Bayesian Inference of Directional Brain Network Models for Intracranial EEG Data

by

Yinge Sun

A Dissertation Presented to the Graduate Faculty
of University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Statistics

University of Virginia

April 2020

Committee Members: Tingting Zhang

Jianhui Zhou

Guofen Yan

Jeff Changchia Liu

Abstract

The human brain is a network system in which brain regions, as network nodes, constantly interact with each other. The directional effect exerted by one brain component on others is referred to as directional connectivity. Since the brain is also a continuous-time dynamic system, it is natural to use ordinary differential equations (ODEs) to model directional connections among brain regions. We propose a high-dimensional ODE model to explore directional connectivity among many small brain regions recorded by intracranial EEG (iEEG). The new ODE model, motivated by the physical mechanism of the damped harmonic oscillator, is effective for approximating neural oscillation, a rhythmic or repetitive neural activity involved in many important brain functions. To produce scientifically interpretable network results, we assume the sparse structure for the ODE model parameters that quantify directional connectivity among regions. We consider two types of sparse structure: 1. a modular network structure consisting of several functionally independent subnetworks/clusters of lower dimensions which provides an intuitive interpretation of functional specialization of brain regions in different clusters, 2. a small-world network structure consisting of several subnetworks with dense connections within the same cluster and sparse connections between different clusters which reflects two principles of the brain's functional organization: functional integration and segregation, resulting in two ODE models. We develop two Bayesian methods to estimate the model parameters of the proposed ODE models and to identify clusters of strongly connected brain regions. We compare the two proposed ODE models through simulation studies and analysis of iEEG data collected from a patient with medically intractable epilepsy and examine the patient's brain networks around the seizure onset.

Acknowledgements

I would like to thank all people who have helped and supported me during my PhD period of study at University of Virginia (UVa). I would never have been able to finish my coursework and dissertation without the guidance of faculty members at UVa Department of Statistics and my committee members.

I would like to express my very great appreciation to Prof. Tingting Zhang, my advisor, for her attentive guidance, encouragement, patience, caring and valuable critiques of my research work. Her sincere enthusiasm, dedication and earnest manner in research had been always inspiring me. Every time I encountered tough problems in research, she was always accessible and able to provide constructive suggestions, which constantly broadened my mind. In addition to her professional advice for my research work, she was always willing to share her opinions on what she had learned from her past experience in all aspects. She patiently trained my logical thinking and taught me how to deal with problems in both life and work in a suitable and effective way. And she always encouraged me and provided me opportunities to communicate with people from diverse background and learn new things to broaden my horizons. This is truly helpful for my future career and life. And for sure, my life at UVa would not be so smooth and rewarding without her valuable advice.

Besides, I am particularly grateful to Prof. Jianhui Zhou and Prof. Chao Du for their excellent teaching and insightful comments. They were always able to find a way to make the obscure concepts in statistics easy to understand and help set a clear direction for me to follow. And I always got prompt reply and patient guidance whenever I had a problem. Thanks for being responsible, professional and kind.

I would also like to thank my thesis committee members, Prof. Tingting Zhang,

Prof. Jianhui Zhou, Prof. Guofen Yan and Prof. Jeff Liu for their encouragement, brilliant comments and suggestions, and hard questions. It's enjoyable to have a discussion with them on my research work. Besides, I would like to thank my advisor's collaborator, Dr. Seiji Tanabe, for collecting, organizing and explaining data used in my research. My work would be impossible without his help. And thanks my lab members and fellow students in the department. My days would not be so colorful without their encouragement and support.

I would like to offer my special thanks to Mrs. Karen Dalton, who has been helping me in all aspects of life and schoolwork and keeping everything organized in the past five years. She spared no effort in making the department a loving family.

Finally, my deepest gratitude goes to my parents, and my friends at UVa and in China, who have been always there cheering me up and standing by me through the good times and bad.

Contents

Chapter 1

Introduction

1.1 Overview of the Problem

1.1.1 Effective Connectivity in Neuroscience

Functional Segregation and Functional Integration

The human brain is the control center of neural system, constantly performing various functions to sustain life and maintain physical operation. As the most important and the most complex organ of human body, the human brain is responsible for processing information from the sense organs and sending signals to the rest of the body through neuron cells. In the process, it has been discovered that different areas of the brain are specialized to particular functions, such as temporal lobe for memory (Squire and Zola-Morgan, 1991), parietal lobe for motor (Fogassi and Luppino, 2005), parietal lobe and frontal lobe jointly for space perception (Rizzolatti *et al.*, 1997). As a matter of fact, this reveals the structural/anatomical network characteristics of human brain, where the cortical areas at similar distance are densely connected by synapses or fiber pathways to form a network hub and perform specific

functions, and those hubs interact globally in the network to complete a complex task. This pattern of locally dense connection and globally coherent interaction in the structural brain network reflects two fundamental principles of human brain organization, i.e. functional segregation and functional integration. Figure 1.1 shows the partition of cerebral cortex.

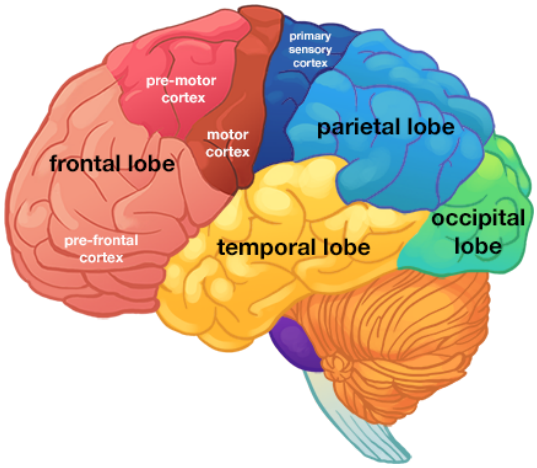


Figure 1.1: The cerebrum is divided into four lobes: frontal, parietal, occipital and temporal. Huntinton's Outreach Project for Education at Stanford. `https://hopes.stanford.edu/the-hopes-brain-tutorial-text-version/brain-lobes/`

In the past decades, human brain mapping has focused on functional segregation and the localization of function to a great extent. Functional segregation emphasizes that a cortical area is specialized for some aspects of information processing and this specialization is anatomically segregated within the cortex (Friston, 2011). And many studies has found the distinct functional roles of different cortex areas (Dehaene-Lambertz *et al.*, 2006; Molnar-Szakacs *et al.*, 2004; Moutoussis and Zeki, 1997; Salmelin *et al.*, 1995; Vaina, 1994; Yu *et al.*, 2011). For example, the color and the direction/orientation of lines are processed separately in the human visual perceptive systems. The segregation of information process in distinct brain areas enables specialized functions in localized brain areas, which motivates the modu-

lar network structure consisting of densely interconnected subnetworks (clusters). The clusters in the network corresponding to localized brain areas, provides intuitive interpretation of functional segregation mechanism in human brain systems, which inspired us to build a modularity based model to characterize the functional organization in human brain (Section 2.2.1).

While in recent years, the relationship between functional segregation and functional integration has aroused great attention of people in many different disciplines, with an emphasize on integration of segregated brain areas. In contrast to segregation of function in local areas, functional integration means functions from segregated areas are globally integrated, e.g., many distinct brain areas can be involved during perception and behavior. Study of the balance of functional segregation and integration in the distributed system, i.e. human brain, helps us better understand how the different brain areas works together in multiple tasks (Tononi *et al.*, 1994). It also will advance our understanding of neuropsychiatric disorders, as well as their diagnosis and treatment.

Functional Connectivity and Effective Connectivity

Functional connectivity and effective connectivity are two approaches to investigate the integration of segregated functions in human brain. *The temporal correlation between spatially remote neurophysiological events* (Friston *et al.*, 1993a) is referred to as functional connectivity in neuroscience, which can be measured in a variety of ways such as cross-correlations, coherence and information-theory based methods (Pereda *et al.*, 2005). However, the correlation-based functional connectivity cannot distinguish causes from different sources. For example, in multiunit electrode recordings, correlations can result from a common stimulus input or stimulus-induced oscillations mediated by synaptic connections (Gerstein and Perkel, 1969)

and effect of the two sources cannot be separated in correlations. Thus, to better characterize the functional integration in such a distributed system, we move to study of effective connectivity.

Effective connectivity is explicitly defined as the *influence one neural system exerts over another* (Friston *et al.*, 1993b). To make it more clear, the distinction between functional connectivity and effective connectivity lies principally in the directionality of influence. Effective connectivity models the direct interaction between neurons (with directionality) and thus can infer causality to some extent, while functional connectivity is observable measurements of statistical dependencies (undirected/symmetric).

1.1.2 Data Acquisition of Human Brain

With the development of technology in neuroscience, various techniques have been applied to generate human brain data with different properties, which can characterize different aspects of brain functions. In the literature, both neuroimaging and electrophysiology are popular techniques to study the effective connectivity among brain regions (Friston, 1994; Greenblatt *et al.*, 2012; Grefkes and Fink, 2011; He *et al.*, 2011; Massimini *et al.*, 2005; Rajapakse and Zhou, 2007; Sakkalis, 2011; Valdes-Sosa *et al.*, 2011; Wu *et al.*, 2011).

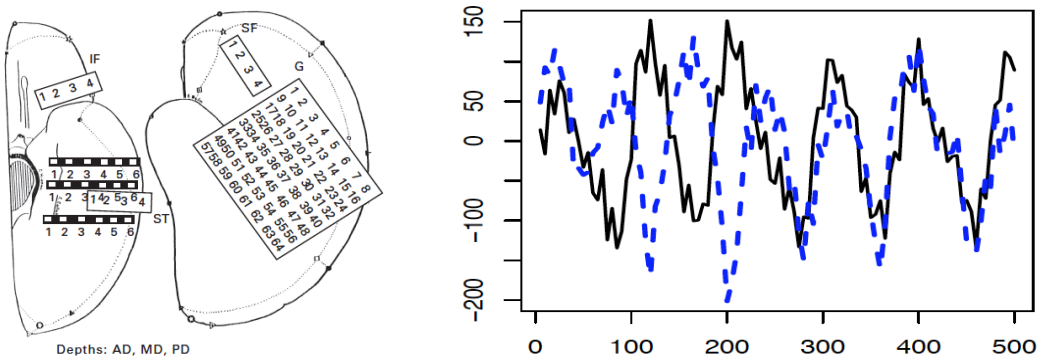
Functional magnetic resonance imaging (fMRI) has stood out from numerous neuroimaging techniques in neurophysiological research since 1990s because of its noninvasiveness, which accelerates people's cognition of functional segregation and integration in human brain. fMRI measures brain activity indirectly through blood oxygen level dependent (BOLD) contrast (Huettel *et al.*, 2004), relying on the fact that cerebral blood flow and neuronal activation are coupled (Logothetis *et al.*,

2001). Despite its noninvasiveness, fMRI has serious limitations in studying the effective connectivity network of human brain due to the data properties. Specifically, instead of directly measuring the neuronal activities, fMRI records the BOLD signal with low temporal resolution (repetition time of two sequential scans at second scale) and a low signal-to-noise ratio (SNR, Welvaert and Rosseel, 2013) due to thermal and physiological (e.g., cardiac and respiratory pulsations, motion) noise. In addition, the latency of BLOD signal relative to immediate neuronal activities is significant and slice by slice scanning across the whole brain makes it difficult to construct a simultaneous effective connectivity network of the brain. Considering that the neuronal activities are instantaneous and the effective connectivity among brain regions is dynamic, all these extra noises introduced cause fMRI data not suitable for effective connectivity analysis.

