(0, \sigma_g^2) \) with zero mean and \( \sigma_g^2 \) variance. Consequently, the group-level testing of conditional independence is to used to test the null hypothesis \( z_g = 0 \).

Because \( z_i \) is unknown and can only be estimated, the inference of \( z_g \) should be conducted with \( \hat{Z}_i = Z(\hat{r}_i) \). If \( X_a(i), X_b(i) \) and \( X_C(i) \) jointly follow a multivariate Gaussian distribution, then \( \hat{Z}_i \) asymptotically follows a Gaussian distribution.
\( N(z_i, \sigma^2_i) \) with \( \sigma^2_i = 1/(N_i - p - 3) \), where \( N_i \) is the sample size of subject \( i \)’s data and \( p \) represents the number of variables in \( X_C(i) \). Therefore, based on Eq. (2.6), we have

\[
\hat{z}_i = z_g + e_i + \varepsilon_i,
\]

(2.7)

where \( \varepsilon_i \) follows \( N(0, \sigma^2_i) \), and \( e_i \) follows \( N(0, \sigma^2_g) \). This is a mixed-effect model where \( \varepsilon_i \) denotes the within-subject randomness and \( e_i \) denotes the inter-subject variability. At the group level, \( \hat{z}_i \) follows a Gaussian distribution \( N(z_g, \sigma^2_g + \sigma^2_i) \).

Note that unlike regular mixed-effect models, the within-subject variance \( \sigma^2_i \) in this model is known, because \( N_i \) and \( p \) are known given the data \( X(i) \) and \( C \). In general, \( \sigma^2_i = 1/(N_i - p - 3) \) is not necessarily equal to \( \sigma^2_j \) for \( i \neq j \), and the inference of \( z_g \) should be conducted in the manner of mixed-models, such as estimating \( \sigma^2_g \) with the Restricted Maximum Likelihood (ReML) approach. However, if the sample size of each subject’s data is the same, then \( \sigma^2_i \) equals \( \sigma^2_j \). For this balanced case, which is typically true in fMRI applications and as well the case in this study, we can simply apply a t-test to \( \hat{z}_i \)’s to test the null hypothesis \( z_g = 0 \).

Replacing Step 7 of the single-subject PCfdr algorithm (i.e. the within-subject hypothesis test) with the test of \( z_g = 0 \), we can extend the single-subject version of the algorithm to its group-level version.

Combined with two extensions, the \( gPC^{+}\text{fdr} \) algorithm is described in Algorithm.

---

**Algorithm 1** The \( gPC^{+}\text{fdr} \) algorithm.

**Input:** the multisubject data \( D \), the undirected edges \( E_{\text{must}} \) that are assumed to exist in the true undirected graph \( G_{\text{true}} \) according to prior knowledge, the undirected edges \( E_{\text{test}} \) \((E_{\text{must}} \cap E_{\text{test}} = \emptyset)\) to be tested from the data \( D \), and the FDR level \( q \) for making inference about \( E_{\text{test}} \).

**Output:** the recovered undirected graph \( G_{\text{stop}} \).

**Notations:** \( a, b \) denote the vertices. \( E, C \) denote the vertex set. \( a \sim b \) denotes an undirected edge. \( \text{adj}(a, G) \) denotes vertices adjacent to \( a \) in graph \( G \). \( a \perp b|C \) denotes the conditional independence between \( a \) and \( b \) given \( C \).

1: Form an undirected graph \( G \) from \( E_{\text{test}} \cup E_{\text{must}} \).
2: Initialize the maximum $p$-values associated with the edges in $E_{\text{test}}$ as $P^{\text{max}}_{a \sim b} = \{-1| a \sim b \in E_{\text{test}}\}$.
3: Let depth $d = 0$.
4: repeat
5:   for each ordered pair of vertices $a$ and $b$ that $a \sim b \in E \cap E_{\text{test}}$ and $|\text{adj}(a, G) \setminus \{b\}| \geq d$ do
6:     for each subset $C \subseteq \text{adj}(a, G) \setminus \{b\}$ and $|C| = d$ do
7:       Test hypothesis $a \perp b|C$ for each subject and calculate the $p$-value $p_{a \perp b|C}$ at the group level.
8:       if $p_{a \perp b|C} > P^{\text{max}}_{a \sim b}$, then
9:         Let $P^{\text{max}}_{a \sim b} = p_{a \perp b|C}$.
10:        if every element of $P^{\text{max}}$ has been assigned a valid $p$-value by step 9, then
11:           Run the FDR procedure, Algorithm 2, with $P^{\text{max}}$ and $q$ as the input.
12:           if the non-existence of certain edges are accepted, then
13:              Remove these edges from $G$.
14:           Update $G$ and $E$.
15:           if $a \sim b$ is removed, then
16:              break the for loop at line 6.
17:           end if
18:       end if
19:   end if
20: end for
21: end for
22: Let $d = d + 1$.
23: until $|\text{adj}(a, G) \setminus \{b\}| < d$ for every ordered pair of vertices $a$ and $b$ that $a \sim b$ is in $E \cap E_{\text{test}}$. 

24
Algorithm 2 FDR-stepup [12]

Input: a set of \( p \)-values \( \{p_i| i = 1, \ldots, H\} \), and the threshold of the FDR \( q \)
Output: the set of rejected null hypotheses

1: Sort the \( p \)-values of \( H \) hypothesis tests in the ascendant order as \( p(1) \leq \ldots \leq p(H) \).
2: Let \( i = H \), and \( H^* = H \) (or \( H^* = H(1 + 1/2, \ldots, +1/H) \), depending on the assumption of the dependency among the test statistics).
3: while \[ \frac{H^*}{i} p(i) > q \text{ and } i > 0, \] (2.8)
   do
4: Let \( i = i - 1 \).
5: end while
6: Reject the null hypotheses associated with \( p(1), \ldots, p(i) \), and accept the null hypotheses associated with \( p(i+1), \ldots, p(H) \).

2.3 Real Application

In order to assess the real-world application performance of the proposed method, we apply the \( \text{gPC}^{+}_{fdr} \) algorithm for inferring group brain connectivity network to fMRI data collected from twenty subjects. All experiments were approved by the University of British Columbia Ethics Committee. Ten normal people and ten Parkinson’s disease (PD) patients participated in the study. During the fMRI experiment, each subject was instructed to squeeze a bulb in their right hand to control an “inflatable” ring so that it smoothly passed through a vertically scrolling a tunnel. The normal controls performed only one trial, while the PD subjects performed twice, once before L-dopa medication, and the other approximately an hour later, after taking medication. Because L-dopa has dramatic clinical effects on motor performance, we expected that the extended \( \text{PC}_{fdr} \) algorithm could detect the normalizing effect of L-dopa over brain connectivity.

Three groups were categorized: group \( N \) for the normal controls, group \( P_{pre} \) for the PD patients before medication, and group \( P_{post} \) for the PD patients after taking L-dopa medication. For each subject, 100 observations were used in the network modeling. For details of the data acquisition and preprocessing, please
Table 2.1: The name and abbreviation of 12 selected brain ROIs.

<table>
<thead>
<tr>
<th>Full Name of Brain Region</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left/Right lateral cerebellar hemispheres</td>
<td>lCER, rCER</td>
</tr>
<tr>
<td>Left/Right globus pallidus</td>
<td>lGLP, rGLP</td>
</tr>
<tr>
<td>Left/Right putamen</td>
<td>lPUT, rPUT</td>
</tr>
<tr>
<td>Left/Right supplementary motor cortex</td>
<td>lSMA, rSMA</td>
</tr>
<tr>
<td>Left/Right thalamus</td>
<td>lTHA, rTHA</td>
</tr>
<tr>
<td>Left/Right primary motor cortex</td>
<td>lM1, rM1</td>
</tr>
</tbody>
</table>

“l” or “r” in the abbreviations stand for “Left” or “Right”, respectively.

Refer to [111]. 12 anatomically-defined Regions of Interest (ROIs) were chosen based on prior knowledge of the brain regions associated with motor performance (Table 2.1).

We utilized the two extensions of the PC_{fdr} algorithm and learned the structures of first-order group dynamic Bayesian networks from fMRI data. Because the fMRI BOLD signal can be considered as the convolution of underlying neural activity with a hemodynamic response function, we assumed that there must be a connection from each region at time $t$ to its mirror at time $t+1$. We also assumed that there must be a connection between each region and its homologous region in the contralateral hemisphere. The TR-interval (i.e., sampling period) was a relatively long 1.985 seconds; we restricted ourselves to learn only connections between ROIs without time lags. In total, there are $12 + 6 = 18$ pre-defined connections, and $12 \times (12 - 1) \div 2 - 6 = 60$ candidate connections to be tested. The brain connectivity networks (with the target FDR of 5%) learned for the normal (group $N$) and PD groups before (group $P_{pre}$) and after (group $P_{post}$) medication are compared in Figure 2.1. Note the connection between the cerebellar hemisphere and contralateral thalamus in the normal subjects, and between the Supplementary Motor Area (SMA) and the contralateral putamen, consistent with prior knowledge. Interestingly, in $P_{pre}$ subjects, the left cerebellum now connects with the Right SMA, and the Right SMA $\leftarrow$left putamen connection is lost. Also, there are now bilateral primary motor cortex (M1) $\leftarrow$putamen connections seen in the $P_{pre}$ group, presumably as a compensatory mechanism. After medication ($P_{post}$), the Left SMA $\leftarrow$Left thalamus connection is restored back to be normal.
2.4 Conclusion and Discussion

Graphical models to infer brain connectivity from fMRI data have relied on the assumption that if a model accurately represented the overall activity in several ROIs, the internal connections of such a model would accurately reflect underlying brain connectivity. The PC_{fdr} algorithm was designed to loosen this overly-restrictive assumption, and asymptotically control the FDR of network connections inferred from data.

In this chapter, we first presented the PC_{fdr}^{+} algorithm, an extension of the PC_{fdr} algorithm, that allows for incorporation of prior knowledge of network structure into the learning process, greatly enhancing its flexibility in practice. The PC_{fdr}^{+} algorithm handles prior knowledge with two inputs: $E_{must}$, i.e. the set of edges that is assumed to appear in the true graph, and $E_{test}$, the set of edges that are to be tested from the observed data. Another extension to PC_{fdr} algorithm we described here was the ability to infer brain connectivity patterns at the group-level, with inter-subject variance explicitly taken into consideration. Combined with two extensions, the proposed gPC_{fdr}^{+} algorithm is able to make inference at the group level by incorporating the prior knowledge and controlling the error rate.

When applying the proposed gPC_{fdr}^{+} to fMRI data collected from PD subjects performing a motor tracking task, we found group evidence of disease changes (e.g. loss of Left cerebellar $\leftarrow$ SMA connectivity), compensatory changes in PD (e.g. bilateral M1 $\leftarrow$ contralateral putamen connectivity) and evidence of restoration of connectivity after medication (Left SMA $\leftarrow$ Left thalamus). The tremendous variability in clinical progression of PD is likely due to variability not only in disease rate progression, but also in variability in the magnitude of compensatory changes. This highlights the importance of the proposed method, as it allows robust estimation of disease effects, compensatory effects and effects of medication, all with a reasonable sample size, despite the enhanced intersubject variability seen in PD.
Figure 2.1: (a) Learned brain connectivity for the normal group (group N). (b) Learned brain connectivity for the PD group before medication (group $P_{pre}$). (c) Learned brain connectivity for the PD group after medication (group $P_{post}$). Here “L” and “R” refer to the left and right sides respectively. The solid lines are predefined connectivity, and the dashed lines are learned connectivity.
Chapter 3

A Genetically-informed, Group Brain Connectivity Modeling Framework

3.1 Introduction

Schizophrenia, a chronic, disabling mental disorder, has profound health care impact. While the exact pathoetiology of the condition is unknown, genetics plays a vital role, with inheritability estimated to be as high as 80% [122], and environmental and genetic/environmental interactions explaining the remainder [20]. Thus the ability to merge genetic information and their phenotypes, and intermediate endophenotypes based on imaging features, may play a key role in understanding the disease [3].

Exploring the genetic influences on structural and functional brain imaging is important for the investigation of partially-inherited psychiatric diseases including Schizophrenia [8]. However, how to deal with high-volume datasets inherent to both genetic and imaging studies remains a challenge. Early research usually involved candidate gene studies. For example, genetic variations coding for the serine-threonine protein kinase, AKT1, were shown to alter human prefrontal-striatal structure and function [139]. This has rapidly been extended to multivariate
approaches allowing for voxel-wise/genome-wide association studies [51,65,151]. These have led to e.g., the observation that there is strong genetic influence in default-mode connectivity [53].

Schizophrenia appears to be particularly suited for assessing the associations between genetics and brain connectivity. Neonates at genetic risk for Schizophrenia have reduced overall anatomical connectivity compared to neonates without the same risk [129]. Previous research has suggested that functional brain connectivity appears to be more sensitively altered in Schizophrenia than amplitude-based fMRI univariate voxel-wise methods, such as entropy analysis [8,95]. Additionally, by representing brain activity as a connectivity network, graph theoretical analysis can be applied to study structure and topology properties [15], providing a way to characterize brain activity at the system level.

A number of mathematical models have been introduced for studying brain connectivity [132] including graphical models, such as Dynamic Causal Models [47] and Bayesian Networks [83]. As biomedical experiments typically involve groups of subjects in order to make inferences about a population, several group methods have been proposed, such as Bayesian model selection [136], Group Covariance estimation [149], multi-subject search algorithm [119], and the group $PC_{fdr}$ algorithm [88]. These group analyses attempt to stratify subjects based on e.g., disease states and treat the subjects equally within the group. However, in neuroimaging studies, simply grouping subjects based on disease states alone may neglect heterogeneity in the population, possibly violating statistical assumptions used in subsequent analyses.

In the following proposed approach, we place emphasis on group fMRI connectivity estimation, as we are interested in effectively modelling intra-group diversity while still allowing for differentiation between subjects in a graphical model framework. We propose to first categorize subjects within each group according to ancillary (e.g. genetic) information, allowing the diversity in brain connectivity among subjects to be modeled more accurately, yet still fitting the overall multivariante data well at the group level. This approach is essentially equivalent to incorporating prior knowledge into the brain connectivity analysis.

We propose the use of prior information, in this case, Single nucleotide polymorphisms (SNPS) in the D-amino acid oxidase activator (DAOA) gene, to inform
group brain connectivity networks. In brain connectivity modeling, prior knowledge can be used for such purposes as variable/model selection which can benefit the computational efficiency, and/or specifying which ROIs to use, easing final biological interpretation [47, 54, 132]. Incorporating prior knowledge in modelling is usually done with confirmatory models, and could conceivable have a detrimental effect on models employing a purely exploratory approach [84]. However, incorporating a priori knowledge into a data driven method may actually provide more flexibility [88]. While prior knowledge is a useful tool in the variable/model selection process, information from other modalities, such as genetic information provides complementary information to guide brain connectivity pattern estimation.

Since the neuroscientific interpretation of brain connectivity analysis is largely based on the pattern of inferred connections rather than just ensuring that the model that fits the overall data well, imposing such prior knowledge into a connectivity model will greatly assist the final biological interpretation.

The proposed approach is different from previous approaches for genetically influenced brain interactions. Previous approaches can be considered as separate-model approaches, where a possibly common strategy is used to analyze brain connectivity and genetic features separately, and then univariate/multivariate associations between features are explored. In contrast, our proposed approach is a direct joint-model approach where the known genotypic information is explicitly modeled as prior knowledge in brain connectivity modeling. A penalized group fused regression model is used based on the time courses from a priori specified ROIs. Each ROI time course is in turn predicted from all other ROI time courses at zero lag using a group regression model with a penalty term incorporating genotypic dissimilarity (Fig. 3.1). While we emphasize genetic information as prior knowledge for informing the connectivity model, we note that it is also possible to combine the proposed group modeling approach with other prior information, such as clinical indices.