Electroencephalogram (EEG) is also a noninvasive method, which records the electrical activities of human brain to reflect neuronal activities. Although EEG has excellent temporal resolution (256-512 Hz for clinical recording, up to 20k Hz in some research applications; Wilson and Corlett, 2005), it has poor spatial resolution and a low SNR due to the low conductivity of human skull. Each scalp electrode collects a mixture of signal from the underlying brain sources (Burle *et al.*, 2015) and thus it is hard to locate the origin of neuronal activities.

Considering the limitations mentioned above, we use intracranial electrocorticography (iEEG) data to study the effective connectivity among brain areas. iEEG uses multiple electrodes placed on the exposed surface of the human brain (inside the skull) to record neuronal activities of many small brain regions. Figure 1.2(a) shows the placement of iEEG electrodes on the exposed brain of an epileptic patient under study. The acronyms IF, SF, ST, AD, MD, PD, and G stand for inferior frontal, superior frontal, superior temporal, anterior depth, medial depth, posterior depth,

and grid electrodes. Figure 1.2(b) illustrates two iEEG time series recorded on two regions, respectively. iEEG data have two unique properties. First, each iEEG electrode directly records the neuronal electrical activity in one small region (about 10 mm in diameter) at millisecond scale (200-4k Hz). As such, iEEG data have high spatial and temporal resolutions for brain activities within the recording field. Second, iEEG produces highly reliable and reproducible measurements of brain activity with a strong SNR. Therefore, iEEG data are ideal for examining the brain network.



(a) iEEG Electrode Placement

(b) iEEG Time Series

Figure 1.2: (a) Spatial placement of a subdural grid and several strip electrodes on the patient's left hemisphere. The acronyms IF, SF, ST, AD, MD, PD, and G stand for inferior frontal, superior frontal, superior temporal, anterior depth, medial depth, posterior depth, and grid electrodes. The strips in black and white are depth electrodes used to record activity from deeper brain structures close to the hippocampus. (b) Illustration of a short segment of two iEEG time series of two regions.

1.2 Existing Studies of Effective Connetivity

Many existing studies of the brain's effective connectivity are model-based, and popular models include but are not limited to dynamic causal modeling (DCM, Daunizeau *et al.*, 2011; David and Friston, 2003; David *et al.*, 2006; Friston *et al.*,

2003; Kiebel *et al.*, 2006), Granger causality modeling (GCM, Roebroeck *et al.*, 2005, 2011), and structural equation modeling (SEM, Protzner and McIntosh, 2006). The first model uses ODEs to characterize effective connectivity among only a few large regions of interest, and the latter two models are based on discretization of the brain activity changes over time.

1.2.1 Dynamic Causal Model (DCM)

The dynamic causal model (DCM) is a system which describes the influence of common experimental inputs on d different brain regions and on the coupling/effective connectivity among different brain regions. It models the neural mechanisms underlying the observed brain activity measures so as to explain the effective connectivity changes under different experimental conditions. In the sense, the DCM is a causal model (Pearl, 2000).

The dynamics of hidden states (including neuronal states and hemodynamic states in fMRI, and neuronal states only in EEG, explained later), $\mathbf{x}(t) = (x_1(t), \ldots, x_d(t))'$, are described by ordinary differential equations (ODEs) in DCM:

$$\frac{d\mathbf{x}(t)}{dt} = F(\mathbf{x}(t), u, \boldsymbol{\theta}), \tag{1.1}$$

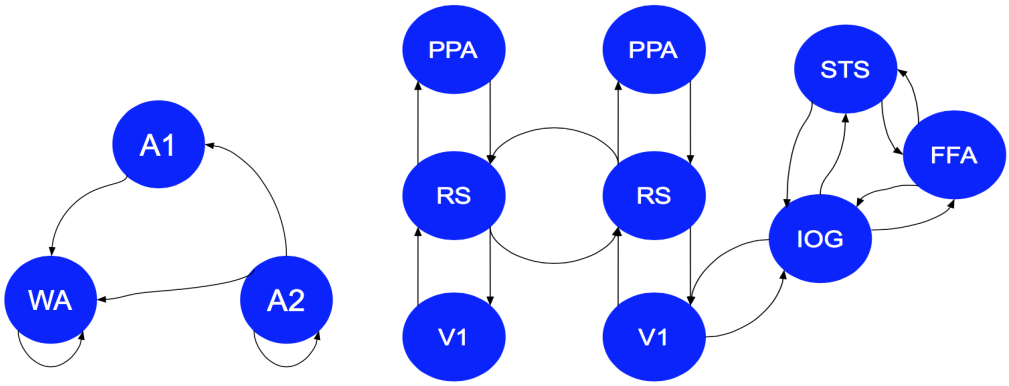
where F is a nonlinear function describing influences of hidden states $\mathbf{x}(t)$ of all d brain regions and experimental inputs $u(t)$ on changes in hidden states $\frac{d\mathbf{x}(t)}{dt}$. $\boldsymbol{\theta}$ are all the parameters in the hidden state model.

Since the hidden states are not observed directly, the DCM also maps $\mathbf{x}(t)$ to the observed measurements $\mathbf{y}(t)$ by an observation equation:

$$\mathbf{y}(t) = H(\mathbf{x}(t), u, \boldsymbol{\varphi}) + \boldsymbol{\epsilon}(t), \tag{1.2}$$

where $\boldsymbol{\varphi}$ denotes all observation parameters and $\boldsymbol{\epsilon}(t)$ accounts for random error and the other nuisance effect.

The DCM can be used to infer effective connectivity among a few large brain regions for fMRI and EEG/MEG data. Typically, there will be no more than 10 brain regions under study (3 for fMRI, Friston *et al.*, 2003, and 9 for EEG, David *et al.*, 2006, see Figure 1.3), and those regions are pre-selected by conventional analysis, which makes the results highly dependent on the prior knowledge of specialized brain regions. And a Bayesian framework is imposed to incorporate constraints on the large amount of parameters in DCM.



(a) Effective Connectivity Network for fMRI

(b) Effective Connectivity Network for EEG

Figure 1.3: Illustration of the effective connectivity network by DCM for both fMRI and EEG data, where the directionality of connection is denoted by arrows between different brain regions: (a) fMRI – A1: primary auditory area, A2: secondary auditory area, WA: Wernicke's area. (b) EEG – V1: primary visual cortex, RS: retrosplenial cortex, PPA: parahippocampal place area, IOG: inferior occipital gyrus, STS: superior temporal sulcus, FFA: fusiform face area. V1, RS, PPA on the left in on the left hemisphere

Specifically, the hidden state model for fMRI combines a simplified bilinear system for neural states, whose parameters captures three aspects of the system: 1. influence of experimental inputs on neuronal states (linear). 2. coupling/effective connectivity among neuronal states (linear). 3. influence of experimental inputs on

coupling (bilinear), and a complex nonlinear hemodynamic model (Friston *et al.*, 2000) linking neuronal states (neuronal activities) to hemodynamic states (hemodynamic responses such as blood volume and deoxyhemoglobin content), which introduces a large amount of free parameters and increases the complexity of the full model. Because fMRI records the BOLD signal instead of neuronal activities directly, the hemodynamic model is necessary here to associate neuronal states with observed measurements through hemodynamic states. As such, the DCM is computationally intensive for high dimensional data.

Similarly, the DCM for EEG incorporates a rather sophisticated neural mass model (David and Friston, 2003) into the hidden state model, which describes the mechanism of membrane potential - current - synaptic firing rate interaction in great detail. Thus, it is not computationally affordable for high dimensional data either.

In addition to the complex modeling, the analysis of brain network effective connectivity in DCM is not straightforward, which is actually done through the Bayesian model comparison and relies on the prior specification of the model parameters.

1.2.2 Granger Causality Model (GCM)

The Granger causality model (GCM) models the directional influence from one brain region to the other in the context of Granger causality (Granger, 1969, 1980) by a vector autoregressive (VAR) process, which does not require prior knowledge of pre-selecting regions of interest. Instead of modeling the coupling per se through a complicated biophysical model like hemodynamical model and neural mass model in DCM, the GCM quantifies the influence that one brain region exerts on the others by the time series' predictive capability of that brain region on the future values in time series of the other brain regions.

The GCM also consists of state equations and observation equations in a state-space formulation. Unlike DCM, the states in GCM may not have definite biophysical meanings because the effective connectivity in GCM is characterized by Granger causality. This leads to a much simpler representation (in the matrix form):

$$\begin{aligned} \mathbf{x}(t) &= \sum_{p=1}^{P} \mathbf{A}_p \cdot \mathbf{x}(t-p) + \boldsymbol{\upsilon}(t) \\ \mathbf{y}(t) &= \mathbf{C} \cdot \mathbf{x}(t) + \boldsymbol{\epsilon}(t), \end{aligned} \tag{1.3}$$

where t takes discrete values in $1, 2, \cdots, T$. The one above is the P-th order VAR state equations, where $\mathbf{x}(t) = (x_1(t), \ldots, x_d(t))'$ are state functions at d different brain regions, $\boldsymbol{\upsilon}(t) = (\upsilon_1(t), \ldots, \upsilon_d(t))'$ are intrinsic noise in the states, and parameter $\mathbf{A}_p$ capture the effective connectivity among all brain regions at time lag p. The one below is the observation equations, where $\mathbf{y}(t) = (y_1(t), \ldots, y_d(t))'$ are observed measurements, $\boldsymbol{\epsilon}(t) = (\epsilon_1(t), \ldots, \epsilon_d(t))'$ accounts for measurement errors, and parameter $\mathbf{C}$ are unknown normalization constants.

Despite the concise form of GCM, study has found that the divergence between Granger causality and true coupling increases with the sampling period (Friston, 2011), which may be relevant to the discretization process of VAR models. To illustrate the problem of discretization process, let's consider the GCM with first-order VAR, then the predictive functions at time lag 2 becomes

$$\mathbf{x}(t+2) = \mathbf{A}^2 \cdot \mathbf{x}(t) + \mathbf{A} \cdot \boldsymbol{\upsilon}(t+1) + \boldsymbol{\upsilon}(t+2), \text{ (see Figure 1.4)}$$

where subscript 1 of matrix $\mathbf{A}$ is omitted for clarity. Suppose the data available is at half frequency of the original data, then the effective connectivity is no longer defined by $\mathbf{A}$ but $\mathbf{A}^2$. As is shown in Figure 1.4, the effective connectivity network

changes for data with different sampling frequencies, where region 2 Granger causes region 1 for data with lower frequency (with a regression coefficient of 0.12). That indicates the directional connections in the underlying GCM always exist with or without true coupling, caused by directed influence between two regions at a finer time-scale than that at which the data are observed. Such problem of discretization is not well addressed in the GCM.

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\cdot\mathbf{x}(t)+\boldsymbol{v}(t+1) \\ \mathbf{x}(t+2) &= \mathbf{A}\cdot\mathbf{x}(t+1)+\boldsymbol{v}(t+2) \end{aligned} \Rightarrow \mathbf{x}(t+2)=\mathbf{A}^2\cdot\mathbf{x}(t)+\mathbf{A}\cdot\boldsymbol{v}(t+1)+\boldsymbol{v}(t+2)$$

$$\mathbf{A}=\begin{pmatrix} -0.5 & 2 & 0 \\ 0.3 & -0.3 & 0 \\ 0.6 & 0 & -0.5 \end{pmatrix} \Rightarrow \mathbf{A}^2=\begin{pmatrix} 0.31 & -0.16 & 0 \\ -0.24 & 0.15 & 0 \\ -0.6000 & 0.12 & 0.25 \end{pmatrix}$$

Figure 1.4: The effect of different sampling frequencies on the effective connectivity network.

1.2.3 Structural Equation Model (SEM)

Structural equation modeling (SEM) avoids the problem of the estimated effective connectivity dependent on sampling frequency in GCM by modeling the dynamic brain system as a static one. Instead of considering state variables at different time points, the SEM focuses on the analysis of the covariance structure of multivariate time series (hidden or observed) since covariance reflects the association between multiple brain regions, i.e., functional connectivity in brain network. More precisely, the SEM compares the observed covariance structure with the one implied by a underlying structural model, where causal connection (effective connectivity) among different brain regions is described by the model parameters. Figure 1.5 is

a graphical display of structural model in SEM, where x_1, x_2, x_3 are state variables (three univariate time series) and v_1, v_2, v_3 are independented random noise affecting the state variables. And directional connections among the state variables, i.e., the effective connectivity, are denoted by the black arrows.

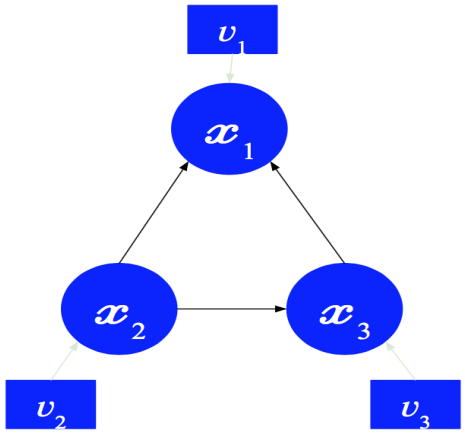


Figure 1.5: The structural model in SEM reveals the directional connection among the hidden state variables, i.e., the effective connectivity network in SEM.

The equations for the structural model are:

$$\mathbf{x} = \mathbf{A} \cdot \mathbf{x} + \boldsymbol{v}. \tag{1.4}$$

And thus,

$$\begin{aligned} \mathbf{x} &= inv(I - \mathbf{A}) \cdot \boldsymbol{v}, \\ \mathbf{x}^T \cdot \mathbf{x} &= (inv(I - \mathbf{A}) \cdot \boldsymbol{v})^T \cdot (inv(I - \mathbf{A}) \cdot \boldsymbol{v}) \\ &= inv(I - \mathbf{A})^T \cdot \mathbf{S} \cdot inv(I - \mathbf{A}), \end{aligned} \tag{1.5}$$

where $\mathbf{A}$ captures the effective connectivity between the state variables and $\mathbf{S} = \boldsymbol{v}^T \cdot \boldsymbol{v}$. $\mathbf{A}$ and $\mathbf{S}$ are model parameters to be estimated. The hidden state variables $\mathbf{x}$ are even not needed to be estimated in SEM.

The parameters in SEM are estimated by minimizing a minus maximum likelihood function (Büchel and Friston, 1997) of the observed covariance $\frac{1}{T-1}\mathbf{y}^T \cdot \mathbf{y}$ (T is the number of time points), denoted as Σ_o, and the implied covariance $\frac{1}{T-1}\mathbf{x}^T \cdot \mathbf{x} = inv(I - \mathbf{A})^T \cdot \mathbf{S} \cdot inv(I - \mathbf{A})$, denoted as Σ_s:

$$log(|\Sigma_o|) - log(|\Sigma_s|) - tr(\Sigma_o \cdot inv(\Sigma_s)) + p \tag{1.6}$$

Although Protzner and McIntosh, 2006 claims that SEM can assess changes in effective connectivity across tasks or between groups, it has serious limitations in characterising the effective connectivity in brain networks. First, by assuming the staticness of brain system, it requires the instantaneous neuronal activities can be captured at multiple time points so that the estimated effective connectivity from SEM is meaningful in interpreting the neural activities. However, even the up-to-date techniques cannot do that, especially for fMRI, which has temporal resolution at second scale. Second, analysis of covariance structure instead of the individual time points may lead to information loss in a great deal while high-dimensional systems actually need more data to be charaterized. As such, it is difficult to generalize SEM into high-dimensional space directly. Third, the reciprocal and cyclic connections cannot be captured efficiently because the conditional dependencies among different brain regions are difficult to be separated in covariance (Friston, 2011).

1.2.4 Dynamic Directional Model (DDM)

Considering the problems mentioned above, the DCM, GCM and SEM cannot be generialized to study the effective connectivity in iEEG data, a high-dimensional system. Many existing high-dimensional network studies of iEEG data have focused on network patterns of pairwise relationships (functional connectivity) between regions

(Burns *et al.*, 2014; Kramer *et al.*, 2008, 2010, 2012; Mormann *et al.*, 2005; Netoff and Schiff, 2002; Schiff *et al.*, 2005; Schindler *et al.*, 2006, 2010, 2008; Wendling *et al.*, 1996; Wu and Gotman, 1998). To date, ODE models for the brain's effective connectivity based on iEEG data have been very limited in the literature. Zhang *et al.* (2015,2017) built high-dimensional ODE models, called Dynamic Directional Model (DDM), for iEEG data. The DDM, consisting of a large number of first-order bilinear ODEs, has a much simpler mathematical formulation and higher dimension than the DCM (David *et al.*, 2006; Friston *et al.*, 2003), does not suffer the sampling frequency problem in GCM and the information loss problem in SEM (Friston, 2011), and can be generally used to explore connectivity pattern among many different regions.

In the DDM, the state equations are characterized by a bilinear approximation of the nonlinear system:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}_1 \cdot \mathbf{x}(t) \cdot (1 - u(t)) + \mathbf{A}_2 \cdot \mathbf{x}(t) \cdot u(t) + \mathbf{B} \cdot u(t) + \mathbf{D}, \tag{1.7}$$

where $\mathbf{A}_1$ represents the directional connection among brain regions without stimulus, $\mathbf{A}_2$ represents the stimulus-dependent directional connection among brain regions, $\mathbf{B}$ represents the effect of the stimulus and $\mathbf{D}$ is the intercept. And the observation equations are simply:

$$\mathbf{y}(t) = \mathbf{x}(t) + \boldsymbol{\epsilon}(t), \tag{1.8}$$

where $\boldsymbol{\epsilon}(t)$ accounts for measurement errors.

To estimate such a high-dimensional system (d is up to dozens or hundreds for iEEG data) with extensive free parameters, a Potts model (Graner and Glazier,

1992; Potts, 1952) has been used to regularize the DDM as well as to characterize the modularity in brain networks (Zhang *et al.*, 2015, 2017), the existence of which has been commonly reported (Newman, 2006). But DDM is not without its limitations, explained later in Section 1.3.1.

1.3 Motivation

The human brain is a continuous-time dynamic system, in which each component (brain region) constantly exert effects over the others. It is biophysically natural to use ordinary differential equations (ODEs) to describe the dynamic mechanism of the brain and use model parameters to quantify the directional/effective connectivity among brain regions. As mentioned earlier in Section 1.2.1, the most popular ODE model for the brain's effective connectivity is the DCM, which characterizes the directional connectivity among only a few large regions (usually no more than 10) based on fMRI and EEG data. The formulation of the DCM is highly complex and relies on the prior knowledge of the existence and directionality of the connections among the regions under study. Since iEEG typically records neuronal activities of more than 50 small regions, among which the relationship is unknown, it is difficult to scale the DCM to iEEG data in terms of both computation and model building.

1.3.1 Oscillatory Activities in iEEG

The DDM (see Section 1.2.4) has a much simpler formulation and can be used to characterize the directional connectivity among many small brain regions. Despite the advantages of the DDM over other existing ODE models, the first-order bilinear ODEs used in the DDM are inefficient to capture the brain's oscillatory activity— periodic or rhythmic up-and-down temporal behaviour of the neuronal

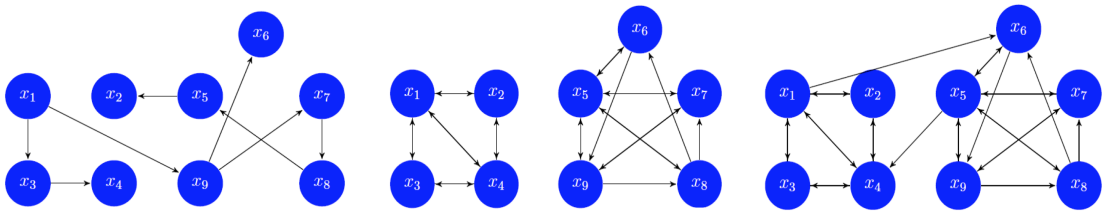
activity, which has been reported to be involved in various important brain functions (Fell and Axmacher, 2011; Fries, 2005; Schnitzler and Gross, 2005). Figure 1.2(b) shows two brain regions' oscillatory activity recorded by iEEG. What the oscillations tell us about the brain has been under intense scientific investigation for almost a century in the neuroscience field. However, the DDM has a great limitation in producing oscillatory time series, because it needs many strong restrictions on its model parameters in order to fit the oscillatory dynamic system, resulting in a serious computational difficulty especially for analyzing high-dimensional iEEG data.

To address the limitations of existing ODE models for iEEG data, we propose a new ODE model. Our model is motivated by damped harmonic oscillator (DHO, Serway and Jewett, 2003), a well-established physical mechanism for oscillatory activity. The DHO is a second-order ODE for one-dimensional oscillatory physical system. Because of its physical implication, the DHO has been used extensively in biophysics (Schuster, 1983) and neuroscience (Daunizeau *et al.*, 2011; David and Friston, 2003; David *et al.*, 2006; Friston *et al.*, 2003; Kiebel *et al.*, 2006) to characterize dynamic systems with oscillatory mechanisms. However, the DHO in these applications are physical models, meaning that they are effective for describing specific systems, but difficult to apply to other systems. As mentioned above, the DDM (Zhang *et al.*, 2015, 2017) and several other ODE models developed (Chen and Wu, 2008; Lu *et al.*, 2011; Wu *et al.*, 2014a,b) are statistical models and have **NO** physical-based components suitable for characterizing oscillatory physical systems. The proposed new method incorporates the physical model DHO into the statistical model. Accordingly, our model provides a substantially better fit to the brain's oscillatory time-series data than the previous statistical models. We refer to this new ODE model as Oscillatory Dynamic Directional Model (ODDM) (details

of the model construction are provided in Section 2).