In the rest of this chapter, we first introduce a group fused brain connectivity modeling approach including the model, objective function, and the implementation method. We name the proposed penalized regression model as Overlapped group fused model (OGFM). We then incorporate simulations to compare the per-
formance of the proposed method with other state-of-the-art modeling approaches. Finally, we apply the proposed approach to a joint resting-state fMRI/genetic data set obtained from eleven subjects with Schizophrenia and nine normal control subjects and demonstrate that brain connectivity patterns are jointly modulated by disease states and genotypes.

3.2 Methods
We first introduce the group modeling approach by incorporating genetic information as a priori knowledge. Asymptotic analysis of large samples, given in the
appendix, is further used to justify the proposed method.

3.2.1 Overlapped Group Fused Model

Prior studies have proposed regression approaches to model brain connectivity, such as autoregressive models [144] or the group LASSO (Least Absolute Shrinkage and Selection Operator) model [23]. In the approach used in the group LASSO model, and the one we adopt here, the fMRI time course of an ROI is regarded as the response variable, and is predicted from the time courses of all other ROIs at zero-lag as,

\[ Y^j = X^{-j} \beta + e, \]  

(3.1)

where \( Y^j \) is the time course of the \( j^{th} \) ROI, \( X^{-j} \) is the predictor matrix based on the timecourse of all other ROIs except the \( j^{th} \), \( \beta \) is the coefficient vector and \( e \) is the Gaussian noise term. Note that since both \( Y^j \) and \( X^{-j} \) have similar autocorrelation properties, it is not necessary to whiten the regressors to ensure iid Gaussian residuals. Also based on previous studies, connections between brain regions can be considered as a sparse network [23]. One computationally efficient approach to promote sparsity in the coefficient vector is to use an \( l_1 \) penalty on the regression coefficients – i.e. the regularized LASSO method [141]. Various extensions of \( l_1 \) penalized regression methods have been proposed, for example, the group robust LASSO method alluded to earlier [23], the two-graph guided multi-task LASSO for eQTL mapping [26], and the variable selection model with elastic net penalty [161].

In this study, we introduce an overlapped group fused model (OGFM) which is able to control the variability in the group modeling by adopting a fusion penalty into the LASSO model at the group level. In the following discussion, for simplicity, we will represent \( Y^j \) as \( Y \) and \( X^{-j} \) as \( X \). Considering a group with \( S \) subjects, for each subject \( i \in \{1, 2, \cdots, S\} \), \( Y(i) \) and \( X(i) \) are the response vector and predictor matrix, where dimension of \( Y(i) \) is \( N \times 1 \) and dimension of \( X(i) \) is \( N \times K \) with \( N \) being the number of time points and \( K \) representing the number of features (= number of ROIs - 1). Therefore, in the group design as shown in Equ. 3.2, \( X_{tot} \) and \( Y_{tot} \) is composed of \( S \) blocks and coefficient \( B_{tot} = diag(B(1), B(2), \cdots, B(S)) \) is a
block matrix of dimension $SK \times S$. $B(i)$ of $B$ represents the coefficients between $Y(i)$ and $X(i)$ for subject $i \in \{1, 2, \cdots, S\}$.

\[
Y_{tot} = (Y(1), Y(2), \cdots, Y(S))
= (X(1), X(2), \cdots, X(S)) \times 
\text{diag}(B(1), B(2), \cdots, B(S)) + E
= X_{tot}B + E.
\] (3.2)

In this model, the off-diagonal elements in $B_{tot}$ are all zeros and we only need to infer the diagonal blocks $B(1), \cdots, B(S)$. To promote a fusion effect across the group structures, the grouping relationships are applied to control the structural similarities within the group. Suppose $G = (V, E)$ is a grouping relationship graph applied on the subjects, where $V$ is a set of vertices representing the subjects and $E$ is a set of edges representing the relationships between subjects. Then the objective function of the overlapped group fused model that we need to minimize is defined as:

\[
f(B) = \sum_{i=1}^{S} ||Y(i) - X(i)B(i)||_F^2 + \lambda \sum_{i=1}^{S} ||B(i)||_{l_1} + \\
\gamma \sum_{e_{lm} \in E} W(e_{lm}) ||B(l) - B(m)||_{l_1}
= \sum_{i=1}^{S} \sum_{j=1}^{N} (Y_j(i) - \sum_{k=1}^{K} X_{jk}(i)B_k(i))^2 + \lambda \sum_{i=1}^{S} \sum_{k=1}^{K} |B_k(i)|
+ \gamma \sum_{e_{lm} \in E} W(e_{lm}) \sum_{k=1}^{K} |B_k(l) - B_k(m)|
\] (3.3)

where $Y(i)$ and $X(i)$ are the response vector and predictor matrix for subject $i \in \{1, 2, \cdots, S\}$ and $Y_j(i)$ is the $j^{th}$ sample in $Y(i)$. $X_{jk}(i)$ and $X_{jk}(i)$ stands for the $j^{th}$ sample row and $k^{th}$ feature column in the predictor matrix $X(i)$. $B(i)$ is the coefficient vector between $Y(i)$ and $X(i)$, and $B_k(i)$ represents the $k^{th}$ coefficient corresponding to $X_{jk}(i)$. $W(e_{lm})$ is a weight assigned to edge $e_{lm} \in E$ to control the grouping effect between subject $l$ and $m$. Here we set the weight $W(e_{lm})$ of edge $e_{lm}$ as the square value of similarity score which is calculated based on the prior
information (the genetic variations in this study). With this formulation, the higher grouping score leads to less discrepancy in models between subjects \(l\) and \(m\). The first \(l_1\) penalty is designed to control the sparsity on the learned coefficients, and the second penalty is designed to control the sparsity on subject differences. The optimization problem in Equ. 3.3 can be efficiently solved by a coordinate-descent algorithm [26, 75]. Optimization of the non-differentiable objective function in Equ. 3.3 can be transferred to an iterative minimization of its surrogate differentiable function (Equ. 3.4):

\[
\begin{align*}
\text{minimize} \quad & \sum_{i=1}^{S} \sum_{j=1}^{N} (Y_j(i) - \sum_{k=1}^{K} X_{jk}(i) B_k(i))^2 \\
& + \lambda \sum_{i=1}^{S} \sum_{k=1}^{K} (B_k(i))^2 / D_{ki} \\
& + \gamma \sum_{e_{lm}\in E} W^2(e_{lm}) \sum_{k=1}^{K} (B_k(l) - B_k(m))^2 / C_{klm} \\
\text{subject to} \quad & \sum_{k,i} D_{ki} = 1, \sum_{k,l,m} C_{klm} = 1, \text{and} \ D_{ki}, C_{klm} \geq 0
\end{align*}
\] (3.4)

We can optimize Equ. 3.4 by its Lagrangian form and obtain the solutions as,

\[
D_{ki} = \frac{|B_k(i)|}{\sum_{k'=1}^{K} \sum_{i'=1}^{S} |B_{k'}(i')|} \quad \text{(3.5)}
\]

\[
C_{klm} = \frac{W(e_{lm}) |B_k(l) - B_k(m)|}{\sum_{k'=1}^{K} \sum_{\ell',m'} \in E W(e_{\ell',m'}) |B_{k'}(\ell') - B_{k'}(m')|} \quad \text{(3.6)}
\]
\[ B_k(i) = \left( \sum_{j=1}^{N} X_{jk}(i) R_{jk}(i) + \gamma \sum_{e_{il} \in E} W^2(e_{il}) B_k(l) / C_{kil} + \gamma \sum_{e_{im} \in E} W^2(e_{im}) B_k(m) / C_{kmi} \right) / \left( \sum_{j=1}^{N} (X_{jk}(i))^2 + \lambda / D_{ki} \right) + \gamma \sum_{e_{il} \in E} W^2(e_{il}) / C_{kil} + \gamma \sum_{e_{im} \in E} W^2(e_{im}) / C_{kmi} \right) \]

where \( R_{jk}(i) = [Y_j(i) - \sum_{k' \neq k} X_{jk'}(i) B_{k'}(i)] \).

Here \( k' \) represents the feature index and \( j' \) represents the sample index. The least square solution between \( Y(i) \) and \( X(i) \) is used as initial value for \( B(i) \) of \( B \). From the initial value \( B \), we iteratively update \( D \), \( C \) and \( B \) until \( B \) converges. The summary of the algorithm implementation is shown in Table 3.1. Note that when \( \lambda \) is zero, it becomes regression with the group fused penalty only. The asymptotic properties of this estimate are given in the appendix.

### 3.2.2 Model Selection and Degree of Freedom

A key issue of our OGFM is to determine the tuning parameters \( \lambda \) and \( \gamma \), which will affect the subsequent model selection and parameter estimation. We propose to select the optimal \( \lambda \) and \( \gamma \) according to the Bayesian information criterion (BIC), which works as follows. We first specify a grid of the tuning parameter pairs \((\lambda, \gamma)\). For each pair of \((\lambda, \gamma)\), we calculate the Degree of freedom \( (DF) \) of the proposed OGFM estimate \( \hat{d}(\lambda, \gamma) \) defined in Equ. 3.12. In essence, the degree of freedom quantifies the number of free parameters in the estimated OGFM (Table 3.1). Then, the estimated \( \hat{d}(\lambda, \gamma) \) is used for calculating the BIC,

\[ \text{BIC}(\lambda, \gamma) = N \cdot \ln(\hat{\sigma}_e^2) + \hat{d}(\lambda, \gamma) \cdot \ln(N), \]

where \( \hat{\sigma}_e^2 \) is the estimated variance of the residuals. Finally, the optimal tuning parameters are defined as,
### Table 3.1: Implementation of overlapped group fused model.

**Input**

The number of subjects \( S \), \( \{X(i)\}_{i=1,2,\ldots,S} \in \mathbb{R}^{N \times K} \), \( \{Y(i)\}_{i=1,2,\ldots,S} \in \mathbb{R}^{N} \), graph structure \( G \), the corresponding weight matrix \( W \), a pair of tuning parameter \((\lambda, \gamma)\) and threshold of convergence \( \theta \).

**Initialization**

Set \( B(i) = \min \|Y(i) - X(i)B(i)\|_2^* \)

**Iterate**

1. Eq. 3.5
   \[
   D_{ki} = \frac{|B_k(i)|}{\sum_{k'=1}^{K} \sum_{i'=1}^{S} |B_{k'}(i')|} \text{ for } k = 1, 2, \ldots, K \text{ and } i = 1, 2, \ldots, S
   \]
2. Eq. 3.6
   \[
   C_{klm} = \frac{W(e_{lm})|B_k(l) - B_k(m)|}{\sum_{k'=1}^{K} \sum_{i'=1}^{S} \sum_{e_{l'm'} \in E} W(e_{l'm'})|B_{k'}(l') - B_{k'}(m')|} \text{ for } k = 1, 2, \ldots, K \text{ and } l, m = 1, 2, \ldots, S
   \]
3. Eq. 3.7
   \[
   B_k(i) = \frac{\sum_{j=1}^{N} X_{jk}(i)R_{jk}(i) + \gamma \sum_{e_{l'm'} \in E} W^2(e_{l'm'})B_k(l)/C_{klm} + \gamma \sum_{e_{l'm'} \in E} W^2(e_{l'm'})B_k(m)/C_{klm}}{\sum_{j=1}^{N} (X_{jk}(i))^2 + \lambda / D_{ki} + \gamma \sum_{e_{l'm'} \in E} W^2(e_{l'm'})/C_{klm}} \text{ for } k = 1, 2, \ldots, K \text{ and } i = 1, 2, \ldots, S
   \]

**Until** convergence of \( B \) (reaches the threshold of convergence \( \theta \)).

**Output**

The coefficient \( B(\lambda, \gamma) = B \)

The least square solution.

\[
(\hat{\lambda}_{opt}, \hat{\gamma}_{opt}) = \arg\min_{\lambda, \gamma} \text{BIC}(\lambda, \gamma). \tag{3.9}
\]

It remains to determine the df for the OGFM. For the regression model, the degree of freedom can be estimated as \([41]\),

\[
d_f = \frac{\sum_{i=1}^{N} \text{cov}(y_i, \hat{y}_i)}{\sigma^2} \tag{3.10}
\]

where \( \hat{y}_i \) is the estimated values of \( y_i \), and \( \sigma^2 \) is the variance of \( y_i \) with \( X \) fixed.

In a general linear model with \( N > K \), the degree of freedom is \( K \). In the LASSO model, the degree of freedom is usually considered as the number of non-zero coefficients and the upper bound is \( \min(K, N) \). In the fused LASSO model \([142]\), the degree of freedom is defined as the number of non-zero coefficient blocks. Following this line, one straightforward way to define the degree of freedom for
Figure 3.2: The comparison of estimated degree of freedom (df) and true degree of freedom. X axis represents the estimated df and y axis represents the true df. The solid line is the least square regression between true df and estimated df, and the dashed line is the 45° degree line.

the proposed model is to count the number of unique non-zero values of coefficients across the subjects for each corresponding feature. For instance, in our study, $B(i)$ is a $K \times 1$ coefficient vector for subject $i \in \{1,2,\cdots,S\}$ and $K$ is the number of features. Then for feature $j \in \{1,2,\cdots,K\}$, each unique non-zero values in $B(j,1),B(j,2),\cdots,B(j,S)$ counts one degree of freedom,

$$df = \sum_{j=1}^{K} \# \{ \text{unique nonzero coefficients in } \ B(j,1),\cdots,B(j,S) \}.$$  \hfill (3.11)

It is clear that the degree of freedom depends on the true coefficient matrix $B$ and it is unknown. However, we can estimate it by the natural plug-in estimate defined as,
Table 3.2: Implementation of model selection.

<table>
<thead>
<tr>
<th>Input</th>
<th>Number of subjects $S$, ${X(i)}_{i=1,2,...,S}$, $Y(i) \in R^N$, graph structure $G$, the corresponding weight matrix $W$, a set of tuning parameter pairs ${\lambda, \gamma}$, and threshold of convergence $\theta$.</th>
</tr>
</thead>
<tbody>
<tr>
<td>for each pair of tuning parameter $(\lambda, \gamma)$</td>
<td>1. Input $S$, ${X(i)}_{i=1,2,...,S}$, $Y(i) \in R^N$, $G$, $W$, $(\lambda, \gamma)$ and $\theta$ into OGMF algorithm (Table 3.1), and get the estimated coefficient $B(\lambda, \gamma)$.</td>
</tr>
<tr>
<td>end for</td>
<td>2. Calculate $BIC(\lambda, \gamma)$ according to Eq. 3.8.</td>
</tr>
<tr>
<td>Select the optimal parameters according to Eq. 3.9,</td>
<td>$(\hat{\lambda}<em>{opt}, \hat{\gamma}</em>{opt}) = \text{argmin}_{\lambda, \gamma} BIC(\lambda, \gamma)$.</td>
</tr>
<tr>
<td>Output the optimal tuning parameter pair $(\hat{\lambda}<em>{opt}, \hat{\gamma}</em>{opt})$</td>
<td></td>
</tr>
</tbody>
</table>

To verify the estimated degree of freedom defined in Eq. 3.12, we compare the estimated and true degree of freedom using the simulated data as described in the Simulation Section. We generate the data set with heterogeneous overlapped group structures. As shown in Fig. 3.2, we note that the estimated values are quite close to the true values. Based on the comparison results, the estimated degree of freedom could be used to represent the approximate degree of freedom in the model selection procedure. The implementation of the model selection is shown in Table 3.2.

3.2.3 Implementation of Brain Connectivity Modeling and Statistical Inference

To make inference of brain connectivity networks using linear regression model as shown in Eq. 3.1, we treat each ROI in turn as a response vector and all other ROIs as predictor matrix. This corresponding coefficient vector would give the
strength of connectivity from all other ROIs to the target ROI. For each ROI, the coefficient vector would be estimated one by one and we could obtain the whole brain networks for all ROIs.

To incorporate the genetic information into brain connectivity estimation, we propose the framework for genetically-informed group brain connectivity modeling as shown in Fig. 3.1. The known genotypic information is incorporated as prior information in brain connectivity modeling. In the first step, the similarities of SNPs between subjects serve to categorize them within the group, and generate the relationship graph. Then the similarity score between any two subjects would be squared as the weight in $W$ as described in the Method section. In the second step, the similarity graph is incorporated into the proposed overlapped group fused model. As discussed before, the coefficient of connectivity for each ROI is estimated sequentially for all subjects to finally obtain the whole group brain connectivity network.

To robustly assess the estimated coefficients of brain connectivity networks, we utilize the permutation test, as has been previously suggested for fMRI data [55, 107]. In brief, for the fixed parameters chosen in the optimal model, the temporal order of time courses of each subject is permuted, and then the coefficients are re-estimated using the permuted signals. This is repeated $L$ times and we count the number of times $M$ that the permuted results are larger than the learned coefficients. The significance of the coefficient is then estimated by $p_{val} = M/L$. In our study, we choose the significance level at $p_{val} = 0.05$.

### 3.2.4 Graph Theoretical Analysis

Based on the learned connectivity networks, we will apply graph theoretical analysis to extract the network features. In the brain connectivity networks, the nodes represent brain Regions of Interest (ROIs) and the edges represent the relationship/cooperation between ROIs. Though the brain connectivity networks themselves are of great interest, the graph theoretical measures could help us quantitatively characterize the networks. It has been widely used to explore the structural and functional properties of brain connectivity networks [15].

In this study, we utilize several popular graph measures in terms of density,
global efficiency and node degree [15]. Global efficiency describes communication efficiency of the entire graph which is calculated by the average of the inverse shortest path between any pairs of nodes in the graph. Density is defined as the fraction of present connections to all possible connections. For a local network measure, we will employ node degree, which is the total number of edges connected with a given node.

3.3 Simulations

We performed simulations to compare the performance of the proposed overlapped group fused model with that of original LASSO and group LASSO methods. The multi-subject time courses were generated from a Gaussian model. In each simulation, the number of subjects was chosen as \( S = 12 \) and the number of features was set to \( K = 50 \). The data were generated as follows:

(1). The group structures were assigned for 12 subjects. All the subjects are categorized into several overlapped subgroups with a few shared coefficients across subgroups (Fig. 3.3).

(2). The coefficient vector \( b \) was generated for each subject according to their grouping effects. For ease of interpretation, the vector \( b \) was binary.

(3). The design matrix \( X \) for each subject was randomly generated containing \( N \) observations and \( K \) predictors. The error \( e \) was Gaussian noise \( \sim N(0, 1) \). To generate the relationship graph \( G \), each element \( G_{ij} \) in \( G \) was determined by the structure similarity between subject \( i \) and \( j \) by calculating their structure overlaps. The weight of \( G_{ij} \) was the square value of the corresponding similarity score.

In the simulations, we tested the performances of the algorithms as a function of the number of time points \( N \). For reliable assessment, each procedure was repeated fifty times and we compared the averaged performances of the different algorithms. In the simulation, large variability was considered in the group modeling where the subjects could be categorized into several partially overlapped subgroups, and the type I error rate, detection power, FDR and F1 score were compared at sample sizes of 70 to 300 time points.

FDR is defined as the expected ratio of spurious connections to all the learned connections. FDR can be more reasonable in biomedical studies, since it’s directly
related with the uncertainty of the reported positive results. F1 score was used to evaluate the general performance by considering both the Type I and Type II error rates,

$$F1 = \frac{2 \times \text{True positive}}{2 \times \text{True positive} + \text{False negative} + \text{False positive}}$$  \hspace{1cm} (3.13)

**Figure 3.3:** One example of the heterogeneous overlapped group structure used in the simulation.

Simulation results are shown in Fig. 3.4. In the simulation, the proposed method steadily controlled the type I error rate at < 0.05. The detection power of the proposed method increased with increasing number of sample points. In our proposed overlapped group fused model, the dimension of the group design matrix is $N \times KS$. Considering $KS > N$, the detection power was $< 1$ and improved with $N$ increasing. However, the LASSO method recovered the coefficients of each subject separately; and in the group LASSO model, the dimension of the design matrix is $NS \times KS$ where the number of samples is larger than the number of features. As a result, they can obtain a higher detection power. However, the FDR and F1 scores shown in Fig. 3.4 demonstrated that our proposed model had generally better performance in recovering the heterogeneous group structures.

The results here show that the overlapped group fused model can accurately
Figure 3.4: Simulation results: Simulation is designed to recover a heterogeneous group network based on measures such as Type I error rate, detection power, FDR and F1 score as the number of data points increase. OGFM represents the proposed overlapped group fused model. (a). Comparison of the Type I error rate. (b). Comparison of the detection power. The lines of detection power of Lasso and Group Lasso models are identical. (c). Comparison of the FDR. (d). Comparison of the F1 score.
recover the heterogeneous group structures with relatively high detection power. When applied to the real application, it provides a way to investigate the diversity in the population.

3.4 Real Application in Schizophrenia

3.4.1 Subjects

Eleven subjects with Schizophrenia and nine control subjects were recruited in this study. All the experiments were approved by the local institutional Ethics Committee, and all subjects provided written, informed consent prior to participation. All the patients were diagnosed as having Schizophrenia by a psychiatrist using standard clinical criteria and all control subjects did not have any history of neurological illness and/or brain trauma. A full description of subject demographics can be found in [22].

3.4.2 Selection and Genotyping of SNPs

Strong evidence suggests that the glutamatergic system may contribute to the pathophysiology of Schizophrenia [71]. The DAOA gene is related to the glutamatergic system and our previous study already confirmed its associations with Schizophrenia [22]. Here we investigate the diversity in brain connectivity under the control of the DAOA gene. Genomic DNA was extracted from peripheral blood leukocytes using a standard phenol/chloroform method [22]. Eight SNPs present in the DAOA gene were selected including rs3916966, rs2391191, rs3918342, rs3918341, rs778294, rs9558562, rs3916967, and rs1421292.

We first used the polymorphisms of the DAOA gene to categorize the subjects by calculating their genotype similarity graph $G$. The similarity score for each pair of subjects was determined by their overlapped percentage of all 8 SNPs and the threshold was set to 0.5 where the small similarity score was removed to zero. The similarity graph was then used in the overlapped group fused model to guide the brain connectivity modeling in the normal and patient groups. BIC was applied to choose the optimal model. The permutation test was employed finally to determine the significant brain connections.
Table 3.3: The index and name of 52 selected brain ROIs. ‘L’ represents the brain left side and ‘R’ represents the brain right side.

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Index</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L-ACC</td>
<td>27</td>
<td>R-ACC</td>
</tr>
<tr>
<td>2</td>
<td>L-Amygdala</td>
<td>28</td>
<td>R-Amygdala</td>
</tr>
<tr>
<td>3</td>
<td>L-Calcarine</td>
<td>29</td>
<td>R-Calcarine</td>
</tr>
<tr>
<td>4</td>
<td>L-Caudate</td>
<td>30</td>
<td>R-Caudate</td>
</tr>
<tr>
<td>5</td>
<td>L-Cingulum Mid</td>
<td>31</td>
<td>R-Cingulum Mid</td>
</tr>
<tr>
<td>6</td>
<td>L-Cingulum Post</td>
<td>32</td>
<td>R-Cingulum Post</td>
</tr>
<tr>
<td>7</td>
<td>L-Cuneus</td>
<td>33</td>
<td>R-Cuneus</td>
</tr>
<tr>
<td>8</td>
<td>L-Frontal Mid</td>
<td>34</td>
<td>R-Frontal Mid</td>
</tr>
<tr>
<td>9</td>
<td>L-Hippocampus</td>
<td>35</td>
<td>R-Hippocampus</td>
</tr>
<tr>
<td>10</td>
<td>L-Insula</td>
<td>36</td>
<td>R-Insula</td>
</tr>
<tr>
<td>11</td>
<td>L-Occipital Inf</td>
<td>37</td>
<td>R-Occipital Inf</td>
</tr>
<tr>
<td>12</td>
<td>L-Occipital Mid</td>
<td>38</td>
<td>R-Occipital Mid</td>
</tr>
<tr>
<td>13</td>
<td>L-Occipital Sup</td>
<td>39</td>
<td>R-Occipital Sup</td>
</tr>
<tr>
<td>14</td>
<td>L-Pallidum</td>
<td>40</td>
<td>R-Pallidum</td>
</tr>
<tr>
<td>15</td>
<td>L-Parahippocampus</td>
<td>41</td>
<td>R-Parahippocampus</td>
</tr>
<tr>
<td>16</td>
<td>L-Parietal inf</td>
<td>42</td>
<td>R-Parietal inf</td>
</tr>
<tr>
<td>17</td>
<td>L-Parietal sup</td>
<td>43</td>
<td>R-Parietal sup</td>
</tr>
<tr>
<td>18</td>
<td>L-Putamen</td>
<td>44</td>
<td>R-Putamen</td>
</tr>
<tr>
<td>19</td>
<td>L-Rectus</td>
<td>45</td>
<td>R-Rectus</td>
</tr>
<tr>
<td>20</td>
<td>L-Temporal Inf</td>
<td>46</td>
<td>R-Temporal Inf</td>
</tr>
<tr>
<td>21</td>
<td>L-Temporal Mid</td>
<td>47</td>
<td>R-Temporal Mid</td>
</tr>
<tr>
<td>22</td>
<td>L-Temporal Sup</td>
<td>48</td>
<td>R-Temporal Sup</td>
</tr>
<tr>
<td>23</td>
<td>L-Thalamus</td>
<td>49</td>
<td>R-Thalamus</td>
</tr>
<tr>
<td>24</td>
<td>L-Total Cingulum</td>
<td>50</td>
<td>R-Total Cingulum</td>
</tr>
<tr>
<td>25</td>
<td>L-Total Frontal Inf</td>
<td>51</td>
<td>R-Total Frontal Inf</td>
</tr>
<tr>
<td>26</td>
<td>L-Cerebellum</td>
<td>52</td>
<td>R-Cerebellum</td>
</tr>
</tbody>
</table>

In addition to the study of combined polymorphisms of the DAOA gene, we also studied the effects of a single SNP on the brain connectivity patterns in the patient group. Two subgroups were categorized by the different genotypes of rs2391191 and connectivity patterns were compared under the modulation of rs2391191.
3.4.3 fMRI Data

A SIEMENS TRIO 3-T scanner was used to collect data in the resting state. Before scanning, all the subjects were instructed to lie on their back in the scanner with their eyes closed and have several minutes to acclimatize themselves to the scanner environment. The resting-state fMRI data was acquired with a 2D echo planar imaging pulse sequence with parameters TR/TE= 2000/30; thickness/gap=4/0 mm; matrix=64×64; FOV=192×192 mm; flip angle=90°. In addition, high resolution T1 scans were also used for localization using a 3D-FLASH sequence: TR/TE=14/4.92ms; thickness/gap = 1.5/0.3 mm; matrix=256×192; FOV=230×230 mm; flip angle=25°; 120 slices. All the image data collected were reconstructed using the software MRIConvert (http://lcni.uoregon.edu/jolinda/MRIConvert/). The first 10 volumes were discarded. Slice timing correction was performed and corrected. Any subjects with > 1.0mm of translation and/or 1.08° of rotation was excluded. Images were then spatially normalized to the MNI (Montreal Neurological Institute) standard EPI template. The fMRI data were then bandpass filtered at 0.01 ∼ 0.08 Hz and spatially smoothed by a 4 × 4 × 4 FWHM Gaussian kernel.

Fifty-two brain ROIs were chosen to learn the brain connectivity networks as shown in Table 3.3. These included representative regions from visual, motor, sensory, attentional, cerebellar, basal ganglia and default mode networks.

3.4.4 Results

We first estimated brain connectivity networks in normal and patient groups based on combined 8 SNPs. The comparison of the Schizophrenia and normal control groups are shown in Fig. 3.5. Common connections and significantly different connections in two groups are shown in Fig. 3.5(a) and (b). A χ² test was employed to decide the significance of the connections related to the disease states. Green connections in Fig. 3.5(b) are unique to normal individuals (p < 0.05) and red connections are unique to the patient group (p < 0.05).

Compared to the patient group, the normal group had higher connectivity density (Normal: 0.0667 ± 0.0058; Patient: 0.0578 ± 0.0008). The global efficiency demonstrated that the patient group had lower communication ability in their brain connectivity networks (Normal: 0.2455 ± 0.0219; Patient: 0.2344 ± 0.0056).
Interestingly, the schizophrenia group had smaller ratio of inter-hemispheric connections as shown in Fig. 3.5(d).

![Graphs showing shared and significant connections between normal and schizophrenia groups](image)

**Figure 3.5:** (a) Common connections for all the subjects in normal and Schizophrenia groups. (b) Significantly different connections associated with normal and Schizophrenia groups. Green connections are associated with the normal control group only and red connections are associated with the Schizophrenia group only. (c) Density and global efficiency observed across groups. (d) Density assessed within hemisphere and across hemispheres across groups.

A number of regions were significantly modulated by the disease states as well as genotypes including the left putamen as shown in Fig. 3.6, the right posterior cingulate gyrus and left middle frontal gyrus \( (p < 0.05) \).

Examples of the subject brain connectivity networks are displayed in Fig. 3.7.
Figure 3.6: (a) The comparison in terms of node degree at the Left Putamen region in both control and Schizophrenia subjects. (b) - (e) Node degree values of the Left Putamen region in Schizophrenia subjects as a function of rs3916966, rs2391191, rs3918341, rs778294 and SNP status respectively.
Figure 3.7: Examples of brain connectivity networks in the group modeling. (a) - (d). Computed brain connectivity for Subjects 01, 03, 05, and 08 in Schizophrenia group. The colors in the figure represent the coefficient weights of the connectivity.

Four subjects (Subject 01, 03, 05, 08) in Schizophrenia group were chosen in Fig. 3.7. According to their genotypes, subjects in Fig. 3.7(a) and (b) are more similar with each other (genotype similarity is 0.875), while subjects in Fig. 3.7(c) and (d) are categorized with higher similarity score (genotype similarity is 1). The colors in the figures represent the coefficient weights which are the regression coefficients estimated from the proposed model. Although the four networks share part of the common connections, we found that the learned networks in Fig. 3.7(a) and (b) are closer in their connectivity structures (structure similarity is 0.8218), and the
Figure 3.8: (a)-(b). The connectivity patterns for the subjects with rs2391191 AG and AA genotypes. The colors in the figure represent the frequency of the learned connections in the subjects. (c). The comparisons of density and global efficiency between two subgroups. (d). The comparisons of density of within hemisphere and across hemisphere between two subgroups.
networks in Fig. 3.7(c) and (d) share more common connections compared with other subjects (structure similarity is 0.8608). This example demonstrated that the proposed method can make inference at the group level, and also can model the diversity in the population modulated by the genetic information.

In addition to the study of combined polymorphisms of the DAOA gene, we also studied the effects of a single SNP on the brain connectivity patterns. In Fig. 3.8, two subgroups were categorized by the different genotypes of rs2391191 (“AA” and “AG” carriers). As described before, we applied the proposed overlapped group fused model and permutation test to determine the brain connectivity network for each subject in the patient group. We summarized the structure of each subject connectivity network within one subgroup together to represent the connectivity networks for each subgroup. In Fig. 3.8(a) and (b), the colors represent the frequency of the learned connections. In each subgroup, the subject connectivity networks maintained high similarity in the structure; while between subgroups, they had different connectivity patterns. As before, global efficiency and density were used in the connectivity patterns comparison. Subjects identified with “AG” carriers have higher connectivity density and larger global efficiency.