1.3.2 Sparse Brain Network Structure — Modularity and Small-worldness

As a high-dimensional ODE model, the ODDM contains many free parameters for quantifying the directional connectivity among small brain regions recorded by iEEG (more than 50 at least). In other words, the ODDM parameters imply a high dimensional brain network, where each network node corresponds to one region and each network edge corresponds to a nonzero directional connection between a pair of brain regions. To increase estimation efficiency of ODDM parameters, we use a prior knowledge of brain networks. Specifically, we consider sparse network structure among many possible network structures for high-dimensional brain system. While, high-dimensional, sparse networks can take different forms. Figure 1.6 shows three different sparse network structures, which may imply different functional organizations of the brain. We focus on two types of sparse network structure, the modular network structure and the small-world network structure, based on a scientific ground.



(a) Sparse Connections at Each Node (b) Modularity (c) Small-worldness

Figure 1.6: Three different forms of sparse networks.

Modular Network Structure

The modular network consists of several functionally independent subnetworks of lower dimensions, as is shown in Figure 1.6(b). The modular structure is in line with the functional specialization of human brain, where the brain areas specialized in the same function tend to be in the same cluster. The modularity has been widely reported in the literature on brain networks (Milo *et al.*, 2002, 2004; Newman, 2006; Sporns, 2011), and has attracted much attention in researching the brain's functional organization. As such, we impose the modular structure on the ODDM parameters that denote the brain's effective connectivity, extending the ODDM to a new ODE model, called modular oscillator dynamic directional model (MODDM). The MODDM simultaneously describes the oscillatory activity of brain regions and characterizes the modular structure of the brain network, which leads to scientifically meaningful network results.

Small-world Network Structure

Unlike functionally independent subnetworks in the modular structure, where the connections among network nodes within the same cluster are dense and between different clusters are void, different clusters in small-world network (Figure 1.6(c)) can also be connected by a few connections. Further, in the small-world network, nodes within the same cluster are densely connected, and between different clusters are sparsely connected. Such small-world structure reflects two fundamental principles of the functional organization of the brain: functional integration and functional segregation. The brain regions in different clusters are specialized for different brain functions, while the connections among clusters ensure integration among these specialized regions. In addition, the small-world networks (Watts and Strogatz, 1998)

are in an efficient functional organization, in which all the nodes are connected directly or indirectly through a few edges, and have been found in many real-world phenomena, including brain networks (Sporns *et al.*, 2004; Yu *et al.*, 2008), human social networks (Milgram, 1967), Internet networks, and many others. Thus, we propose a new high-dimensional ODE model, called small-world oscillatory dynamic directional model (SWODDM), that characterizes the small-world network structure of the brain for iEEG data. Existing studies of effective connectivity mainly focus on functional integration among a few large brain regions (low-dimensional), or functional segregation among many small brain regions without considering their integration. Whereas, the SWODDM, a high-dimensional model in the small-world network structure, can characterize functional integration and functional segregation of human brain simultaneously.

1.3.3 Model Estimation Consideration

We identify clusters, select connected brain regions, and evaluate the strength of effective connectivity by estimating the ODDM (i.e., MODDM and SWODDM) parameters based on the observed data. To address the difficult in specifying the complex interactive relationship among many regions, we use a linear approximation in ODDM to model the complex mechanism of the high-dimensional brain system, an idea similar to the linear regression. Like many other statistical models, ODDM has a discrepancy from the underlying true mechanism of the brain because of the approximation. We develop a Bayesian framework to estimate the model parameters of the ODDM while accounting for the discrepancy between the proposed models and the underlying true brain. The quantification of the ODE model discrepancy has rarely been addressed in the literature. In addition, the Bayesian approach offers an

inference framework for evaluating the statistical significance of estimated effective connectivity. As such, we not only address a pressing need for statistical modeling of the biophysical mechanism of the brain but also introduce a new approach to inferring high-dimensional ODE models with many free parameters.

The rest of the thesis is organized as follows. Section 2 introduces the ODDM that simultaneously accommodates the systems' oscillatory activity and characterizes the directional connections among components of high-dimensional dynamic systems in two types of network structures (Section 2.2.1 and 4.2 respectively). Section 3 presents two Bayesian hierarchical methods to estimate the MODDM and SWODDM parameters based on basis representation of brain regions' state functions and a parallel Markov chain Monte Carlo (MCMC) algorithm to make posterior inference. Section 4 presents simulation studies of the proposed Bayesian method in comparison with existing network methods. Section 5 applies the ODDM to analyze a real iEEG study. Section 6 discusses analysis results and future research directions.

Chapter 2

ODE Models for iEEG Data

Let $\mathbf{y}(t) = (y_1(t), \ldots, y_d(t))'$ be the observed iEEG measurements of d regions at time t and $\mathbf{x}(t) = (x_1(t), \ldots, x_d(t))'$ be the neuronal state functions of the d brain regions at time t. Since each iEEG electrode directly records one brain region's neuronal electrical activity, we assume the following observation model that links observed data $\mathbf{y}(t)$ to the underlying states $\mathbf{x}(t)$:

$$\mathbf{y}(t) = \mathbf{x}(t) + \boldsymbol{\epsilon}(t), \tag{2.1}$$

where $\boldsymbol{\epsilon}(t) = (\epsilon_1(t), \ldots, \epsilon_d(t))'$ is a d-dimensional vector of measurement errors with mean zeroes. The observed data, $\mathbf{y}(t)$, are measured at discrete time points $t = 1, 2, \ldots, T$.

Since brain regions interact with each other through neuron firing, the model for the brain's directional activity is constructed at the regions' neuronal level, i.e., for $\mathbf{x}(t)$. Existing high-dimensional ODE models for a dynamic system with many interactive components, including the first-order linear or bilinear ODEs (Zhang *et al.*, 2015, 2017) and several other first-order ODEs (Chen and Wu, 2008; Lu *et al.*,

2011; Wu *et al.*, 2014a,b), do not accommodate oscillatory activity of the system. To address this limitation, we propose to use the damped harmonic oscillator (DHO, Serway and Jewett, 2003), a one-dimensional oscillatory physical system, to build our model for $\mathbf{x}(t)$.

2.1 Oscillatory Dynamic Directional Model (ODDM)

The DHO is a one-dimensional second-order ODE given by

$$\frac{d^2z(t)}{dt^2} = F(t) + A\ z(t) + G\ \frac{dz(t)}{dt}, \tag{2.2}$$

where $z(t)$ is the state of the spatial location of a one-dimensional space (called oscillator) at time t, $F(t)$ is an externally applied force, and the parameters A and G determine the oscillator's oscillation amplitude and period(see Chapter 2 in Fitzpatrick, 2013, for detailed explanation). Simple one-dimensional oscillatory systems described by the DHO include a spring-mass system and pendulum. Figure 2.1 shows temporal activities of three DHOs with $F(t) = 0$ and different combines of parameters A and G, which lead to different frequencies of the time series.

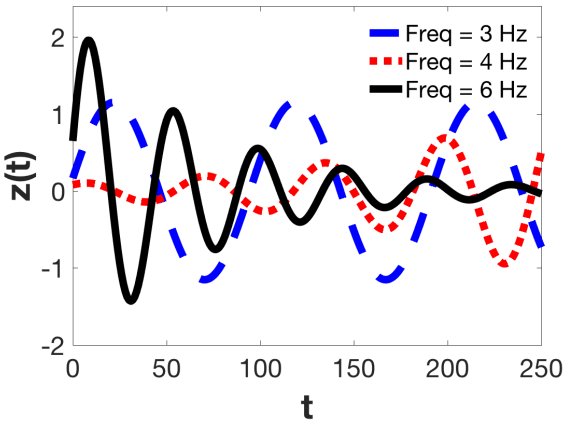


Figure 2.1: Time series of different DHOs

We extend the one-dimensional DHO to the high-dimensional brain system consisting of many interactive brain regions. Specifically, we model d brain regions under study as a set of interactive oscillators—each corresponding to one region—influenced by the effect exerted by others. For region i,

$$\frac{d^2x_i(t)}{dt^2} = F_i(\mathbf{x}(t)) + A_{ii}\ x_i(t) + G_i\ \frac{dx_i(t)}{dt}, \tag{2.3}$$

where $F_i(\mathbf{x}(t))$ is the directional effect exerted by other regions over region i, and parameters A_{ii} and G_i are associated with region i's oscillation amplitude and period.

The functions $F_i(\mathbf{x}(t))$, $i = 1, 2, \ldots, d$, represent the directional/effective connectivity among d regions and are difficult to specify due to the limited understanding of the brain's biophysical mechanism. Moreover, for the proposed high-dimensional ODE model, it is difficult to get around identifiability issue for nonlinear F_is, that is, many different combinations of model parameters may lead to the same or similar state functions $\mathbf{x}(t)$. To address this issue and also to reduce the model complexity, we use a first-order Taylor expansion, $\sum_{j\neq i}^{d} A_{ij} \cdot x_j(t) + D_i$, to approximate the complex function $F_i(\mathbf{x}(t))$, which leads to the following model:

$$\frac{d^2x_i(t)}{dt^2} = D_i + \sum_{j\neq i}^{d} A_{ij} \cdot x_j(t) + A_{ii} \cdot x_i(t) + G_i\ \frac{dx_i(t)}{dt}. \tag{2.4}$$

where D_i is the intercept in the first order Taylor approximation of $F_i(\mathbf{x}(t))$ and A_{ij} represents the directional effect exerted by $x_j(t)$ on $x_i(t)$.

The model formulation in (2.4) brings three benefits in practice. First, the second-order ODEs with a physical foundation are more suitable for characterizing the brain's oscillatory activity. The amplitude and frequency of the oscillations can be estimated from the model parameters associated with d damped harmonic

oscillators. Second, this simple form provides a flexible way to model directional interactions, i.e., effective connectivity, among different brain regions, because each model parameter A_{ij} denotes the directional effect exerted by region j on region i. Third, the linear form facilitates fast computation for high-dimensional data with a large d.

The state model (2.4) together with the observation model (2.1) is referred to as the *Oscillatory Dynamic Directional Model* (ODDM). Model parameters $\mathbf{A} = \{A_{ij}, i = 1, \ldots, d,\ j = 1, \ldots, d\}$, $\mathbf{G} = \{G_i, i = 1, \ldots, d\}$ and $\mathbf{D} = \{D_i, i = 1, \ldots, d\}$ are unknown and to be estimated based on the observed time series $\mathbf{y}(t),\ t = 1, \ldots, T$.

Note that the ODDM is **NOT** an extension of the first-order ODE model to second-order ones despite their mathematical similarity. Second-order ODEs have several different formulations. For example, a comprehensive second-order linear ODE model should include all possible first-order derivatives, $\frac{dx_j(t)}{dt}$, $j = 1, \ldots, d$. We do not include those terms because the ensuing model no longer has a direct physical interpretation. In short, the formulation of the ODDM stems from its integration of statistical modeling and scientific modeling rather than first-order linear ODEs.

Under the ODDM, inference about effective connectivity among the d regions is equivalent to estimating parameters $\mathbf{A}$, and mapping the brain's directional network is equivalent to identifying statistically significant nonzero A_{ij}s. Note that because the two directional effects between each pair of regions i and j are characterized by two separate parameters, A_{ij} and A_{ji}, the proposed method indeed produces separate estimates of the directional effects in two directions. As such, the total number of parameters for quantifying directional connections among d regions is d^2. This is different from many association studies (Kramer *et al.*, 2008, 2010,

2012; Mormann *et al.*, 2005; Netoff and Schiff, 2002; Schiff *et al.*, 2005; Schindler *et al.*, 2006, 2010, 2008; Wendling *et al.*, 1996; Wu and Gotman, 1998) in which only one parameter is used to characterize the association relationship, i.e., functional connectivity, between each pair of regions.