3.5 Conclusion and Discussion

In this chapter, we have proposed a framework to incorporate genetic factors into group brain connectivity modeling. We used a fused regression group model that was capable of recovering different connectivity patterns that could effectively deal with individual variability as well as group similarity in both simulations and real fMRI data. In our study, the coordinate descent approach is adopted to solve the optimization problem. It’s worth noting that other algorithms can also be used including the primal-dual method to achieve better computational efficiency [33]. Using the proposed approach, we demonstrated the changes in connectivity patterns under the effect of a specific genetic variation and how this was modulated by disease.

The proposed approach is a novel way for incorporating genetic information into imaging analysis. Unlike previous separate-model approaches where genetic and imaging data were examined separately and then later joined in post-hoc fash-
ion, we incorporated prior genetic knowledge directly into a data-driven model which makes the model more flexible and assists in the final biological interpretation. Furthermore, we have made special efforts to account for intersubject variability. Nevertheless, a potential weakness of our approach is that we have reduced the high-dimensional genetic data to a scalar similarity score. However, incorporating a multidimensional similarity score into the objective function is a direct extension of our work. It is worth to mention that incorporation of the genetic information into fMRI signal modeling can be considered as the data fusion approach for multi-modal analysis. The popular data fusion methods include the canonical correlation analysis (CCA) based [27, 35] and Independent Component Analysis (ICA) based [28, 91, 103] approaches with different statistical assumptions. The data fusion analysis would gather complementary information and benefit the biological interpretation in neuroimaging studies. Another limitation of the proposed approach is the usage of the permutation test in the brain network estimation which may potentially break down the temporal information contained in the data. The alternative ways are to permute the time courses in the Fourier domain or wavelet domain by randomly resampling the phase of the time series [115].

A myriad of factors can affect brain connectivity, and the single SNP examined here will likely only account for a small fraction of observed inter-subject variability. Nevertheless, the genetic information in this model used as prior knowledge still resulted in significant connectivity alterations that were also a function of disease states. We found changes in connectivity in the left putamen, consistent with prior PET studies demonstrating asymmetric $D_2$ receptor changes in the putamen [44], and an $^{18}F2\beta - 3\beta - (4 - \text{fluoro})tropane$ ($^{18}F$ FCT) study that found asymmetric binding potentials in the putamen of chronic schizophrenic patients [78]. Similarly, we found changes in the posterior cingulate connectivity, where decreased grey matter volumes are associated with a poor prognosis in Schizophrenia [102]. A meta-analysis suggested decreased grey matter in a number of brain regions in schizophrenic subjects, including the left middle frontal gyrus [52]. It is tempting to speculate, based on our results, if some or all of the above results would have been affected by knowing DAOA SNP information of the subjects examined.

This framework is our first attempt to integrate the genetic variation into the
brain connectivity modeling. A large number of participants will be recruited in our future study. Finally, we note that the proposed framework is sufficiently general that it can also be combined with other prior information (e.g. clinical indices) to control the brain connectivity modeling. It is a general framework to estimate the group brain connectivity that can model the diversity in a population.

3.6 Asymptotic Properties

Here we justify the proposed method by studying the asymptotic properties of our estimate in the classical framework where $S$, the number of subjects, and $K$, the features are fixed and the sample size $n \to \infty$. Theorem 3.6.1 provides a weak convergence result for the fusion estimate if the regularization strength parameters $\lambda$ and $\gamma$ are chosen properly, implying consistency at the parametric efficient rate.

**Theorem 3.6.1.** For each subject $i = 1, \cdots, S$, we assume that there exist positive definite matrices $C_i = \text{plim}_{n \to \infty} n^{-1}X(i)X(i)^\top$. Suppose further $\lambda/\sqrt{n} \to \lambda_0$ and $\gamma/\sqrt{n} \to \gamma_0$. Then the proposed estimate $\hat{B} = \text{diag}(\hat{B}(1), \cdots, \hat{B}(S))$ satisfies

$$\sqrt{n}(\hat{b} - b) \Rightarrow \text{arg min}(V),$$  

(3.14)

where $\hat{b}_{(KS)\times 1}$ and $b_{(KS)\times 1}$ are vectorized versions of $\hat{B}$ and $B$ by stacking their non-zero columns, respectively, $V : \mathbb{R}^{(KS)\times 1} \to \mathbb{R}$ is given by

$$V(u) = \sum_{i=1}^{S} \left( u_i^\top C_i u_i - 2u_i^\top v_i \right) + \lambda_0 \sum_{i=1}^{S} \sum_{k=1}^{K} \left[ u_k(i) \text{sign}(B_k(i))\mathbb{I}(B_k(i) \neq 0) \\
+ |u_k(i)|\mathbb{I}(B_k(i) = 0)) \right] \\
+ \gamma_0 \sum_{e_{lm} \in E} W(e_{lm}) \left\{ \sum_{k=1}^{K} \left[ (u_k(l) - u_k(m)) \text{sign}(B_k(l) - B_k(m)) \\
- B_k(m)\mathbb{I}(B_k(l) \neq B_k(m)) \\
+ |u_k(l) - u_k(m)|\mathbb{I}(B_k(l) = B_k(m)) \right] \right\}$$

and $\{v_i\}_{i=1}^{S}$ are independent random vectors with densities $N(0, \sigma_i^2 C_i)$. In partic-
ular, if \( \max(\lambda, \gamma) = o(\sqrt{n}) \), then \( \hat{B} \) is a \( \sqrt{n} \)-consistent estimate of \( B \).

Proof. The proof is based on a similar argument as the one used in proving Theorem 4.1 in [26]; so here we only outline the differences. By definition, \( \sqrt{n}(\hat{b} - b) \) minimizes the function

\[
V_n(u) = \sum_{i=1}^{S} \sum_{t=1}^{n} \left( \frac{u_{i}^\top X_t(i)}{\sqrt{n}} - e_{ti} \right)^2 - e_{ti}^2 \\
+ \lambda \sum_{i=1}^{S} \sum_{k=1}^{K} \left| \frac{u_{k}(i)}{\sqrt{n}} + B_k(i) \right| - |B_k(i)| \\
+ \gamma \sum_{e_{lm} \in E} W(e_{lm}) \sum_{k=1}^{K} |B_k(l) - B_k(m)| + \\
\frac{|u_{k}(l) - u_{k}(m)|}{\sqrt{n}} - |B_k(l) - B_k(m)| \\
= I + II + III.
\]

Observe that

\[
I = \sum_{i=1}^{S} \left[ u_{i}^\top (n^{-1} \sum_{t} X_t(i)X_t(i)^\top) u_{i} \right] \\
- 2 \sum_{i=1}^{S} \sum_{t=1}^{n} \frac{u_{i}^\top X_t(i)}{\sqrt{n}} e_{ti}.
\]

So by our assumption and the central limit theorem, it follows that \( I \Rightarrow \sum_{i=1}^{S} \left( u_{i}^\top C_i u_{i} - 2u_{i}^\top v_i \right) \), where \( v_i \sim N(0, \sigma^2 C_i) \) are independent. Then, arguing the limit of term II and III as in the proof of Theorem 4.1 in [26] and applying Slutsky’s convergence together lemma, (3.14) follows from standard epi-convergence results. \( \square \)
Chapter 4

A Sticky Weighted Time Varying Model for Resting State Brain Connectivity Estimation

4.1 Introduction

Inferring brain connectivity networks from fMRI plays an important role for understanding brain functioning both normally, and in disease states. As mentioned in Section 1, many mathematical methods have been developed for brain connectivity modeling. However, most current approaches assume that the connectivity structure is time-invariant, i.e., without considering temporal variations of the underlying neural activity, and thus the inferred brain connectivity possibly a temporally averaged connectivity pattern [92]. Assessing the temporal dynamics of connectivity patterns may therefore represent an additional dimension to assess brain activity [68].

Several strategies have thus far been proposed to investigate brain connectivity dynamics. Lagged interaction based approaches, such as dynamic Bayesian network modeling [83] and auto-regressive (AR) models [54], examine brain interactions simultaneously and over adjacent time steps. State space model based approaches, by combining lagged interaction and filtering theory, estimate non-
stationary brain connectivity at each time point [74]. In addition, time-frequency based approaches, such as wavelet transform based coherence analysis, infer resting state dynamic brain connectivity from places in the time-frequency plane [21]. Wavelet based time varying Granger causality analysis has also been used to produce evolving brain connectivity maps that are modulated by task performance [124].

If brain connectivity networks can be assumed to change slowly and smoothly over time, a sliding window approach is appropriate. By specifying a fixed window length and shifting the window by a given number of data samples, different network learning methods such as correlation [61, 67, 140], covariance [1] and ICA [72, 123] have been applied to estimate the time dependent interactions within each window. However, determining the appropriate window length is critical and difficult: with too small a window length, estimated connectivity patterns suffer from large fluctuations due to noise and thus may not truly reflect the underlying dynamic properties of brain activity; in contrast, too large a window will result in insensitivity to possibly important brain state changes. In order to avoid the assumption that changes occur slowly over time, several studies have reported that functional networks inferred by stationary approaches may be unduly influenced by changes at a few critical time points [92]. These critical time points can be used to segment the entire signal into quasi-stationary sections for the purposes of brain connectivity estimation [37, 86, 92]. Nevertheless, change point detection may be particularly susceptible to artifacts (e.g. due to head motion) in the data. Another important characteristic of both sliding window and change point detection multivariate models is that they assume that different brain regions have temporal variations at the same time scale so that the entire brain dynamics are assumed to switch simultaneously, while in practice, different pairs of brain regions may interact at different temporal scales [68].

Beyond the specific area of fMRI brain connectivity modeling, several time-varying frameworks have been proposed to discover multivariate interactions over time. These include time varying regularized graphical structural estimation [160], linear regression-based Bayesian Network (BN) approaches [120, 133, 153] and change point detection approaches [76]. In a BN framework, both network structure and parameter changes are treated as random processes whose values at each
time epoch are modelled via a BN approach. In one change point detection model, a fused penalty used in a preliminary linear regression model is used to detect change points, and then multivariate regression is separately applied to each segment [76]. Another change point detection approach uses penalized regression and Gaussian mixture models [114].

Based on the above discussion, it is clear that modeling brain dynamics often requires fairly strict assumptions be made, such as assuming the networks change smoothly, change suddenly, or change in a piece-wise stationary fashion [68]. Moreover, learning dynamic changes in brain interactions may be complicated by factors such as head movement, measurement noise and other randomized fluctuations. To reduce the influence of random noise, it is reasonable to assume that brain connectivity patterns mostly change smoothly except at critical change points. Based on this assumption, temporally adjacent networks are more likely to share common patterns than temporally distant networks, as assumed in a weighted time-varying regression model [133], yet abrupt changes can still occur at specific change points.

Therefore, in this chapter, we propose a Sticky weighted time varying (SWTV) model that estimates the non-stationary process of brain interactions in a temporally penalized, weighted regression fashion. We incorporate a fused penalty [142] into the weighted regression model [133]. The fused penalty is added into a weighted regression model in which we can estimate both smooth changing coefficients and abrupt changing structures so that the change point detection problem will not be separated from the network estimation problem. More importantly, the proposed method allows different pairs of brain regions to exhibit fluctuations at different time scales, as illustrated in simulations in Section 4.3. Finally, we assume connections between spatially disparate brain regions are relatively sparse to facilitate meaningful biological interpretation.

In the remainder of this chapter, we will introduce the SWTV model in Section 4.2 and perform simulations to validate the proposed method in Section 4.3. Also, in Section 4.4 the proposed SWTV model is applied to resting state fMRI data sets from both Parkinson’s Disease (PD) and control subjects, and significant different temporal and spatial patterns are found to be associated with the disease state.
4.2 Methods

In this section, we will briefly introduce regression models used in the brain connectivity network modeling at the single subject level, and then we will present the proposed sticky weighted time varying model to estimate dynamic interactions between different brain regions in the resting state.

4.2.1 Sticky Weighted Time Varying Model

Multivariate linear regression models have been widely used to infer neural interactions [89, 121]. As introduced in Section 3.2, the fMRI time course of a ROI is regarded as the response variable, and is predicted from the time courses of all other ROIs at zero-lag as,

$$Y = X\beta + e,$$

(4.1)

where vector $Y$ with length $T$ means the time course of one brain ROI, $X$ is $T \times K$ predictor matrix based on the time courses of all other ROIs with $K + 1$ representing the entire number of ROIs, $\beta$ is the coefficient vector and $e$ is the Gaussian noise term. Due to the non-stationary nature of the brain activity, the time dependent regression model becomes,

$$Y_t = X_t\beta_t + e,$$

(4.2)

where $t$ represents the time index and we need to estimate the regression coefficient vector at each time point respectively. $Y_t$ is the response sample at time point $t$ and $X_t$ is the $t^{th}$ sample row in the predictor matrix. In order to make the connections sparse, one could use an $l_1$ penalty on the regression coefficients. However, with only one sample point, the estimator of coefficients would be extremely unstable. Thus in order to estimate time-varying structures/coefficients, yet still allow sparsity, we assume that the underlying networks are changing smoothly over time. Following prior work [133], we could estimate the coefficients for each time point separately in a Weighted time varying (WTV) model as,

$$\hat{\beta}^{r*} = \arg \min_{\beta^{r*}} \sum_{t=1}^{T} W^{r*}(t) \left( Y_t - X_t \beta^{r*} \right)^2 + \lambda \left\| \beta^{r*} \right\|_{l_1},$$

(4.3)
where $\beta^{t^*}$ is the coefficient vector we need to estimate at time $t^*$, $\hat{\beta}^{t^*}$ is the estimator of $\beta^{t^*}$, and $\lambda$ is the parameter for the $l_1$ penalty. $W^{t^*}(t)$ is the weighting of observations from time $t$ when we estimate the coefficients at time $t^*$. In general, $W^{t^*}(t)$ can be defined as any normalized kernel function. In this chapter, $W^{t^*}(t)$ is defined as,

$$W^{t^*}(t) = \frac{\exp\left(\frac{-(t-t^*)^2}{h}\right)}{\sum_{t=1}^{T} \exp\left(\frac{-(t-t^*)^2}{h}\right)}. \tag{4.4}$$

This is a normalized Gaussian Radial basis function (RBF) kernel, with $h$ representing the kernel band-width. Note that this model is essentially a sparse weighted regression model that allows us to estimate the coefficients at each time point separately by reweighting the observations. With the smoothly changing assumption, temporally adjacent coefficients are more likely to be similar than temporally distant coefficients.

A “sticky” weighted time varying model is therefore introduced as,

$$\begin{align*}
\text{minimize} & \sum_{t^*} \sum_{t=1}^{T} W^{t^*}(t) \left( Y_t - X_t \beta^{t^*} \right)^2 + \lambda \sum_{t^*=1}^{T} \| \beta^{t^*} \|_{l_1} \\
& + \gamma \sum_{t^*=2}^{T} \| \beta^{t^*} - \beta^{t^*-1} \|_{l_1},
\end{align*} \tag{4.5}$$

where $\gamma$ is the parameter to control the fused penalty and serves to keep the coefficients temporally consistent except at (possibly several) abrupt change points.

To efficiently solve this optimization problem, we can rewrite the response vector and predictor matrix as $\tilde{Y}^{t^*} = \sqrt{W^{t^*}(t)} Y_t$ and $\tilde{X}^{t^*} = \sqrt{W^{t^*}(t)} X_t$. Let $Y^{t^*} = (Y_1^{t^*}, Y_2^{t^*}, \cdots, Y_T^{t^*})'$, $X^{t^*} = (X_1^{t^*}, X_2^{t^*}, \cdots, X_T^{t^*})'$, then the weights can be incorporated into the square loss function directly. We can further simplify the objective function by expressing them in a matrix format. Suppose $\tilde{Y} = (Y^1, Y^2, \cdots, Y^T)'$ is a response vector with length $TT$, $\tilde{X} = \text{diag}(X^1, X^2, \cdots, X^T)$ is a block diagonal matrix with dimension $TT \times TK$, and $\tilde{\beta} = (\beta^1, \beta^2, \cdots, \beta^T)'$ is the concatenated time varying coefficient vector with length $TK$. Each $\beta^t$ corresponds to the coefficient vector at time point $t$. The objective function can be formulated as,
\[
\text{minimize } \left\| \tilde{Y} - \tilde{X}\hat{\beta} \right\|_F^2 + \lambda \left\| \hat{\beta} \right\|_{l_1} + \gamma \left\| C\hat{\beta} \right\|_{l_1}.
\]  

(4.6)

where \( C \) is a sparse \((T - 1)K \times TK \) matrix with two nonzero elements in each row. More specifically, \( C((T - 1) \ast (i - 1) + j, K \ast (j - 1) + i) = -1 \) and \( C((T - 1) \ast (i - 1) + j, K \ast j + i) = 1, i = 1, 2, \cdots, K, j = 1, 2, \cdots, T - 1 \). By this formulation, it becomes a generalized fused LASSO problem which can be solved using the smoothing proximal gradient (SPG) method [25]. SPG is one efficient algorithm with the convergence rate of \( O(\frac{1}{\varepsilon}) \) where \( \varepsilon \) is the precision of the algorithm; Per-iteration complexity of SPG is linear with the number of nonzero elements of the constructed sparse network \( C \) [25]. In this study, we use the SPG optimization toolbox [25], and the implementation of the SWTV model is described in Table 4.1. It’s worth noting that other algorithms can also be used including the primal-dual method to solve this optimization problem [33].

4.2.2 Model Selection

The parameters of stationary regression model are usually determined by cross validation (CV) which separates the data into training and testing sets. However, the standard CV approach cannot be employed directly in our time-varying case, since each sample corresponds to a specific time point and the structures and coefficients may be different across time.

Therefore, in order to apply cross validation, we first up-sample the data by a factor of two: the odd samples represent the original data points and even samples are the interpolated data points. For the purpose of model selection, we assume that the corresponding even samples have the same temporal properties as those of odd samples. In the following simulation studies, by treating the odd samples as the training set and even samples as the testing set, we can select optimal parameters of the model.

For each fixed set of parameters \( \lambda, \gamma \) and the bandwidth \( h \), we can estimate time-varying coefficients as described in Table 4.1 and use the cross validation to select the optimal values. However, for large scale data sets, cross validation can be time consuming, and it may not be feasible for the large scale problems. Instead, the gradient descent approach can be applied to iteratively update each
Table 4.1: Implementation of sticky weighted time varying model.

<table>
<thead>
<tr>
<th>Input:</th>
<th>$Y \in \mathbb{R}^T$, $X \in \mathbb{R}^{T \times K}$, regularization parameters $\lambda$, $\gamma$ and band width $h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: Weighting the response vector $Y$ and the predictor matrix $X$.</td>
<td></td>
</tr>
</tbody>
</table>
1. Based on $h$, calculate $W^t(t)$ according to Eq. (4.4), for $t = 1, 2, \cdots, T$.
2. $Y_i^t \leftarrow \sqrt{W^t(t)}Y_i$, $t = 1, 2, \cdots, T$.
3. $X_i^t \leftarrow \sqrt{W^t(t)}X_i$, $t = 1, 2, \cdots, T$.

| Step 2: Constructing the objective function. |  
1. $\tilde{Y} \leftarrow (Y^1, Y^2, \cdots, Y^T)'$.
2. $\tilde{X} \leftarrow \text{diag}(X^1, X^2, \cdots, X^T)$.
3. Construct a sparse matrix $C \in \mathbb{R}^{(T-1)K \times TK}$ as,
   
   $C((T-1) * (i-1) + j, K * (j-1) + i) \leftarrow -1,$
   
   $C((T-1) * (i-1) + j, K * j + i) \leftarrow 1,$
   
   $i = 1, 2, \cdots, K, \quad j = 1, 2, \cdots, T-1.$
4. Formulate the objective function as in Eq. (4.6) by inputting $\tilde{Y}, \tilde{X}, \tilde{\beta}, C, \lambda, \gamma$.

| Step 3: Estimating time-varying coefficients $\tilde{\beta} = (\beta^1, \beta^2, \cdots, \beta^T)'$. |  
Apply SPG toolbox [25] to solve the optimization problem in Eq. (4.6) and obtain the estimated time varying coefficients $\tilde{\beta}$.

| Output: | Time varying coefficients $\tilde{\beta} = (\beta^1, \beta^2, \cdots, \beta^T)'$. |

4.2.3 Statistical Analysis

To perform inference of brain connectivity networks, we utilize a linear regression approach which has been described in Section 3.2.3. Generally speaking, we treat each ROI in turn as the response vector and all other ROIs as constituting the parameter as described in [26, 75]. It sequentially applies three line searches along each descent direction to minimize the corresponding mean square error of current CV until the error convergences. For the static sparse regression model used in the simulation part, we utilize the stability selection approach which is proved to enhance selection accuracy [101].
predictor matrix. In this way, the time varying coefficient vector for each ROI is estimated one by one until we obtain the whole brain network.

To quantify and compare the temporal variability of the inferred networks, we define the network variation as,

\[ V = \frac{1}{T-1} \sum_{t=2}^{T} \| G(t) - G(t-1) \|_F^2, \]  

(4.7)

where \( t \) represents the time index and \( G(t) \) represents the brain connectivity network which is a matrix estimated using SWTV model as described before at time point \( t \). Network variation calculates the average of distance between two brain connectivity networks at adjacent time points. It quantifies the changes of network structures as well as connectivity strengths. This metric measures the ability of switching or oscillation of the networks across the time. While a relatively simple measure, we found \( V \) an intuitive way to quantify and compare the temporal changes of brain connectivity networks: with a fixed \( \gamma \), a higher \( V \) implies higher moment-to-moment variability in the networks (Fig. 4.5).

### 4.3 Simulations

To validate the proposed method, we performed simulations to compare the performance of SWTV model with both that of the weighted time varying model and a static sparse regression model. We considered different simulation settings where the different variables changed in the same time scale with and without autocorrelation, and also when they changed in different time scales.

In brief, the simulated data were generated from a Gaussian model with changing structures and coefficients as \( Y_t = X_t \beta_t + e_t \). \( X_t \) was a randomly generated sample row at time point \( t \) with \( K \) variables (i.e., a \( 1 \times K \) row vector), \( \beta_t \) was a time dependent coefficient vector with same length \( K \) (\( K = 20 \)) and \( e_t \) was white Gaussian noise.

More specifically:

1. We first generated the changing coefficients \( \beta_t \). In the first and second simulations, we assumed that all the variables changed in the same time scale as \( N \) and the total length of sample size was \( T = 3 \times N \) (an example of this is shown
Figure 4.1: Results for the first simulation. (a). The true Model. (b)-(d). Models learned by static LASSO, Weighted Time Varying model, and Sticky Weighted Time Varying model respectively in the first simulation. The time index is along the x-axis, the variable index is along the y-axis and the color bar represents the coefficients’ strength.

In Fig. 4.1 (a)). In the third simulation, different coefficients could have different time scales as shown in Fig. 4.3 (a). The averaged time scale and sample size were set to $N$ and $T = 3 \times N$ respectively.

(2) The design matrix $X$ was randomly generated containing $T$ observations and $K$ predictors. The error vector $e$ was Gaussian noise $\sim N(0,1)$. The response vector $Y$ was generated by $Y_t = X_t \beta^t + e_t$ with $t = 1, \cdots, T$. In the second simulation, to generate the autocorrelation structures on data, the Gaussian smooth filter
Figure 4.2: Results for the second simulation. (a). The true Model. (b)-(d). Models learned by static LASSO, Weighted Time Varying model, and Sticky Weighted Time Varying model respectively in the second simulation. The time index is along the x-axis, the variable index is along the y-axis and the color bar represents the coefficients’ strength.

with variance 1 was applied on X and Y separately.

We compared the proposed SWTV model with the weighted time varying model [133] and the static LASSO model. The cross validation was used for parameter selection in SWTV and WTV models as discussed before, and stability selection was used for static LASSO model.

In the simulations, we tested the performance of the algorithms as a function of the number of time scales $N$. For reliable assessment, each procedure was re-
Figure 4.3: Results for the third simulation. (a). The true Model. (b)-(d). Models learned by static LASSO, Weighted Time Varying model, and Sticky Weighted Time Varying model respectively. The time index is along the x-axis, the variable index is along the y-axis and the color bar represents the coefficients’ strength.
Figure 4.4: Simulation Results. (a) F1 scores of the first simulation. (b) F1 scores of the second simulation. (c) F1 scores of the third simulation. Red lines represent the F1 scores of the proposed method, blue lines represent the F1 scores of weighted time varying model and the green lines represent the F1 scores of the static model.

As expected, the recovered coefficients by the static model were indeed determined in part by the critical samples, and larger fluctuations were observed in the estimated coefficients in WTV model when compared with those of SWTV model in all the simulations.

The F1 score was employed to quantitatively evaluate the general performance by considering both the Type I and Type II error rates.
As shown in Fig. 4.4, we compared F1 scores at the time scale (or averaged time scale) \( N = 15, 20, 30, 50 \) and 70 respectively. The results demonstrated that with the increasing of time scales, the proposed SWTV and WTV models had better accuracy in recovering time varying structures. When the time scale was smaller than 15 time points, it may be unreliable to estimate the true changing structures. Compared with the other two methods, the simulation results demonstrated that the proposed SWTV model yield higher accuracy in recovering time-dependent structures in all the simulations. After adding the autocorrelation structures into the data in the second simulation, we observed that the accuracy of recovering the underlying time varying structures decreased compared with those of data without autocorrelation. In the third simulation, although the averaged time scale was fixed, the structures may change within a shorter time period. As a result, the F1 scores were lower compared with those of first simulation. By having the capability of estimating both smoothly changing and abrupt changes, the proposed SWTV model more accurately estimated the underlying time-varying brain connectivity patterns.

### 4.4 Real Application

In this section, we apply the proposed method to a real resting state fMRI data set and study the dynamic properties of brain connectivity networks in subjects with Parkinson’s Disease (PD). We first estimate the time varying brain connectivity networks for each subject and then compare the temporal and spatial patterns of the inferred connectivity networks.

#### 4.4.1 Subjects and fMRI Resting State Data Set

Twelve PD subjects and ten healthy control subjects were recruited from Pacific Parkinson’s Research Center (PPRC) at the University of British Columbia (UBC). All the experiments were approved by the Ethics Board at UBC, and all the subjects have provided informed consent prior to experiment participation.

A 3 Tesla scanner (Philips Gyroscan Intera 3.0T; Philips Medical Systems, Netherlands) equipped with a head-coil was used to collect data in the resting state. Before scanning, all the subjects were instructed to lie on their back in the
Table 4.2: The index and name of 48 selected brain ROIs. ‘L’ represents the brain left side and ‘R’ represents the brain right side.

<table>
<thead>
<tr>
<th>Index</th>
<th>Name</th>
<th>Index</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>L-Cerebellum</td>
<td>R1</td>
<td>R-Cerebellum</td>
</tr>
<tr>
<td>L2</td>
<td>L-PMd</td>
<td>R2</td>
<td>R-PMd</td>
</tr>
<tr>
<td>L3</td>
<td>L-PMv</td>
<td>R3</td>
<td>R-PMv</td>
</tr>
<tr>
<td>L4</td>
<td>L-Pre-SMA</td>
<td>R4</td>
<td>R-Pre-SMA</td>
</tr>
<tr>
<td>L5</td>
<td>L-SMA-proper</td>
<td>R5</td>
<td>R-SMA-proper</td>
</tr>
<tr>
<td>L6</td>
<td>L-ACC</td>
<td>R6</td>
<td>R-ACC</td>
</tr>
<tr>
<td>L7</td>
<td>L-Caudate</td>
<td>R7</td>
<td>R-Caudate</td>
</tr>
<tr>
<td>L8</td>
<td>L-Cerebellum-Cortex</td>
<td>R8</td>
<td>R-Cerebellum-Cortex</td>
</tr>
<tr>
<td>L9</td>
<td>L-PFC</td>
<td>R9</td>
<td>R-PFC</td>
</tr>
<tr>
<td>L10</td>
<td>L-Pallidum</td>
<td>R10</td>
<td>R-Pallidum</td>
</tr>
<tr>
<td>L11</td>
<td>L-Putamen</td>
<td>R11</td>
<td>R-Putamen</td>
</tr>
<tr>
<td>L12</td>
<td>L-Somatosensory</td>
<td>R12</td>
<td>R-Somatosensory</td>
</tr>
<tr>
<td>L13</td>
<td>L-Thalamus-Proper</td>
<td>R13</td>
<td>R-Thalamus-Proper</td>
</tr>
<tr>
<td>L14</td>
<td>ctx-L-caudalmiddle-facing</td>
<td>R14</td>
<td>ctx-R-caudalmiddle-facing</td>
</tr>
<tr>
<td>L15</td>
<td>ctx-L-cuneus</td>
<td>R15</td>
<td>ctx-R-cuneus</td>
</tr>
<tr>
<td>L16</td>
<td>ctx-L-inferiorparietal</td>
<td>R16</td>
<td>ctx-R-inferiorparietal</td>
</tr>
<tr>
<td>L17</td>
<td>ctx-L-inferiortemporal</td>
<td>R17</td>
<td>ctx-R-inferiortemporal</td>
</tr>
<tr>
<td>L18</td>
<td>ctx-L-lateraloccipital</td>
<td>R18</td>
<td>ctx-R-lateraloccipital</td>
</tr>
<tr>
<td>L19</td>
<td>ctx-L-middletemporal</td>
<td>R19</td>
<td>ctx-R-middletemporal</td>
</tr>
<tr>
<td>L20</td>
<td>ctx-L-precentral</td>
<td>R20</td>
<td>ctx-R-precentral</td>
</tr>
<tr>
<td>L21</td>
<td>ctx-L-precuneus</td>
<td>R21</td>
<td>ctx-R-precuneus</td>
</tr>
<tr>
<td>L22</td>
<td>ctx-L-superiorparietal</td>
<td>R22</td>
<td>ctx-R-superiorparietal</td>
</tr>
<tr>
<td>L23</td>
<td>ctx-L-superiortemporal</td>
<td>R23</td>
<td>ctx-R-superiortemporal</td>
</tr>
<tr>
<td>L24</td>
<td>ctx-L-supramarginal</td>
<td>R24</td>
<td>ctx-R-supramarginal</td>
</tr>
</tbody>
</table>

scanner and have several minutes to acclimatize themselves to the scanner environment with eyes closed. Blood oxygenation level-dependent (BOLD) contrast echo-planar (EPI) T2*-weighted images were taken with the following specifications with a repetition time of 1985 ms, echo time of 37 ms, flip angle 90°, field of view (FOV) 240.00 mm, matrix size 128×128, with pixel size 1.9 mm×1.9 mm. The SENSE acceleration was used in EPI acquisition. The duration of each functional run was 4 mins during which we obtained 36 axial slices with 3 mm thickness and 1 mm gap thickness. The FOV was set to include the cerebellum.
ventrally and also include the dorsal surface of the brain. 48 Freesurfer-derived ROIs in total were chosen in this study as shown in Table. 4.2.