2.2 Oscillatory Dynamic Directional Model for Sparse Brain Networks

For a high-dimensional brain system with a large d, estimates of many ODDM parameters can be unstable and have large variances. To improve estimation efficiency of the ODDM, we assume that many parameters A_{ij}s are zero. A motivation for sparsity lies in the established idea that directional connections are energy consuming (Anderson, 2005; Földiák and Young, 1995; Olshausen and Field, 2004), and biological systems tend to minimize energy consuming activities (Bullmore and Sporns, 2009; Micheloyannis, 2012).

Among different sparse network structures, we are particularly interested in two types of network structures, the modular network structure and the small-world network structure, both of which reflect the functional organization of brain networks. The two network structures both rely on the existence of clusters with densely interconnected nodes, corresponding to functionally specialized brain areas, with the difference that the small-world one assumes sparse connections between clusters, corresponding to the integration of segregated areas, while the other does not. Either of the two network structures has been widely reported in the literature on brain networks (Milo *et al.*, 2002, 2004; Newman, 2006; Sporns, 2011 for modularity and Sporns *et al.*, 2004; Yu *et al.*, 2008 for small-worldness).

For clarity, we refer to the densely connected part common in both network structures as cluster, which may have or may not have connections with other parts of the network.

2.2.1 Modular Oscillatory Dynamic Directional Model (MODDM)

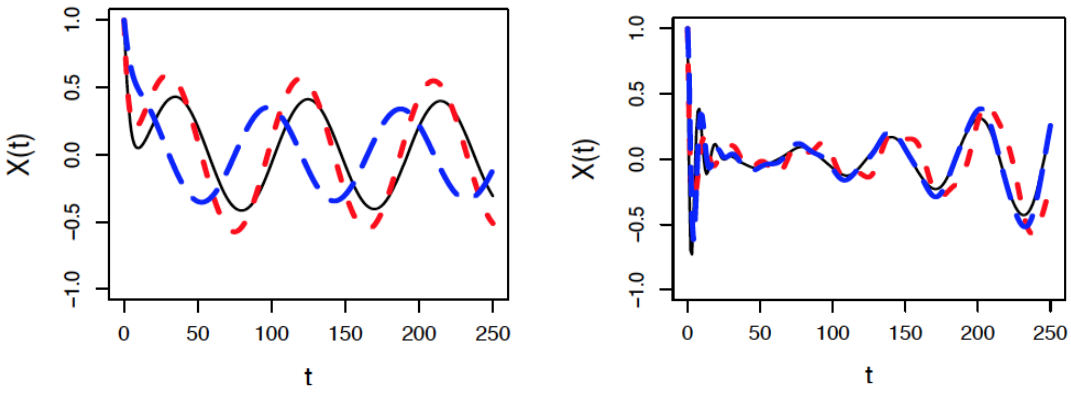
To characterize modularity, we introduce cluster labels $\boldsymbol{m} = \{m_1, m_2, \ldots, m_d\}$, which take integer values from 1 to d, to denote the clusters of the d brain regions. As such, the brain network can have at most d clusters, each consisting of one region only. We use indicator γ_{ij} with a value of 1 to indicate the significant effective connectivity from region j to region i and 0 otherwise. We propose the following ODE model, as an extension of the ODDM (2.4), for the sparse brain network in modularity:

$$\frac{d^2x_i(t)}{dt^2} = \sum_{j=1}^{d} \delta(m_i, m_j) \cdot \gamma_{ij} \cdot A_{ij} \cdot x_j(t) + D_i + G_i \, \frac{dx_i(t)}{dt}, \tag{2.5}$$

where the delta function, $\delta(m_i, m_j)$, equals 1 if $m_i = m_j$ and 0 otherwise. Model (2.5) implies that a directional effect from region j onto region i is nonzero, if and only if the two regions fall into the same cluster and the associated indicator γ_{ij} is nonzero.

Model (2.5) with the observation model (2.1) is referred to as the *Modular Oscillatory Dynamic Directional Model* (MODDM). To illustrate the MODDM, we generated $\mathbf{x}(t)$ from a dynamic system with two clusters. The regions in the same cluster are all pairwise connected. Figure 2.2(a) shows $x(t)$s of three regions in one cluster and Figure 2.2(b) for the other cluster. The MODDM produces oscillatory state functions, and the state functions of the regions within the same cluster have

the same or similar oscillation frequencies.



(a) $\mathbf{x}(t)$ in the First Cluster (b) $\mathbf{x}(t)$ in the Second Cluster

Figure 2.2: Simulated time series from MODDM.

2.2.2 Small-world Oscillatory Dynamic Directional Model (SWODDM)

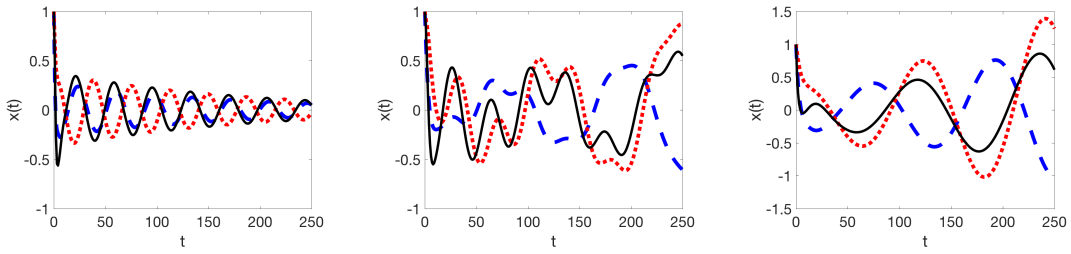
In order to characterize small-world network structure, we keep the cluster labels $\boldsymbol{m}$ and indicators γ_{ij}s introduced in MODDM, to characterize the significant effective connectivity within clusters, and add another set of indicators γ_{ij}^{B}s, combined with cluster labels $\boldsymbol{m}$ to characterize the significant effective connectivity between clusters. For clarity, we replace γ_{ij}s by γ_{ij}^{W}s to distinguish within-cluster connections γ_{ij}^{W}s with between-cluster connections γ_{ij}^{B}s. The indicators γ_{ij}^{B}s have the same usage as γ_{ij}s in MODDM, with value of 1 to indicate the significant effective connectivity from region j to region i and 0 otherwise. We propose the model for small-world

brain network as follows:

$$\frac{d^2x_i(t)}{dt^2} = \sum_{j=1}^{d} \left[\delta(m_i, m_j) \cdot \gamma_{ij}^W + (1 - \delta(m_i, m_j)) \cdot \gamma_{ij}^B\right] \cdot A_{ij} \cdot x_j(t) + D_i + G_i \frac{dx_i(t)}{dt}. \tag{2.6}$$

Model (2.6) indicates that region j has a nonzero directional effect on region i either if the two regions are in the same cluster, i.e., $m_i = m_j$, and γ_{ij}^W is nonzero, or if $m_i \neq m_j$ and γ_{ij}^B is nonzero.

The model (2.6) together with the observation model (2.1) is referred to as the *Small-world Oscillatory Dynamic Directional Model* (SWODDM). To illustrate the difference between SWODDM and MODDM, i.e., the effect of between-cluster connections on the entire system, we generated $\mathbf{x}(t)$ from a dynamic system with two clusters. The regions in the same cluster are pairwise connected and one region in the first cluster has a directional effect on the other one region in the second cluster. Figure 2.3(a) shows $x(t)$s of three regions in the first cluster and Figure 2.3(b) shows



(a) $\mathbf{x}(t)$ in the First Cluster (b) $\mathbf{x}(t)$ in the Second Cluster (c) $\mathbf{x}(t)$ in the Second Cluster without between cluster connections

Figure 2.3: Simulated time series from SWODDM.

$x(t)$s of three regions in the second cluster, in which the region corresponding to the black curve in the first cluster (Figure 2.3(a)) has a a directional connection with the the region corresponding to the black curve in the second cluster (Figure

2.3(a)). As a comparison, Figure 2.3(c) shows the $x(t)$s of three regions in the second cluster when the directional connection between the two clusters is taken off. Since the direction of the between-cluster connection is from the first cluster to the second cluster, the simulated time series in the first cluster do not change with or without the between-cluster connection. While it is clear to see that, with the directional connection between the two clusters, the time series in the second cluster are affected by the first cluster, which tend to have a transition from a low frequency to a higher frequency, as those have in the first cluster. Moreover, not only the black curve, corresponding to the direct affected region, the other curves in the second cluster have also changed (red and blue curves in Figure 2.3(b)) because of the pairwise connection within the second cluster, but with a delay and slight degree. That indicates the information flow from the first cluster has spread to the entire second cluster through the directional connection between them. SWODDM is expected to capture such pattern of information flow to illustrate the functional integration of separated brain regions.

Chapter 3

Bayesian Estimation Methods for MODDM and SWODDM

Under the ODDM, the focus is on identifying clusters, selecting network edges within clusters in MODDM, corresponding to nonzero $\delta(m_i, m_j) \cdot \gamma_{ij}$s, or network edges within clusters and between clusters in SWODDM, corresponding to nonzero $\delta(m_i, m_j) \cdot \gamma_{ij}^W + (1 - \delta(m_i, m_j)) \cdot \gamma_{ij}^B$s. We develop two Bayesian methods to estimate these parameters.

3.1 Estimation Methods for ODE models

Two types of approaches are proposed for estimating ODE models in the literature: discretization approaches which numerically $\mathbf{x}(t)$ based on the assumed ODE model (Bard, 1974; Biegler *et al.*, 1986; Campbell, 2007; Cao *et al.*, 2012; Gelman *et al.*, 1996; Girolami, 2008; Hemker, 1972; Huang and Wu, 2006; Huang *et al.*, 2006; Li *et al.*, 2005; Mattheij and Molenaar, 1996; Xue *et al.*, 2010), and basis-function-expansion approaches which represent $\mathbf{x}(t)$ with functional bases (Bhaumik and

Ghosal, 2014; Brunel *et al.*, 2008; Deuflhard and Bornemann, 2012; Poyton *et al.*, 2006; Qi *et al.*, 2010; Ramsay, 2004; Ramsay *et al.*, 2007; Varah, 1982). The discretization approaches usually rely on the numerical methods to approximate the solution of ODEs given parameters by linearization, such as Runge-Kutta algorithm, where nonlinear systems are replaced by first-order Taylor expansions, so that they are restricted to a short time period. Besides, the discretization approaches usually estimate parameters through optimization algorithms. Thus, we take the latter approach as it accounts for the model error, as explained in detail below.

The authors (Ramsay *et al.*, 2007) used third-order B-spline bases to represent the state functions that follow first-order ODE models. In this article, since the second-order derivatives of $\mathbf{x}(t)$ are smooth, we represent state functions $\mathbf{x}(t)$ with fifth-order B-spline basis functions, $\boldsymbol{b}(t) = (b_1(t), \ldots, b_L(t))'$ defined on an equally spaced partition $\{t_1 = 1, t_2, \ldots, t_q = T\}$ (we use $q = T * 5$) of the interval $[1, T]$:

$$x_i(t) = \tilde{\eta}_i' \, \boldsymbol{b}(t), \tag{3.1}$$

where $\tilde{\eta}_i = (\eta_{i1}, \ldots, \eta_{iL})'$ is the vector of the basis coefficients of $x_i(t)$.