4.4.2 Results

To apply the proposed method on the subjects with PD and control subjects, we need to choose the band width \( h \), the sparse penalty parameter \( \lambda \) and the fused penalty parameter \( \gamma \) in the SWTV model. We conducted parameter selection using gradient descent approach for each subject. The optimal parameters for each subject varied across a broad range of values (\( h = 38.9091 \pm 26.1769, \lambda = 0.5117 \pm 0.4382, \gamma = 1.5341 \pm 0.8308 \)). Using the optimal values, the density of learned connectivity networks varied from 0.0294 to 0.3326, making it difficult to compare two groups. To fairly compare the connectivity patterns of patient and control groups, we chose fixed parameters/densities for all the subjects in this study.

Although a few studies have been conducted on time variation in connectivity networks, the exact time scale of brain activities is unclear, and varies between subjects. This is especially true in resting state studies where subjects are asked to lie quietly and not think of anything in particular, so the exact temporal patterns of brain activity may vary across the population. Similar to the choice of sliding window length, a small bandwidth \( h \) will suffer from large fluctuations while a large band width \( h \) may reduce sensitivity to fluctuations in the signal. Following prior work [61] as well as our preliminary studies, we set the bandwidth to 32s (16 points). A comprehensive comparisons of brain connectivity variation scales will be conducted in future work.

Fig. 4.5 (a) demonstrates the relationship between the averaged number of connections and the sparsity parameter \( \lambda \) as applied to one control subject. Suppose \( \lambda_0 = 0, \lambda_1 = 0.1, \lambda_2 = 0.2, \ldots, \lambda_{20} = 2 \), it is apparent that the averaged number of connections decreases with sparsity parameter \( \lambda \) increasing. We compare the averaged common connections between the inferred networks with \( \lambda_i \) and \( \lambda_{i-1} \) (\( i = 2, \ldots, 20 \)), as shown in Fig. 4.5 (a). Reassuringly, we observed that the connectivity inferred with a larger \( \lambda \) is mostly contained in the estimated networks with a smaller \( \lambda \). In other words, important connections will always be selected. In our study, we learned the networks with a fixed sparsity parameter \( \lambda \) of 0.5 for
Figure 4.5: (a) The averaged number of detected connections within the networks as sparsity parameter $\lambda$ increases, with a fixed fused penalty parameter $\gamma = 0.5$. The blue line represents the averaged number of detected connections as a function of $\lambda$. The red dashed line represents the averaged number of common connections with network detected by a smaller $\lambda$. E.g., the first point of the red line is the number of averaged common connections between networks with $\lambda = 0.1$ and $\lambda = 0$. (b) The network variations of networks as fused penalty parameter $\gamma$ increases, with a fixed $\lambda = 0.5$. 
a fair comparison at the population level. We also compared the temporal patterns with fixed sparsity (0.1) across all the subjects. Fig. 4.5 (b) demonstrates the relationship between the value of network variation and the fused penalty parameter $\gamma$ when applied to one control subject. The network variation generally decreases with increasing values of $\gamma$. However, since we are interested in the relative differences between control and patient groups, we set $\gamma = 1.5$ for all subjects. We also compared the connectivity networks between two groups when $\gamma = 0.5$.

Fig. 4.6 and Fig. 4.7 demonstrate the examples of time varying brain connectivity networks of typical normal and PD subjects at different time points where the networks are learned with fixed parameters $h = 32s, \lambda = 0.5, \gamma = 1.5$. The proposed method could estimate the brain connectivity networks with both changing structures and coefficients. When compared with the normal subject, we note that the PD subject shows a sparser network. In addition, the PD subject has more distributed connections while the normal control subject tends to incorporate more hub regions in brain connectivity networks.

To measure temporal properties of the learned time varying brain connectivity networks, we compared the network variations between control and PD groups in Fig. 4.8. The averaged network variations were significantly lower in the PD group, whether or not the sparsity parameter $\lambda = 0.5$ (Fig. 4.8(a)) or fixed sparsity (0.1) (Fig. 4.8(b)). Fig. 4.9 compares the averaged time period, defined as the duration of non-zero values, between control and PD groups. We note that the PD group has a larger time period with different parameters compared with that of control group. If we consider “switching ratio”, defined as number of time points with switching from zero to non-zero states to the total length of time points, we note that the PD group had a significantly smaller switching ratio as shown in Fig. 4.10.

We have also investigated the spectrum properties of the inferred brain connectivity networks. We note that the most dominant low frequency connectivity fluctuations are below 0.02 Hz, and specifically at around 0.005 and 0.015 Hz, which are consistent with previous studies [1]. While we found no significant differences between groups in the mean frequency, we suspect this could be due to the relatively small subject size in our study.

In addition to the temporal dynamics, we also studied the spatial patterns learned
by the fixed sparsity (0.1) by examining “consistent” connections over time. We define consistent connections as those connections that appear at least once at one time point in all subjects within a given group. As shown in Fig. 4.11, the PD group has fewer cortico-basal ganglia connections and more cortico-cortical connections compared to the control group. The alterations in cortico-cortical and cortico-basal connectivity may reflect compensatory connections to ameliorate the effects of the diseased basal ganglia [109–111].

4.5 Conclusion and Discussion

It is clear that the brain is inherently non-stationary. Therefore, studying dynamic properties of brain connectivity networks could extend our understanding of brain functioning. In this chapter, a penalized weighted regression model is presented to estimate both smooth and abrupt changes in the time dependent brain connectivity patterns. Compared with previous multivariate time varying approaches introduced for fMRI brain connectivity modeling, the proposed SWTV model is more flexible and allows different pairs of brain regions to have different dynamic time scales. While the proposed method is designed for the time evolving networks estimation, when the underlying models are static, the proposed method could still accurately estimate networks with appropriate parameters.

When applied to real fMRI resting state data consisting of 12 subjects with PD and 10 control subjects, PD subjects had significantly reduced network variation, likely related to impaired cognitive flexibility in PD. This highlights the importance of establishing dynamic properties in PD subjects.

While the proposed method appears promising, there are a number of limitations. The synthetic time courses generated in the simulations could be more realistic by embedding the long memory component using the fractional Gaussian noise model [30, 31, 63]. We had to estimate certain parameters, such as sparsity and temporal bandwidth. Further work will be required to more comprehensively investigate time varying brain connectivity patterns over a broad range of time scales. Resting state data is particularly challenging in this regard, as different subjects will undoubtedly have different temporal patterns. We used a very simple metric to estimate temporal variability of the network, this could be expanded in future work.
A previous study has suggested that larger brain regions tend to show greater connectivity variability, while the smaller regions are more stable [68]. Nevertheless, the disease related changes of the time-varying patterns in brain connectivity such as we observed might be the potential biomarker for future studies [72, 85].
Figure 4.6: Time varying brain connectivity networks learned with fixed parameters ($h = 32$, $\lambda = 0.5$, $\gamma = 1.5$) for one typical control subject at (a) $t = 70$ s (35 points), (b) $t = 80$ s (40 points) and (c) $t = 90$ s (45 points). The blue lines and red lines represent the positive and negative coefficients respectively. The thicknesses of the lines represent the absolute values of the coefficients.
Figure 4.7: Time varying brain connectivity networks learned with fixed parameters \( h = 32s, \lambda = 0.5, \gamma = 1.5 \) for one typical PD subject at (a) \( t = 70s \) (35 points), (b) \( t = 80s \) (40 points) and (c) \( t=90s \) (45 points). The blue lines and red lines represent the positive and negative coefficients respectively. The thicknesses of the lines represent the absolute values of the coefficients.
Figure 4.8: The comparison of network variation between the normal and PD group with either (a) fixed sparsity penalty parameter (\( \lambda = 0.5 \)) or (b) fixed sparsity (0.1).

Figure 4.9: The comparison of averaged time period between the normal and PD group with either (a) fixed sparsity penalty parameter (\( \lambda = 0.5 \)) or (b) fixed sparsity (0.1).
Figure 4.10: The comparison of switching ratio between the normal and PD group with either (a) fixed sparsity penalty parameter ($\lambda = 0.5$) or (b) fixed sparsity (0.1).
Figure 4.11: Connections that consistently appear in at least one time point in all subjects in (a) the control group and (b) the PD group.
Chapter 5

A Combined Static and Dynamic Model for Resting State Brain Connectivity Estimation

5.1 Introduction

As the brain is inherently non-stationary, assessing the temporal dynamics of brain connectivity patterns, represents an additional dimension through which to gain deeper insights into brain activity [17, 68].

The dynamics of brain connectivity networks are particularly important as they are associated with a variety of neurodisorders such as Schizophrenia [124], multiple sclerosis [80], Parkinson’s Disease [90] and post-traumatic stress disorder [85]. For instance, altered contributions of brain connectivity dynamic patterns have been reported in subjects with multiple sclerosis [80]. The network variations of subjects with Parkinson’s Disease was decreasing compared with that of control subjects, an observation that may be related with the cognitive rigidity – i.e. difficulty between switching tasks – that is frequently observed in PD [90].

A few studies have investigated brain connectivity dynamics [7, 42, 60, 77, 124] and have demonstrated that connectivity can be mediated by learning and/or task performance [7, 42, 124]. In addition to task-related connectivity changes, var-
ious groups have assessed dynamic changes during resting state fMRI. Frequently, the assumption is made that the brain networks change slowly and smoothly with time, so a sliding window based approach is used [1, 49, 61, 67, 72, 123, 140]. An alternative to assume smoothly changing connectivity is that the brain states are relatively stable between a few critical time points [92]. These critical time points thus can be used to segment the entire brain signals into quasi-stationary sections for the purposes of brain connectivity estimation [37, 86, 92].

It is becoming increasingly apparent that neither a continually changing nor quasi-stationary model is adequate for brain connectivity modelling. For example, when studying functional coupling between brain regions across different task states, it is found that central tendencies dominate the coupling configurations, though different task states show significant differences in brain connectivity patterns [77]. Similar conclusions have been made when comparing the network patterns in resting and different task states: A similar intrinsic architecture is present across the tasks and resting state [32]. Thus the brain very likely maintains relatively stable connections with superimposed flexible alterations. The central dominant architecture may be due to the underlying brain structures, and should play key roles in executing intrinsic functions of the brain [77].

Therefore, in this chapter, we propose a Combined static and dynamic model (CSDM) for brain connectivity estimation. Instead of only investigating the static or dynamic features separately, we aim to jointly make inference of time-invariant and time-dependent connectivity patterns. The proposed combined approach is inspired by the so-called Dirty model (DM) which was originally developed for multitask learning [70]. Multitask learning is statistically effective when recovering the related features across different tasks jointly. For instance, the group Lasso approach has been used to estimate brain connectivity networks at the group level [154]. Different from group Lasso which only encourages the shared features, dirty model is able to recover the individual task dependent features in addition to the shared common features. It is more realistic and outperforms both Lasso and group Lasso models [70].

Though being a powerful method, DM has never been studied for time-varying brain connectivity network modeling. Therefore in this chapter, we introduce the dirty model to the brain network modelling area and extend it to the time-varying
settings based on the sliding window framework. In our formulation, each time window is treated as one task and we simultaneously make inference of all time windows (multitask learning). It can model the temporal brain connectivity patterns by jointly estimating the time invariant and time dependent connections. In our proposed approach, the multitask learning model mainly serves as the feature selection tool in the time-varying setting. With the selected static and dynamic variables, a least square model is further adopted to better estimate the coefficients.

In the remainder of this chapter, we will present the combined static and dynamic model and discuss the interpretations of the learned temporal brain connectivity networks in Section 5.2. The proposed approach is compared with the Lasso and group Lasso models in the time-varying setting in Section 5.3. In Section 5.4, the proposed CSDM approach is applied to a resting state fMRI study in PD. The different roles of the static and dynamic connectivity features will be discussed.

5.2 Methods

The multivariate linear regression model is one of the most commonly used methods in brain connectivity network modeling as we discussed in Section 3.2. With the extension to the multitask setting, multiple data sets could be simultaneously inferred. It allows multiple tasks to share some common features such as sparsity, yet still fit the individual multivariate data well. In this section, multivariate regression models will be first briefly described, and we then will focus on the proposed combined static and dynamic model for resting state temporal brain connectivity modeling.

5.2.1 Combined Static and Dynamic Brain Connectivity Network Estimation

In the regression model used for brain connectivity network estimation, the fMRI time course of one brain Region of Interest (ROI) is regarded as the response vector, and the time courses of all other ROIs are treated as predictor variables,

\[ Y = X\beta + e, \]  

where the response variable \( Y \) denotes the time course of one brain ROI with sam-
ple length $T$, $X$ is predictor matrix with dimension $T \times K$ based on the time courses of all other ROIs, $\beta$ is the coefficient vector and $e$ is the Gaussian noise term. To infer time varying connectivity, the time dependent regression model should be estimated at each time point. However, with only one sample point, the estimator would be extremely unstable. Thus in order to estimate time-varying structures/coefficients, the weighted regression and sliding window strategies are usually employed with the assumption that the underlying connectivity networks are changing slowly and smoothly across the time. In this chapter, we adopt the sliding window approach. Suppose that the window length is $L$ and the step size of moving the window is one, the sliding window based time dependent regression model can be represented as,

$$Y_t = X_t \beta_t + e_t, \quad (5.2)$$

where $t$ represents the window index and we need to estimate the regression coefficient vector within each window respectively. $Y_t$ is the response vector at window $t$ with length $L$ and $X_t$ is $L \times K$ predictor matrix. The total number of windows is $M = T - L + 1$. $\beta_t$ is the coefficient vector we need to make inference at window $t$.

Determining the window length $L$ is challenging. Too large a window may result in insensitivity to possible important temporal features. While with too short a window, the estimated connectivity may suffer from large fluctuations and inference difficulties. Simultaneously modeling the regression models of all windows may improve statistical efficiency and the multitask learning model is suitable to jointly make inference of all windows.

Another realistic assumption is that brain networks will maintain certain time-invariant connectivity patterns while still allow specific transient features. One solution to the overlapped feature selection in the multitask setting is the dirty model which leverages the constraint on the common features across the tasks [70]. When extending to the time varying setting, we can denote the model as,

$$Y_t = X_t S + X_t D_t + e_t, \quad (5.3)$$

where $S$, which is time-invariant, represents the static coefficients, $D_t$ means the temporal dynamic coefficients and the combined coefficients are denoted as $B_t =$
The objective function for inferring $S$ and $D_t$ is formulated as,

$$
\begin{align*}
\min_{S, D_t \in \mathbb{R}^M} & \quad \frac{1}{2L} \sum_{t=1}^{M} \left\| Y_t - X_t(S + D_t) \right\|_2^2 \\
& + \lambda \| S \|_{1,\infty} + \gamma \| D_t \|_{1,1},
\end{align*}
$$

(5.4)

where $Y_t$ denotes the time course of one brain ROI and $X_t$ is predictor matrix based on the time courses of all other ROIs at window $t$ ($t = 1, 2, \ldots, M$, with $M$ being the total number of windows based on $M = T - L + 1$, $T$ is the sample length and $L$ is the window length). $\lambda$ is the tuning parameter for controlling the static coefficients’ sparsity and $\gamma$ is the tuning parameter to constrain the time-dependent coefficients’ sparsity. Suppose $E$ is a vector, $C$ is a matrix and $C(j)$ is the $j^{th}$ row of $C$. $\| E \|_2^2$ is defined as the sum of square values of elements in $E$. $\| C \|_{1,1}$ is defined as the sum of absolute values of elements in $C$. $\| C \|_{1,\infty} = \sum_j |C(j)|_\infty$ and $|C(j)|_\infty$ is defined as the maximum value in row $C(j)$. This optimization problem can be solved using the accelerated gradient methods (AGM) implemented in MALSAR (multi-task learning via structural regularization) toolbox [159].

However, in real applications, the penalty could substantially shrink the coefficients and lead to imprecise estimators. Therefore the above multitasking learning model mainly serves as the feature selection tool in our proposed approach. With the selected variables in Eqn. 5.4, the least square model is further adopted to better estimate the coefficients. Suppose the selected variables for static and dynamic features are $X_t^S$ and $X_t^D$ respectively. The objective function of the least square estimation is formulated as,

$$
\begin{align*}
f(S, D) = & \sum_{t=1}^{M} \left\| Y_t - X_t^S S - X_t^D D_t \right\|_2^2
\end{align*}
$$

(5.5)

where $X_t^S$ and $X_t^D$ are the predictor matrices at window index $t$ for the static and dynamic variables respectively.

The derivations with regard to $S$ and $D_t$ are,
\[
\frac{\partial f}{\partial S} = -2 \sum_{t=1}^{M} X_t^S Y_t + 2 \sum_{t=1}^{M} X_t^S X_t^D D_t + 2 \sum_{t=1}^{M} X_t^S X_t^S S
\]  
(5.6)

\[
\frac{\partial f}{\partial D_t} = -2 X_t^{D'} Y_t + 2 X_t^{D'} X_t^S S + 2 X_t^{D'} X_t^D D_t
\]  
(5.7)

Setting Eqn. 5.6 and Eqn. 5.7 to be zeros, we could get the static coefficients \( S \) and time dependent coefficients \( D_t \) as,

\[
S = \left( \sum_{t=1}^{M} X_t^S Z_t X_t^S \right)^{-1} \left( \sum_{t=1}^{M} X_t^S Z_t Y_t \right)
\]  
(5.8)

\[
D_t = (X_t^{D'} X_t^D)^{-1} X_t^{D'} Y_t - (X_t^{D'} X_t^D)^{-1} X_t^{D'} X_t^S S
\]  
(5.9)

where \( Z_t = I_L - X_t^D (X_t^{D'} X_t^D)^{-1} X_t^{D'} \) and \( I_L \) is the \( L \times L \) identity matrix. The implementation of the combined static and dynamic model is summarized in Table 5.1.

To perform inference of brain connectivity networks, we utilize a linear regression approach. We treat each ROI in turn as the response vector and the signals from all other ROIs as the predictor matrix. The combined static and dynamic model as described in Table 5.1 will be employed to infer the coefficients and the corresponding estimated coefficient vector would give the strength of connectivity from all other ROIs to the response ROI, i.e., the estimated directed connections. In this way, the temporal coefficient vector for each ROI is estimated one by one until we obtain the whole brain network. The tuning parameters \( \lambda \) and \( \gamma \) could be determined according to the Bayesian Information Criterion or pre-specified network sparsity. However, in this study, we would either fix the parameters or densities for a fair group comparison.
Table 5.1: Implementation of combined static and dynamic model for time varying coefficients estimation.

<table>
<thead>
<tr>
<th>Input: $Y \in \mathbb{R}^T$, $X \in \mathbb{R}^{T \times K}$, regularization parameters $\lambda$, $\gamma$ and window length $L$.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1:</strong> Defining the response vector $Y_t$ and predictor matrix $X_t$ at each time window, $t = 1, 2, \ldots, M, M = T - L + 1$.</td>
</tr>
<tr>
<td><strong>Step 2:</strong> Selecting the static and dynamic features $X^S$, $X^D$ according to Eqn. 5.4.</td>
</tr>
</tbody>
</table>
| $\begin{align*}
\text{minimize} & \quad \frac{1}{2T} \sum_{t=1}^{M} \| Y_t - X_t (S + D_t) \|_2^2 + \lambda \| S \|_{1,\infty} + \gamma \| D_t \|_{1,1} .
\end{align*}$ |
| **Step 3:** Based on the selected static and dynamic features, estimating coefficients according to Equ. 5.8 and Equ. 5.9. |
| $S = \left( \sum_{t=1}^{M} X^S_t Z_t X^S_t \right)^{-1} \left( \sum_{t=1}^{M} X^S_t Z_t Y_t \right),$ |
| $D_t = (X^D_t X^D_t)^{-1} X^D_t Y_t - (X^D_t X^D_t)^{-1} X^D_t X^S_t S$, |
| $t = 1, 2, \ldots, M$, $Z_t = I_L - X^D_t (X^D_t X^D_t)^{-1} X^D_t$ and $I_L$ is $L \times L$ identity matrix. |
| **Output:** Static and dynamic coefficients. |

5.2.2 Eigenconnectivity Network Extraction

Interpreting the learned group/dynamic brain connectivity networks is challenging due to the temporal-spatial complexity of the connectivity patterns. To facilitate such interpretations, in this study, eigenconnectivity networks are extracted based on the concatenated static and dynamic connectivity networks using the decomposition approach [80]. They serve as the representative spatial connectivity patterns, which can then be compared across groups.

In order to estimate the eigenconnectivity networks, we first vectorize and concatenate the dynamic connectivity networks. Suppose the number of ROIs is $K$, the learned connectivity network for one subject at one window is a square matrix with size $K \times K$ which can be converted to a vector with length $K(K - 1)$ (the diagonal elements are all zeros, and thus can be ignored). Suppose the number of control subjects is $P_1$, the number of PD subjects is $P_2$, the total number of subjects is $P = P_1 + P_2$, and the number of windows for each subject is $M$. Then the vectors for all the subjects across all time windows are concatenated to form an observation matrix $O$ with size $K(K - 1) \times PM$. 

85
Then Principal Component Analysis (PCA), one popular blind source separation method, is applied to the formed observation matrix. It assumes that the observations are the mixtures of the underlying orthogonal sources and PCA tries to project the original data to a space with uncorrelated variables. With applying PCA, we have

\[ O = A \times W', \]  

(5.10)

where \( A \) contains the principal components, each column yielding the eigenconnectivity after reshaping to the \( K \times K \) matrix. \( W \) is the loading weight matrix with each column corresponding to one component. The elements of \( W \) denote the contributions to the estimated eigenconnectivity. To estimate \( A \) and \( W \), Singular value decomposition (SVD) is commonly employed. In real applications, we are often interested in a small number of eigenconnectivities which are able to explain most of the variations in the group/dynamic observation matrix. More details could be found in [80].

In this study, we estimate the eigenconnectivity networks using the static and dynamic connections respectively. We investigate the corresponding eigenconnectivities that contain 75% variance, and we further compare the differences of the averaged contributions between control and PD groups.

### 5.2.3 Dynamic Feature Extraction

To further compare the temporal variability of the estimated networks, the network variation is adopted as described in Section 4.2.3,

\[
V = \frac{1}{M-1} \sum_{t=2}^{M} \|GD(t) - GD(t-1)\|_F^2,
\]  

(5.11)

where \( M \) is the number of windows, \( t \) represents the window index and \( GD(t) \) represents the dynamic brain connectivity network which is a matrix estimated using the proposed approach as described in Section 5.2.1. Network variation calculates the average distance between two brain connectivity networks at adjacent time windows which is an efficient way to measure the variability of the networks across the time. In our real data application, we will compare network variations between
different groups.

5.3 Simulations

To validate the proposed method, we performed simulations on the combined static and dynamic model and compared its performance with that of other regression approaches. In [70], it has been shown that the dirty model outperformed both the Lasso and group Lasso models. Here we focus on the estimation performance improvement of the proposed approach in the time-varying setting.

The data were generated from a Gaussian model with static and dynamic structures as $Y_t = X_t(S + D_t) + e_t$ where $t$ represented the time index. $S$ was the time invariant coefficients and $D_t$ was the time dependent coefficients. In detail, we first randomly generated the static coefficients $S$ and the time varying coefficients $D_t$ with $P$ variables. The static features $S$ remained the same at all time points and the time dependent coefficients $D_t$ changed with the time. More specifically, the non-zero dynamic coefficients were assumed to be present for several successive time points which was the time scale of the changing coefficients (as shown in Fig. 5.1 (b)). In the simulation, the time scale of changing coefficients was $N$ and the total sample length was $T = 3 \times N$. The design matrix $X$ was randomly generated, containing $T$ observations and $P$ variables. The response vector $Y$ was calculated by $Y_t = X_t(S + D_t) + e_t$ with $t = 1, \cdots, T$, and $e_t$ was the Gaussian white noise.

In the first simulation, $P$ was 50, the sparsity of the coefficients was set to be 0.16, and the time scale $N$ was 50. We compared the performance of different methods with changing static feature ratio $r$ (the ratio of static features to all non-zero features) from 0.3 to 0.8. One example of the generated model was shown in Fig. 5.1 (a)-(c). The second simulation tested the performance of the algorithms as a function of the number of time scales $N$. $P$ was set to be 20, the sparsity of the coefficients was 0.25, and the static feature ratio was 0.6.

We compared the estimation performance of the proposed approach with the static lasso (SL) [141] time varying group lasso (TVGL) and time varying lasso (TVL) models. TVGL and TVL extended the original group lasso [157] and lasso model [141] to the time-varying setting based on the sliding window approach as described in Section 5.2. The SL model assumes that the coefficients are the
Figure 5.1: Simulation example. (a). True static coefficients. (b). True dynamic coefficients. (c). True combined coefficients. (d). Estimated static coefficients using the proposed CSDM. (e). Estimated dynamic coefficients using CSDM. (f). Estimated combined coefficients using CSDM. (g). Estimated coefficients using the static lasso model. (h). Estimated coefficients using the time varying group lasso model. (i). Estimated coefficients using the time varying lasso model. The time index is along the x-axis, the variable index is along the y-axis and the color bar represents coefficients’ strength.

same across the time and all time points are used together to estimate the sparse coefficients. TVGL assumes that the temporal features are similar across the time and applies the group estimation to all windows. Finally, the TVL model estimates
Figure 5.2: The comparison of the L2-loss of the estimated coefficients for (a) simulation 1 and (b) simulation 2. Red, black, green, pink lines represent the proposed CSDM, static Lasso [141] time varying group Lasso (modified based on [157]) and time varying Lasso (modified based on [141]) respectively.

the coefficients at each time window separately. The parameters were selected according to the sparsity of the estimated structures for all models.

One example of the estimated results of different approaches in the first simulation was shown in Fig. 5.1. It demonstrated that the proposed CSDM was able to estimate the static features precisely and also yielded relatively good estimation ac-

89
accuracy of dynamic coefficients. In the simulation, $L_2$ loss of estimated coefficients of different models were compared. We tested the performance of the algorithms as a function of static feature ratio $r$ in the first simulation. We have repeated each procedure fifty times and the averaged performance have been compared as shown in Fig. 5.2 (a). It was noted that with the increasing of the static feature ratio $r$, all the models produced better performances. It may due to the fact that the static features were contained in all windows while the dynamic features could only be detected with limited time epochs. Thus, with the increasing number of static features, all the methods produced higher estimation accuracy. In the second simulation as shown in Fig. 5.2 (b), it was noted that the time scale didn't affect the performance of static lasso model much. However, the estimation precision of all other methods was improved with increasing number of time scale $N$. The CSDM model always obtained lower estimation loss as shown in Fig. 5.2.

It’s worth mentioning that our proposed approach was able to detect the static features and dynamic features simultaneously which was promising in describing the mixed static and dynamic models. For other methods, the post-hoc testing was necessary to distinguish the static and dynamic features.

### 5.4 Real Application

In this section, we apply the proposed method to a real resting state fMRI data set and examine the static as well as dynamic properties of brain connectivity networks in subjects with Parkinson’s Disease (PD). We first estimate the brain connectivity networks for each subject and then compare the temporal patterns of the inferred connectivity networks. We are interested in exploring the dynamic and static features in PD study.

#### 5.4.1 Subjects and fMRI Resting State Data Set

In our study, twelve subjects with PD and ten age-matched controls were recruited. Resting state fMRI data were collected and PD subjects participated the experiment before and after receiving L-dopa medication (denoted as PDon and PDoff). All experiments were approved by the appropriate University Ethics board.

The parameter settings of data collection was the same with those of the study
Table 5.2: The index and name of 54 selected brain ROIs. 'L' represents the brain left side and 'R' represents the brain right side.

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>ID</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ctx-L-G-frontmiddle</td>
<td>28</td>
<td>ctx-R-G-frontmiddle</td>
</tr>
<tr>
<td>2</td>
<td>ctx-L-ventral-lat</td>
<td>29</td>
<td>ctx-R-ventral-lat</td>
</tr>
<tr>
<td></td>
<td>-prefrontalcortex</td>
<td></td>
<td>-prefrontalcortex</td>
</tr>
<tr>
<td>3</td>
<td>ctx-L-insula</td>
<td>30</td>
<td>ctx-R-insula</td>
</tr>
<tr>
<td>4</td>
<td>ctx-L-superiortemporal</td>
<td>31</td>
<td>ctx-R-superiortemporal</td>
</tr>
<tr>
<td>5</td>
<td>ctx-L-middletemporal</td>
<td>32</td>
<td>ctx-R-middletemporal</td>
</tr>
<tr>
<td>6</td>
<td>ctx-L-inferiortemporal</td>
<td>33</td>
<td>ctx-R-inferiortemporal</td>
</tr>
<tr>
<td>7</td>
<td>ctx-L-parahippocampal</td>
<td>34</td>
<td>ctx-R-parahippocampal</td>
</tr>
<tr>
<td>8</td>
<td>L-Hippocampus</td>
<td>35</td>
<td>R-Hippocampus</td>
</tr>
<tr>
<td>9</td>
<td>L-Somatosensory</td>
<td>36</td>
<td>R-Somatosensory</td>
</tr>
<tr>
<td>10</td>
<td>ctx-L-superiorparietal</td>
<td>37</td>
<td>ctx-R-superiorparietal</td>
</tr>
<tr>
<td>11</td>
<td>ctx-L-inferiorparietal</td>
<td>38</td>
<td>ctx-R-inferiorparietal</td>
</tr>
<tr>
<td>12</td>
<td>ctx-L-occipital-parietal</td>
<td>39</td>
<td>ctx-R-occipital-parietal</td>
</tr>
<tr>
<td></td>
<td>-visual-assoc-area</td>
<td></td>
<td>-visual-assoc-area</td>
</tr>
<tr>
<td>13</td>
<td>ctx-L-lateraloccipital</td>
<td>40</td>
<td>ctx-R-lateraloccipital</td>
</tr>
<tr>
<td>14</td>
<td>ctx-L-caudalanterior</td>
<td>41</td>
<td>ctx-R-caudalanterior</td>
</tr>
<tr>
<td></td>
<td>-cingulate</td>
<td></td>
<td>-cingulate</td>
</tr>
<tr>
<td>15</td>
<td>ctx-L-posteriorcingulate</td>
<td>42</td>
<td>ctx-R-posteriorcingulate</td>
</tr>
<tr>
<td>16</td>
<td>ctx-L-precuneus</td>
<td>43</td>
<td>ctx-R-precuneus</td>
</tr>
<tr>
<td>17</td>
<td>ctx-L-vent-medial</td>
<td>44</td>
<td>ctx-R-vent-medial</td>
</tr>
<tr>
<td></td>
<td>-prefrontal-cortex</td>
<td></td>
<td>-prefrontal-cortex</td>
</tr>
<tr>
<td>18</td>
<td>L-Cerebellum</td>
<td>45</td>
<td>R-Cerebellum</td>
</tr>
<tr>
<td>19</td>
<td>L-M1</td>
<td>46</td>
<td>R-M1</td>
</tr>
<tr>
<td>20</td>
<td>L-SMA-proper</td>
<td>47</td>
<td>R-SMA-proper</td>
</tr>
<tr>
<td>21</td>
<td>L-Pre-SMA</td>
<td>48</td>
<td>R-Pre-SMA</td>
</tr>
<tr>
<td>22</td>
<td>L-PMd</td>
<td>49</td>
<td>R-PMd</td>
</tr>
<tr>
<td>23</td>
<td>L-PMv</td>
<td>50</td>
<td>R-PMv</td>
</tr>
<tr>
<td>24</td>
<td>L-Thalamus-Proper</td>
<td>51</td>
<td>R-Thalamus-Proper</td>
</tr>
<tr>
<td>25</td>
<td>L-Caudate</td>
<td>52</td>
<td>R-Caudate</td>
</tr>
<tr>
<td>26</td>
<td>L-Putamen</td>
<td>53</td>
<td>R-Putamen</td>
</tr>
<tr>
<td>27</td>
<td>L-Pallidum</td>
<td>54</td>
<td>R-Pallidum</td>
</tr>
</tbody>
</table>

in Section 4.4 The duration of each functional run was 8 mins and fifty-four Freesurfer-derived brain ROIs were chosen to learn the brain connectivity networks as shown in Table 5.2
5.4.2 Results

Figure 5.3: Examples of static eigenconnectivities (fixed density). (a)-(d).
Static eigenconnectivity network 1, 2, 3, and 11 respectively.