3.2 Bayesian Estimation Methods for ODDM

As in Ramsay, 2004, we chose the number of basis functions L comparable to the number of data points T for enough flexibility to fit functional curves. We have tried three different numbers of $L : L = [T]$, $L = [T/2]$, and $L = [T/3]$ and found that the three numbers lead to similar fitted $\mathbf{x}(t)$ and similar accuracy in selecting connected regions by the proposed approach. Following Ruppert, 2002, we used the generalized cross-validation to determine L and got $L = [T/3]$.

We assume the data measurement error $\epsilon_i(t)$ in (2.1) independent with a normal distribution with mean zero and unknown variance σ_i^2. As such,

$$y_i(t) \sim N(x_i(t), \sigma_i^2). \tag{3.2}$$

We show in the simulation study that because of the strong SNR of the data, the proposed method is robust to violations of the model assumptions for $\epsilon_i(t)$. Next, we assign to basis coefficients a prior, also a prior distribution for $\mathbf{x}(t)$.

3.2.1 Model for Basis Coefficients/State Functions

For simplicity, we describe the basis coefficient model for MODDM only, which can be easily generalized to SWODDM by denoting $\{\boldsymbol{\gamma}^W, \boldsymbol{\gamma}^B\}$ as $\boldsymbol{\gamma}$ and replacing $\delta(m_i, m_j) \cdot \gamma_{ij}$ in model-fitting errors with $\delta(m_i, m_j) \cdot \gamma_{ij}^W + (1 - \delta(m_i, m_j)) \cdot \gamma_{ij}^B$.

Let $\boldsymbol{\eta} = \{\tilde{\eta}_i, i = 1, \ldots, d\}$, $\boldsymbol{\gamma} = \{\gamma_{ij}, i, j = 1, \ldots, d\}$ and $\boldsymbol{\theta} = \{\boldsymbol{A}, \boldsymbol{D}, \boldsymbol{G}\}$. Further, denote all the ODDM parameters by $\boldsymbol{\Theta}_I = \{\mathbf{A}, \mathbf{D}, \mathbf{G}, \boldsymbol{m}, \boldsymbol{\gamma}\}$. We propose a model for basis coefficients $\boldsymbol{\eta}$ that is conditional on $\boldsymbol{\Theta}_I$ through the MODDM model-fitting errors:

$$p(\boldsymbol{\eta}|\boldsymbol{\Theta}_I, \boldsymbol{\tau}) \propto \exp\left\{ -\sum_{i=1}^{d} \frac{\mathrm{R}_i(\boldsymbol{\eta}, \boldsymbol{\Theta}_I)}{2\tau_i} \right\}, \tag{3.3}$$

where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)'$ are positive hyperparameters. $\mathrm{R}_i(\boldsymbol{\eta}, \boldsymbol{\Theta}_I)$ is the model-fitting error of region i's state function with the form:

$$\mathrm{R}_i(\boldsymbol{\eta}, \boldsymbol{\Theta}_I) = \int_0^T \left(\frac{d^2 x_i(t)}{dt^2} - \sum_{j=1}^{d} \delta(m_i, m_j) \cdot \gamma_{ij} \cdot A_{ij} \cdot x_j(t) - D_i - G_i \cdot \frac{dx_i(t)}{dt} \right)^2 dt.$$

In the above equation, all state functions and their derivatives are represented by

basis functions: $x_i(t) = \tilde{\eta}_i' \, \boldsymbol{b}(t)$, $dx_i(t)/dt = \tilde{\eta}_i' \, \boldsymbol{b}^{(1)}(t)$, and $d^2x_i(t)/dt^2 = \tilde{\eta}_i' \, \boldsymbol{b}^{(2)}(t)$.

The distribution (3.3) for $\boldsymbol{\eta}$, with a form of the exponential of the negative model-fitting errors, has an intuitive explanation. The hyperparameters τ_is are the variances of the model-fitting errors for different regions' temporal activities. And this probability (3.3) provides a generating model for the basis coefficients $\boldsymbol{\eta}$ and the state functions $\mathbf{x}(t)$ from the MODDM. It allows for the deviation of the state functions from the assumed ODE model while suggesting a preference for the state functions with small ODE fitting errors. Moreover, the formulation of the distribution (3.3), equivalent to a normal distribution for $\boldsymbol{\eta}$, as explained below, leads to normal posterior conditional distributions of model parameters $\boldsymbol{\theta} = \{\boldsymbol{A}, \boldsymbol{D}, \boldsymbol{G}\}$, which are easy to simulate.

With the linear basis representation for $\mathbf{x}(t)$ in (3.1), $\sum_{i=1}^{d} \mathrm{R}_i(\boldsymbol{\eta}, \boldsymbol{\Theta}_I)/\tau_i$ given $\boldsymbol{\Theta}_I$ and $\boldsymbol{\tau}$ is quadratic of $\boldsymbol{\eta}$:

$$\sum_{i=1}^{d} \mathrm{R}_i(\boldsymbol{\eta}, \boldsymbol{\Theta}_I)/\tau_i = \boldsymbol{\eta}' \, \boldsymbol{\Omega}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}} \, \boldsymbol{\eta} - 2\boldsymbol{\Lambda}'_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}} \, \boldsymbol{\eta} + \boldsymbol{\Xi}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}}, \tag{3.4}$$

where $\boldsymbol{\Omega}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}}$, $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}}$, and $\boldsymbol{\Xi}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}}$, respectively, are a $dL \times dL$ matrix, a $dL \times 1$ vector, and a scalar, and their formulas depending on $\boldsymbol{\Theta}_I$ and $\boldsymbol{\tau}$ are provided in the Appendix 7.2. Thus, the probability model (3.3) is multivariate normal distribution

$$\boldsymbol{\eta} | \boldsymbol{\Theta}_I, \tau \sim \mathrm{MN}(\boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}} \, \boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}}, \boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I, \boldsymbol{\tau}}), \tag{3.5}$$

This means that the state functions follow a Gaussian process centered at the MODDM with variability controlled by $\boldsymbol{\tau}$.

3.2.2 Prior Specification for ODDM Parameters

We propose the following joint prior for MODDM parameters $\boldsymbol{\Theta}_I = \{\mathbf{A}, \mathbf{D}, \mathbf{G}, \boldsymbol{m}, \boldsymbol{\gamma}\}$, similar to the prior proposed for ODE model parameters by (Zhang *et al.*, 2017):

$$p(\boldsymbol{\Theta}_I|\boldsymbol{\tau}) \propto \det(\Omega_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}})^{-1/2} \cdot \exp\left\{\frac{1}{2}(\boldsymbol{\Lambda}'_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\ \boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}} - \Xi_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}})\right\} \cdot \exp\left\{-\mu \sum_{i,j=1}^{d} \delta(m_i, m_j)\right\}$$
$$\cdot\, p_0^{\sum_{i,j}\gamma_{ij}} \cdot (1-p_0)^{d^2-\sum_{i,j}\gamma_{ij}} \cdot \prod_{i,j=1}^{d} \phi\left(\frac{A_{ij}}{\xi_0}\right) \cdot \prod_{i=1}^{d} \phi\left(\frac{G_i}{\xi_0}\right) \cdot \prod_{i=1}^{d} \phi\left(\frac{D_i}{\xi_0}\right), \tag{3.6}$$

where $\phi(\cdot)$ is the standard normal density, ξ_0 is a large constant to given an almost flat prior for $\mathbf{A}$, $\mathbf{G}$, and $\mathbf{D}$, μ is a nonnegative constant and p_0 is a given prior probability. We let $\mu = 0$ to give a non-informative prior for the modularity and let $p_0 = 0.9$ to impose the prior belief that within-cluster connections are dense. We have tried different values for p_0 and found that setting $p_0 = 0.9$ produced the highest true positive rate in selecting network edges. This is because a large value of p_0 effectively reflects the prior information that the connections within clusters are dense and facilitates the cluster identification, while smaller p_0 leads to lower selection accuracy.

By minor changes to (3.6), we have the joint prior for SWODDM parameters $\boldsymbol{\Theta}_I = \{\mathbf{A}, \mathbf{D}, \mathbf{G}, \boldsymbol{m}, \boldsymbol{\gamma}\}$, where $\boldsymbol{\gamma} = \{\boldsymbol{\gamma}^W, \boldsymbol{\gamma}^B\}$ in this case:

$$p(\boldsymbol{\Theta}_I|\boldsymbol{\tau}) \propto \det(\Omega_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}})^{-1/2} \cdot \exp\left\{\frac{1}{2}(\boldsymbol{\Lambda}'_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\ \boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}} - \Xi_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}})\right\} \cdot \exp\left\{-\mu \sum_{i,j=1}^{d} \delta(m_i, m_j)\right\}$$
$$\cdot\, p_w^{\sum_{i,j}\gamma_{ij}^W} \cdot (1-p_w)^{d^2-\sum_{i,j}\gamma_{ij}^W} \cdot p_b^{\sum_{i,j}\gamma_{ij}^B} \cdot (1-p_b)^{d^2-\sum_{i,j}\gamma_{ij}^B}$$
$$\cdot \prod_{i,j=1}^{d} \phi\left(\frac{A_{ij}}{\xi_0}\right) \cdot \prod_{i=1}^{d} \phi\left(\frac{G_i}{\xi_0}\right) \cdot \prod_{i=1}^{d} \phi\left(\frac{D_i}{\xi_0}\right), \tag{3.7}$$

where p_w, p_b reflects prior information on within-cluster connection probability and between-cluster connection probability respectively. We set p_w a large value close to 1 to reflect dense connections within clusters and p_b a small value close to 0 to reflect sparse long-range connections between clusters.

3.2.3 Joint Posterior Distribution

Let $\mathbf{Y} = \{\mathbf{y}(t), t = 1, \ldots, T\}$ and $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_d\}$. The model (3.2) together with priors in (3.3) and (3.6, MODDM)/(3.7, SWODDM) define a hierarchical Bayesian model for the ODDM. The joint posterior distribution is

$$p(\boldsymbol{\Theta}_I, \boldsymbol{\eta}, \boldsymbol{\sigma} | \mathbf{Y}, \boldsymbol{\tau}) \propto p(\mathbf{Y} | \boldsymbol{\eta}, \boldsymbol{\sigma}) \cdot p(\boldsymbol{\sigma}) \cdot p(\boldsymbol{\eta} | \boldsymbol{\Theta}_I, \boldsymbol{\tau}) \cdot p(\boldsymbol{\Theta}_I | \boldsymbol{\tau}), \tag{3.8}$$

where $p(\boldsymbol{\sigma}) \propto \prod_{i=1}^{d} 1/\sigma_i^2$, is an uninformative prior for σ_i^2.

3.3 An Alternative Bayesian Estimation Method

Standard approaches simulate from the posterior distribution $p(\boldsymbol{\Theta}_I, \boldsymbol{\eta}, \boldsymbol{\sigma} | \mathbf{Y}, \boldsymbol{\tau})$ and estimate the state functions $\mathbf{x}(t)$ and ODE parameters jointly within the Bayesian framework (Zhang *et al.*, 2017). However, in the problem under study, we focus on the posterior inference of parameters Θ_I only, while $\boldsymbol{\eta}$ contributes to most parameters in the Bayesian model (3.8). In addition, since iEEG data are smooth with a strong SNR (Cervenka *et al.*, 2013), estimated $x_i(t)$s by a nonparametric smoothing method (Ramsay, 2004) are similar to those from the Bayesian model (3.8). Considering this, we propose to first fit $\mathbf{x}(t)$ based on the observed data $\mathbf{Y}$, and use a Bayesian method to estimate ODDM parameters based on the fitted $\mathbf{x}(t)$, i.e., the estimated $\boldsymbol{\eta}$. In this case, we leave out the posterior sampling of parameters $\boldsymbol{\eta}, \boldsymbol{\sigma}$

to enable fast computation. We elaborate the details in the following.

We first estimate $\boldsymbol{\eta}$ by minimizing

$$\sum_{i=1}\sum_{t=1}^{T}(y_i(t)-x_i(t))^2+\lambda\sum_{i=1}^{d}\int_1^T\left(x_i^{(2)}(t)\right)^2dt, \tag{3.9}$$

where $x_i(t)$ is given by (3.1), $x_i^{(2)}(t)=\tilde{\eta}_i'\boldsymbol{b}^{(2)}(t)$, and the smoothing penalty parameter λ is chosen by the generalized cross-validation (Härdle, 1990).

Next, we treat estimated $\boldsymbol{\eta}$ as the observed data with the likelihood (3.3). As such, the joint posterior distribution of ODDM parameters Θ_I is simplified into

$$p(\boldsymbol{\Theta}_I|\boldsymbol{\eta},\boldsymbol{\tau})\propto p(\boldsymbol{\eta}|\boldsymbol{\Theta}_I,\boldsymbol{\tau})\cdot p(\boldsymbol{\Theta}_I|\boldsymbol{\tau}). \tag{3.10}$$

Since the hyperparameters $\boldsymbol{\tau}$ for the variance of model-fitting error can also be treated a variable within the Bayesian framework, we add a prior for it such that $\boldsymbol{\tau}$ can account for the variability of model fitting error variance (we didn't add a prior for $\boldsymbol{\tau}$ in model (3.8) because it leads to an improper posterior). A commonly used non-informative prior for variance parameters is $p(\boldsymbol{\tau})\propto\prod_{i=1}^{d}1/\tau_i$ (Gelman *et al.*, 2013). However, this prior leads to an improper posterior. Thus, we propose the following prior for $\boldsymbol{\tau}$, which is close to the non-informative prior yet leads to a proper posterior:

$$p(\boldsymbol{\tau})\propto\prod_{i=1}^{d}\left(\frac{1}{\tau_i}\right)^{\frac{3}{2}} \tag{3.11}$$

As such, (3.3), (3.11) and (3.6) for MODDM or (3.7) for SWODDM define the second hierarchical Bayesian model for the ODDM, and the ensuing joint posterior distribution of Θ_I is

$$p(\boldsymbol{\Theta}_I,\boldsymbol{\tau}|\boldsymbol{\eta})\propto p(\boldsymbol{\eta}|\boldsymbol{\Theta}_I,\boldsymbol{\tau})\cdot p(\boldsymbol{\Theta}_I|\boldsymbol{\tau})\cdot p(\boldsymbol{\tau}). \tag{3.12}$$

We refer to the Bayesian model with the posterior (3.8) as the full Bayeisian method, and the model with the posterior (3.12) as the Bayesian smoothing method hereafter.

We use a partially collapsed Gibbs Sampler (PCGS, Van Dyk and Park, 2008) to sample from the posterior distributions $p(\boldsymbol{\Theta}_I, \boldsymbol{\eta}, \boldsymbol{\sigma}|\mathbf{Y}, \boldsymbol{\tau})$ and $p(\boldsymbol{\Theta}_I, \boldsymbol{\tau}|\boldsymbol{\eta})$. The posterior sampling steps of the two distributions are similar while the former takes at least three times more computational time than the latter. We present the Markov chain Monte Carlo (MCMC) steps of the latter in the Appendix 7.1.

3.4 Posterior Inference

Let $\varsigma^{(s)}$ be the sth MCMC simulation of the ODDM parameter ς for $s = 1, \ldots, S$, where S is the total number of MCMC iterations excluding the burn-in time. We use Gelman-Rubin statistics (Gelman *et al.*, 2013) to examine the convergence of the MCMC outputs.

Two posterior quantities of particular interest are the posterior clustering probability and the posterior network edge selection probability. The former is the posterior probability that two regions, i and j, are in the same cluster, which is estimated by $\hat{P}_{ij}^m = \frac{1}{S}\sum_{s=1}^{S} \delta(m_i^{(s)}, m_j^{(s)})$, $i, j = 1, \ldots, d$. The latter is the posterior probability that the directional connection from region i to region j is nonzero and is estimated by $\hat{P}_{ij}^{\gamma} = \frac{1}{S} \sum_{s=1}^{S} \delta(m_i^{(s)}, m_j^{(s)}) \cdot \gamma_{ij}^{(s)}$ in MODDM, or $\hat{P}_{ij}^{\gamma^W} = \frac{1}{S} \sum_{s=1}^{S} \delta(m_i^{(s)}, m_j^{(s)}) \cdot \gamma_{ij}^{W^{(s)}}$ (within-cluster) and $\hat{P}_{ij}^{\gamma^B} = \frac{1}{S} \sum_{s=1}^{S} \left[1 - \delta(m_i^{(s)}, m_j^{(s)})\right] \cdot \gamma_{ij}^{B^{(s)}}$ (between-cluster) in SWODDM.

Given a threshold $\hbar^m$ for $\hat{P}_{ij}^m$, we put the pair of brain regions (i, j) with $\hat{P}_{ij}^m > \hbar$ in the same cluster and group brain regions into different clusters accordingly. Based on the identified clusters, we use network edge selection probabilities, $\hat{P}_{ij}^{\gamma}$ or $\hat{P}_{ij}^{\gamma^W}$ together with $\hat{P}_{ij}^{\gamma^B}$, to select the directional network edge (from j to i) if $\hat{P}_{ij}^{\gamma} > \hbar^{\gamma}$

for a given threshold $\hbar^{\gamma}$ for $\hat{P}_{ij}^{\gamma}$, or $\hat{P}_{ij}^{\gamma^W} > \hbar^{\gamma^W}$ for a given threshold $\hbar^{\gamma^W}$ for $\hat{P}_{ij}^{\gamma^W}$ and $\hat{P}_{ij}^{\gamma^B} > \hbar^{\gamma^B}$ for a given threshold $\hbar^{\gamma^B}$ for $\hat{P}_{ij}^{\gamma^B}$.

3.4.1 Choice of thresholds

We propose to use the false discovery rate (FDR) to determine the thresholds for posterior probabilities. To evaluate the FDR, we develop a method to approximate the null distributions of $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma}$s, $\hat{P}_{ij}^{\gamma^W}$s, $\hat{P}_{ij}^{\gamma^B}$s under the null hypothesis that none of the regions are connected. We first generate the data, denoted by $\mathbf{Y}^0$, that satisfies the null. Given long multivariate time-series data, we divide them into short segments of the same length T. We randomly sample the time segment of each region with the pairwise distance between any two regions' segments greater than $10T$. All the segments combined create $\mathbf{Y}^0$ (of the same size as $\mathbf{Y}$). The posterior probabilities $\{\hat{P}_{ij}^{m_0}, i,j = 1,\ldots,d\}$ and $\{\hat{P}_{ij}^{\gamma_0}, i,j = 1,\ldots,d\}$, $\{\hat{P}_{ij}^{\gamma_0^W}, i,j = 1,\ldots,d\}$, $\{\hat{P}_{ij}^{\gamma_0^B}, i,j = 1,\ldots,d\}$ based on $\mathbf{Y}^0$ give the empirical null distributions of $\hat{P}_{ij}^{m}$ and $\hat{P}_{ij}^{\gamma}$, $\hat{P}_{ij}^{\gamma^W}$, $\hat{P}_{ij}^{\gamma^B}$, respectively. Based on these null distributions, we evaluate the significance levels of $\hat{P}_{ij}^{m}$ and $\hat{P}_{ij}^{\gamma}$, $\hat{P}_{ij}^{\gamma^W}$, $\hat{P}_{ij}^{\gamma^B}$ for every pair of regions i and j and determine the thresholds corresponding to 5% FDR using the method by (Benjamini and Hochberg, 1995; Efron and Tibshirani, 2002).

3.4.2 Computational Time

The computational time of the ODE model estimation depends on the dimension of the system d, the length of time points T, the number of clusters K and the number of regions in the largest cluster. If the number of regions in each cluster is roughly the same, the computational time of the proposed Bayesian method is $O(T \cdot \frac{d^4}{K^2})$. We reduce this computation time to $O(T \cdot \frac{d^3}{K})$ by using parallel computing (Caffo

et al., 2010; Suchard *et al.*, 2010).