To fairly compare the connectivity patterns of PD and control groups, we choose fixed densities (sparsity = 0.17)/parameters (\(\lambda = 1900, \gamma = 20\)) for all the subjects in this study. The length of the sliding window was set to be 128s (64 TRs) in order to avoid large fluctuations.

Eigenconnectivities are adopted to investigate the spatial patterns of the estimated time invariant and time evolving connections. For the fixed sparsity setting, 11 static eigenconnectivity networks and 161 dynamic eigenconnectivity networks were identified respectively, explaining 75% variance (Fig. 5.3 and Fig. 5.4).
As shown in Fig. 5.3, the cross-hemisphere bilateral ROIs were strongly maintained in all the static eigenconnectivities which was consistent with prior medical knowledge. However, for the dynamic features in Fig. 5.4, the variability of the connectivity patterns was very large. It was noted that the first dynamic eigenconnectivity had clear bilateral connections (Fig. 5.4(a)) while others had quite different spatial patterns. Similar eigenconnectivity patterns were also identified in the fixed parameters cases.

We further examined the corresponding averaged loading weights, denoting the group contributions to each eigenconnectivity. An unpaired T test was applied to find the eigenconnectivities with significant different averaged time de-
Figure 5.5: (a) The static eigenconnectivity network 2 (fixed density). (b). The comparison of averaged contributions of control and PD groups to the static eigenconnectivity network in (a). (c) The static eigenconnectivity network 3 (fixed parameter). (d). The comparison of averaged contributions of control and PD groups to the static eigenconnectivity network in (c).

As shown in Fig. 5.5 (a) and (b), for the fixed density case, the time dependent weights for second static eigenconnectivity network was significantly different between control and PD groups (P value = 0.0327). The different contributions from two groups may indicate the connectivity network alterations due to PD. When comparing the loading weights of 161 dynamic eigenconnectivity networks, 8 of
them were found to be significant different between groups. The first significant different dynamic eigenconnectivity network for fixed density case has been shown in Fig. 5.6 (a) and (b).

We noted that eigenconnectivity network patterns based on static and dynamic features were quite different. The static connectivity networks preserved the most robust connections while the dynamic connectivity patterns may provide insights.
Figure 5.7: The comparison of network variations between control, PDon, and PDoff groups with either (a) fixed density (density = 0.17) or (b) fixed parameters ($\lambda = 1900, \gamma = 20$).

5.5 Conclusion and Discussion

Studying dynamic properties of brain connectivity networks can extend our understanding of brain functioning. It provides a promising way to assess the evolving of brain organizations. The influence of tasks and external stimulus can also be ex-
explored which is crucial in evaluating the brain during adaption. The brain connectivity dynamics also open new avenues to the exploration of the neurodegenerative disorders. Compared with static connectivity, time varying connectivity patterns may provide a more sensitive and specific biomarker of disease [17].

However, purely assuming that brain fluctuations are static or dynamic may be oversimplified. Previous studies have suggested that some central couplings dominating the brain coordination are largely invariant across specific cognitive states [77]. Relative stable connections may be present with the transient couplings existing. Thus, a mixed temporal connectivity model with both the time invariant and time related connections is likely required.

In this chapter, a combined static and dynamic brain connectivity network modeling approach has been proposed. Compared with previous multivariate time varying approaches introduced for fMRI brain network estimation, our method is more flexible which is able to identify the time invariant and time dependent connectivity features simultaneously.

The proposed method has been applied to a real fMRI resting state data set with Parkinson’s Disease. The estimated static and dynamic brain connectivity networks are found to be quite different which represent the complementary information of the brain connectivity patterns. In addition, the PD subjects have significantly reduced network variation, likely related to impaired cognitive flexibility.

There are still some challenges in studying the brain dynamics. Although a lot of studies have been conducted on the time dependent brain network estimation, there is no consensus on the underlying brain connectivity patterns. Whether the brain coupling changes smoothly or suddenly, is still unclear. What is the time scale of the brain connectivity networks, is still in debate. The multimodality studies such as EEG/fMRI combined studies may be crucial in our future work to explore the connectivity evolving at high temporal resolution [17,145].

While the proposed method appears interesting, there are still a number of limitations. We have to select the parameters for the modeling. Although we have conducted the comparisons of brain connectivity patterns with various window lengths and parameters, similar conclusions have been made. Further work would be required to more comprehensively investigate time varying brain connectivity patterns over a broad range of parameter settings. In this chapter, we have adopted
an eigenconnectivity network estimation to compare the static and dynamic connectivity patterns. It is a promising way to extract the representative fluctuation patterns. However, resting state time varying connectivity patterns are particularly challenging in interpretation due to its spatial and temporal complexity. More measures may be required to describe the full potential of temporal and spatial information which could be expanded in future work.

This is our first attempt to model the static and dynamic connectivity patterns in the resting state simultaneously. Jointly estimation of time variant and invariant connectivity patterns may be important to fully understand the brain robustness and efficiency.
Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, we have developed a set of novel network modeling approaches for brain connectivity estimation using fMRI signals. The proposed algorithms try to address several challenges present in real brain connectivity studies including the group inference for multiple subjects, error rate control, prior knowledge incorporation and dynamic brain connectivity network modeling. All the proposed methods have been applied to real fMRI datasets, with the ultimate goal to investigate the alterations in brain connectivity patterns associated with neurological diseases. The results demonstrate that connectivity features may sever as biomarkers for future studies. The main contributions and findings of this thesis are summarized as follows.

In chapter 2, a false discovery rate controlled, prior knowledge incorporated group level network modeling approach has been introduced by extending the PC\textsubscript{fdr} algorithm to the group level and imposing the prior knowledge on the network structures. The proposed approach was able to control the FDR directly at the group level. It was an efficient and reliable group level inference approach. When applied to fMRI data collected from subjects with Parkinson’s Disease on and off L-dopa medication and normal controls performing a motor task, we found robust group evidence of disease-related changes, compensatory changes and the normalizing effect of L-dopa medication. Our proposed approach was able to detect the
reliable group level connectivity features which made the group comparisons possible.

When incorporating the prior knowledge from other modality and modeling the diversity in the brain connectivity networks instead of conventionally treating all the subjects equally, a genetically informed group brain connectivity network modeling approach was proposed in Chapter 3. While genetic information was used in our study, other prior information such as clinical indices could also be adopted to assist the brain connectivity diversity modeling. The proposed framework was more flexible in dealing with the inter-subject variability in a population, and it efficiently integrated the information from other modalities which greatly assisted the final biological interpretations. Using the proposed approach, we demonstrated the changes in connectivity patterns under the effect of a specific genetic variation and how this was modulated by Schizophrenia.

Modeling the evolution of brain connectivity networks along the time is of great importance due to the fact that the brain is inherently non-stationary. Studying dynamic properties of brain connectivity could extend our understanding of brain functioning. In Chapter 4 a sticky weighted time varying model was developed to investigate the time-dependent brain connectivity networks based on the fused regression model. The proposed method was able to recover both smooth changing coefficients and abrupt changing structures. In addition, the proposed method allowed different pairs of brain regions to exhibit fluctuations at different time scales which made the model more flexible. When applied to a real fMRI resting state study in Parkinson’s Disease, PD subjects were found to have significantly reduced network variation, likely related to impaired cognitive flexibility in PD. The dynamics of brain connectivity patterns have thus provided us more insights into the neurological disorders.

In Chapter 5 we further leveraged assumption on the dynamic brain connectivity networks. Instead of purely assuming static or dynamic connectivity, we proposed learning both time-invariant connections and time-varying coupling patterns simultaneously by employing a multitask learning model followed by a least square approach to precisely estimate the connectivity coefficients. It was our first attempt to jointly model the static and dynamic connectivity patterns in the resting state. It served as a potential tool to fully describe the brain robustness and flexi-
ibility. The proposed method was applied to a real fMRI resting state data set in Parkinson’s Disease. The estimated static and dynamic brain connectivity networks were found to be quite different which represented the complementary information of the brain connectivity patterns. In addition, the PD subjects had significantly reduced network variation which was consistent with our previous study.

In summary, the major contributions of this thesis work are developing a collection of novel brain connectivity modeling approaches that are able to deal with particular challenges present in real fMRI applications. The group comparisons between subjects in control and disease states based on the proposed group level inference have examined the effects of L-dopa medication and enhanced the understanding of disease induced changes on brain connectivity patterns. The second work expanded the imaging genetic studies and provided a new strategy for prior knowledge incorporation into the brain network modeling. The proposed time varying brain connectivity modeling approaches in this thesis assessed the brain functions in the additional temporal dimension which would be promising in the studies of neurological disorders. The results suggest that certain connectivity patterns could sever as the biomarkers and deserve further investigations in the future studies.

6.2 Future Work

Although a lot of methods have been developed in this area, it’s worth noting that each method has its own advantages and limitations, and there are still many more aspects of brain connectivity that need to be further studied. Increasing number of combined methods have been developed which benefit from each other’s complementary advantages. In the following, we will discuss on some possible future directions regarding brain connectivity network estimation.

6.2.1 Brain Connectivity Transition Patterns Estimation

As in our aforementioned discussions, time-dependent brain connectivity has provided a new perspective on brain adaption and reconfiguration. In particular, the dynamics of brain connectivity may be altered in disease states which indicate the importance to investigate the time varying brain connectivity networks [80, 90].
Most existing approaches examine the temporal changes of intrinsic brain connectivity patterns in the resting state. While a few studies have demonstrated that connectivity can be mediated by learning and/or task performance [7, 42, 81, 124]. Therefore, it would be interesting to investigate the reconfiguration of brain connectivity under different cognition states. For instance, the transition patterns from task state to resting state may be vital to describe the brain functioning, and potential to investigate a variety of neurological disorders.

Considering the success of decomposition methods in the area of static brain connectivity network modeling, a possible strategy for state dependent brain network estimation is HMM-IV A (Hidden Markov Models-Independent Vector Analysis) model which is a joint multidimensional multi-state network estimation approach, aiming to extract the state dependent common sources from several data sets simultaneously. Independent Vector Analysis (IVA) is a joint blind source separation method which has been successfully introduced to simultaneously extract the underlying common sources across multiple data sets. Combined with HMM, we could extend it to the dynamic settings where connectivity transition patterns under different conditions are able to identified.

### 6.2.2 Large Scale Brain Connectivity Network Modeling

Due to the scarcity of fMRI samples, current graphical models of brain connectivity estimation usually involve dozens or hundreds of nodes. However, fMRI simultaneously measures around 100,000 voxels which have been traditionally averaged within each ROI. Designing algorithms which are able to recover networks of thousands of variables will be of great interest for brain connectivity modeling.

To recover a large scale brain connectivity networks, the algorithms with high computational efficiency are demanded. The high dimension precision matrix inference is one possible method for estimating large scale voxel-wise brain interactions [29, 93].

The multi-scaling approach is another potential strategy to efficiently estimate the brain connectivity networks hierarchically. The spatial information could be incorporated into the process of network modeling at the coarser and fine levels.
The relationship between small regions contained within certain functional areas could be firstly estimated and the representative signals are extracted at the fine scale. Then the interactions between larger ROIs with the representative signals are studied at a coarser level.

### 6.2.3 Connectivity Based Brain Voxel Selection

Two straightforward approaches to assess brain connectivity networks are voxel-based and ROI-based methods. Voxel-based approaches usually involve large number of variables. In general, the correlation threshold method is adopted to study the correlation fMRI maps with thousands of voxels. However, graphical models which are not able to handle a large number of variables favour more the ROI-based connectivity modeling approaches with a small number of variables involved. In addition, the voxel-based approaches are more sensitive to the noise, which make the results unstable. As a result, ROI-based approaches are usually adopted. However, we have to carefully define the ROIs with accurate voxels selected. Otherwise, the signal fluctuations within a single ROI may be the result of the influence of multiple underlying brain networks.

The parcellation of brain is a challenging problem. The brain regions could be defined based on their anatomical structures or functional specializations [138]. However, labelling of ROIs is a time consuming task and the accurate definition of some region boundaries is still in debate. To relax the requirements of the prior knowledge, several data driven methods have been developed, such as clustering approach [36] and seed region growing approach [94].

Previous research has demonstrated that the properties of brain connectivity networks heavily depend on the parcellation of brain ROIs [152]. For the purpose of connectivity estimation, ROI definition approaches should be able to select functional and spatially homogeneous voxels within each region. The connectivity based ROI parcellation methods have been of great interest. Applying the modularity graph measure on the functional connectivity network, Barnes et.al try to detect the optimal divisions in basal ganglia [5]. Spatially regularized canonical correlation analysis is used to define the ROIs as a set of voxels with similar connectivity patterns to other ROIs [39]. In another interesting study, classifier based approach
is adopted to identify brain ROIs with consistent genotype-phenotype relationships [127].

In our preliminary study, the spatial information is taken into consideration in the brain region definition by a sparse spatially regularized fused lasso regression model. The target brain ROI is defined as the functional consistent and spatially adjacent area [158]. We plan to further develop a framework which is able to conduct whole brain ROIs definition by combining computational-efficient connectivity estimation models with the spatial constraints. To represent each ROI, different strategies could be adopted, such as averaging over all voxels, principal component and other distance measures. It would be interesting to compare the impact of those representing signals on the properties of brain connectivity networks. In addition, using the temporal information as another criterion to segment the functional regions may be a novel way for brain parcellation where the evolutionary clustering approach is a potential method.

6.2.4 Application to Parkinson’s Disease Studies

Parkinson’s Disease is one of the most common movement disorders which is characterized by muscle rigidity, tremor, bradykinesia and etc. In addition, the cognitive and psychiatric changes often accompany with PD. The high prevalence and serious consequences of PD have enormous impact on the whole population. The diagnosis of Parkinson’s Disease heavily depends on the medical history and clinical symptoms. However, since no lab test clearly identifies PD, diagnosing PD at the early stage from other parkinsonian diseases at a sufficiently accurate level is still challenging [2]. As one of the non-invasive neuroimaging technologies, fMRI has been widely used for studying brain activities in PD. [126].

Previous studies have suggested that PD may lead to the alterations of brain connectivity patterns [4, 18, 155, 156]. For instance, the functional connectivity of the motor network in the resting state was found to be disrupted in PD and related with the severity of the disease [155]. In another study, effective connectivity in resting state was employed to assess the effects of deep bran stimulation and demonstrated the distributed effects of stimulation on the resting motor networks [73]. Modeling and comparing brain connectivity networks in health and PD could
thus provide a way to evaluate the disease related connectivity abnormalities. They may further serve as potential tools for the diagnosis of PD and for quantifying the severity of PD.

We have conducted some preliminary studies on Parkinson’s Disease and found interesting results. In the future work, we plan to comprehensively study the brain connectivity networks in PD, in order to characterize the spatial and temporal properties associated with PD, and to extract the particular connectivity patterns of the patient population. The multimodal assessment of PD is another potential direction to jointly probe the disease and to strengthen the biomarkers form combined studies [98].
Bibliography


116


B. Ng. Prior-informed multivariate models for functional magnetic resonance imaging. 2011. → pages 3


117


120