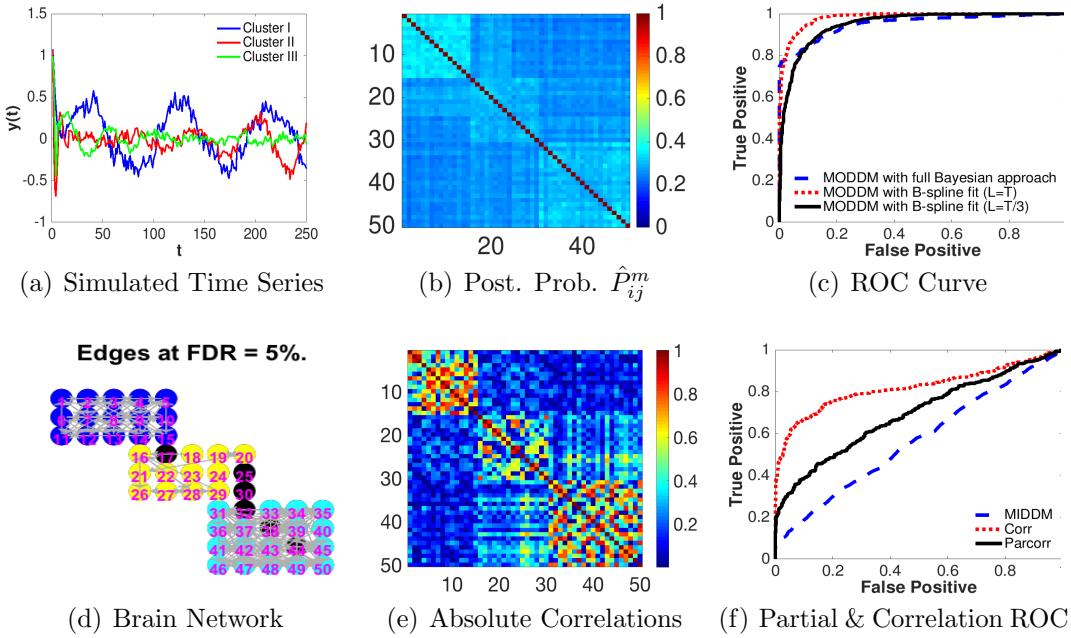
Chapter 4

Simulation Study

4.1 MODDM

We consider a 50-dimensional dynamic system ($d = 50$) that has 3 clusters of size 15, 15, and 20. We first generated $x_i(t), i = 1, \ldots, 50$, from the MODDM (2.5) and (2.1). For simplicity, we let the components within the same cluster be all pairwise connected and let $D_i = 0$ and $x_i(0) = 1$ for all i. Within each cluster l, $l = 1, 2, 3$, we let $A_{ii} = a_l$, $A_{i\ i+1} = (-1)^i \cdot b_l$, and $G_i = c_l$ for region i in cluster l. The rest of A_{ij}s were simulated from a standard normal. We chose different values for a_l, b_l and c_l in different clusters so that the three clusters have different oscillatory features. Specifically, $a_1 = -3.6$, $b_1 = 2.2$, and $c_1 = -20$ for the first cluster; $a_2 = -7.1$, $b_2 = 8.2$, and $c_2 = -15$ for the second cluster; and $a_3 = -4$, $b_3 = 1.8$, and $c_3 = -4$ for the third cluster. We simulated 50 time series $\boldsymbol{\epsilon}(t)$ from an AR(1) model with lag-1 autocorrelation equalling 0.5 and median pairwise spatial correlations between regions equalling 0.2 (the median spatial correlation of real data is no more than 0.2). We chose the variances of $\boldsymbol{\epsilon}(t)$ such that the SNR of each time series—defined as $var(x_i(t))/var(\epsilon_i(t))$—equals 20, which is far below SNRs of typical iEEG data

(Zhang *et al.*, 2015) (the median SNR of real iEEG data is above 100). Finally, we obtained $y_i(t)$ as the sum of $x_i(t)$ and $\epsilon_i(t)$. Figure 4.1(a) shows three representative time series $y_i(t)$, each from one unique cluster.



(a) Simulated Time Series (b) Post. Prob. $\hat{P}_{ij}^m$ (c) ROC Curve

(d) Brain Network (e) Absolute Correlations (f) Partial & Correlation ROC

Figure 4.1: 4.1(a) Simulated three time series from three different clusters. 4.1(b) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 50$ of the Bayesian smoothing method. 4.1(c) The ROC curve of network edge selection by the full Bayesian approach and the Bayesian smoothing approach using the number of basis functions $L = [T]$ and $L = [T/3]$. 4.1(d) Network edges with 5% FDR for $\hat{P}_{ij}^m$ and $\hat{P}_{ij}^\gamma$. Nodes in the same color correspond to components in the same cluster identified by the Bayesian smoothing method. 4.1(e) The matrix of the absolute correlations of the simulated data. 4.1(f) The ROC curves of the network edge selection based on the first-order ODE model MIDDM, correlations and partial correlations of the simulated data.

After standardizing each time series $y_i(t), i = 1, \ldots 50$, to mean zero and norm 1, we applied the developed two Bayesian approaches to the simulated data set. Figure 4.1(b) shows the posterior clustering probabilities $\hat{P}_{ij}^m$ estimated by the Bayesian smoothing method. From the figure, we see that $\hat{P}_{ij}^m$s for the regions truly in the same clusters are consistently greater than $\hat{P}_{ij}^m$s for the regions in different clusters. This also demonstrates that $\hat{P}_{ij}^m$ is a good measure used for clustering regions.

We calculated the true positive rates (TPR) and false positive rates (FPR) of the proposed method using different thresholds for $\hat{P}_{ij}^{\gamma}$. TPR is the percentage of the network edges with $\hat{P}_{ij}^{\gamma}$ above the threshold among all true network edges; FPR is the percentage of the network edges with $\hat{P}_{ij}^{\gamma}$ above the threshold among all void network edges. Figure 4.1(c) shows the ROC curves (pairs of TPRs and FPRs for different thresholds) for the full Bayesian method and the Bayesian smooth method using the number of bases $L = [T]$ and $L = [T/3]$. The figure indicates that the network estimation by the two Bayesian approaches with different numbers of basis functions are similar and the proposed method is able to select network edges with a high TPR and low FPR. Figure 4.1(d) shows the directional network edges selected using 5% FDR thresholds for $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma}$s estimated by the Bayesian smoothing method. The nodes in the same color correspond to components identified to be in the same cluster. The proposed method can identify three clusters.

For comparison, we applied the existing first-order ODE model, called the modular and indicator-based dynamic directional model (MIDDM) (Zhang *et al.*, 2015, 2017), to the data. As shown in Figure 4.1(f), the MIDDM gave mediocre results with an AUC of 0.58, indicating that the first-order ODEs are unable to capture the connectivity among regions with oscillatory activities. Furthermore, we compared with network methods based on correlations and partial correlations. Figure 4.1(e) shows the matrix of the absolute correlations of the simulated data. The calculated pairwise correlations of the regions truly in the same clusters are not consistently large as expected despite their time series have similar oscillatory frequencies. Figure 4.1(f) shows the ROC curves of network edge selection by using different thresholds for the correlations (with an AUC of 0.82) and partial correlations (with an AUC of 0.67). Despite similar oscillatory frequencies of the time series within the same cluster, their correlations can still be small. Thus, the proposed method outperforms

the correlation-based methods.

We also evaluated the mean estimation error for the estimated parameters: $\sum_{ij}(A_{ij} - \hat{A}_{ij})^2/d^2$, which is 1.46. For comparison, we evaluated the mean estimation error by MIDDM, which is 1.48. We believe that the large estimation errors are due to the short time series and the many parameters in the model for quantifying directional connectivity. In practice, the network edge selection is of more interest than model parameter estimation, because the proposed model is an approximation of the underlying complex system and the detected network edges provide valuable information of the existence of directional connections among regions. We showed that the network edge selection by the proposed Bayesian method has high accuracy.

To further demonstrate the efficiency of the proposed method for network edge selection, we independently simulated 100 multivariate time series from the same model and applied the proposed approach to each data set. Table 4.1 lists the summaries of areas under the ROC curve (AUC) and indicates that the proposed method is robust to violation of model assumptions and can consistently detect connected regions with high accuracy.

min	Q1	median	Q3	max	standard deviation
0.9244	0.9453	0.9531	0.9595	0.9658	0.0098

Table 4.1: Summary AUC for 100 Replicates.

4.2 SWODDM

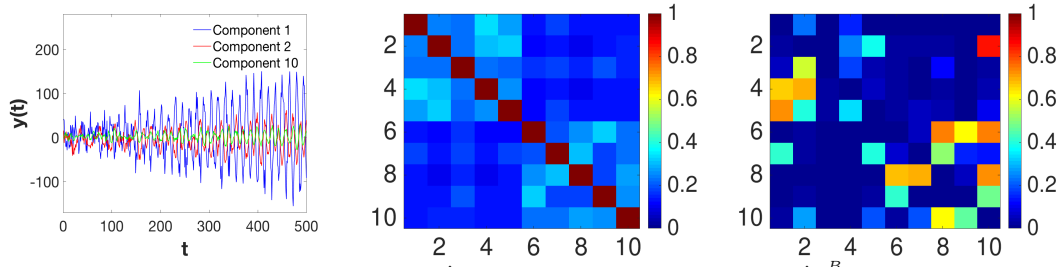
We present several dynamic systems of different dimensions and different between-cluster edge connections generated from the assumed model SWODDM ((2.6) and (2.1)) with a SNR of 10, and compare the performance of SWODDM and MODDM

on the simulated systems.

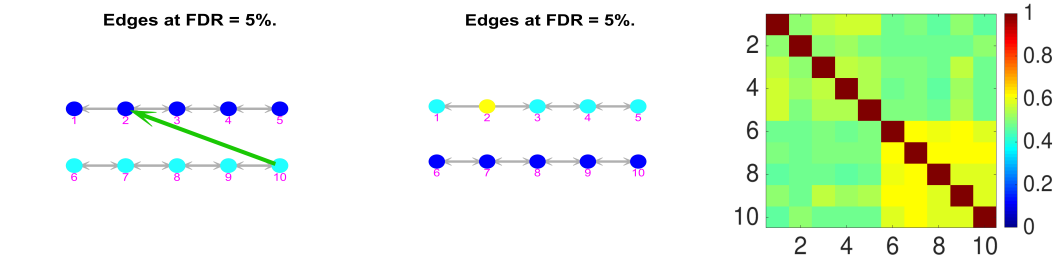
Network Edge Selection by FDR

We start from a low-dimensional dynamic system consisting of 2 clusters of size 5 and 5, where the components within the same cluster are all pairwise connected and moreover, component 10 in the second cluster has a directional effect on component 2 in the first cluster. Figure 4.2(a) shows three time series of the simulated system, two from the first cluster, component 1 and component 2, and one from the second cluster, component 10. We applied both the SWODDM and MODDM to the simulated data.

Figure 4.2(b) and 4.2(c) show the posterior clustering probabilities $\hat{P}_{ij}^{m}$ and the posterior between-cluster network edge selection probabilities $\hat{P}_{ij}^{\gamma^B}$ respectively, estimated by SWODDM using the full Bayesian method. It is clear to see that $\hat{P}_{ij}^{m}$s for regions within the same cluster are generally higher than those for the regions in different clusters, and the $\hat{P}_{2\ 10}^{\gamma^B}$, which denotes the posterior probability of the existence of the directional connection from component 10 to component 2, has the highest value among $\hat{P}_{ij}^{\gamma^B}$s for all pairs of components. Figure 4.2(d) shows the directional network edges selected using 5% FDR thresholds for $\hat{P}_{ij}^{m}$s, $\hat{P}_{ij}^{\gamma^W}$s and $\hat{P}_{ij}^{\gamma^B}$s estimated by the full Bayesian approach. The grey arrows correspond to directional connections within the same cluster while the green arrow corresponds to the between-cluster connection from component 10 to component 2. Specifically, the nodes of different colors, i.e., nodes in different clusters, are identified by the FDR procedure on posterior clustering probabilities $\hat{P}_{ij}^{m}$; the grey arrows, i.e., within-cluster directional connections, are selected by the FDR procedure on posterior within-cluster network edge selection probabilities $\hat{P}_{ij}^{\gamma^W}$ and the green arrow, i.e., between-cluster directional connections, are selected by the FDR procedure on pos-



(a) Simulated Time Series (b) $\hat{P}_{ij}^{m}$ from SWODDM (c) $\hat{P}_{ij}^{\gamma^B}$ from SWODDM



(d) SWODDM Brain Network (e) MODDM Brain Network (f) $\hat{P}_{ij}^{m}$ from MODDM

Figure 4.2: 4.2(a) Simulated three time series from two different clusters. 4.2(b) Posterior clustering probabilities $\hat{P}_{ij}^{m}$ for $i, j = 1, \ldots, 10$ from SWODDM by the full Bayesian method. 4.2(c) Posterior between-cluster network edge selection probabilities $\hat{P}_{ij}^{\gamma^B}$ for $i, j = 1, \ldots, 10$ from SWODDM by the full Bayesian method. 4.2(d) Network edges with 5% FDR for $\hat{P}_{ij}^{m}$, $\hat{P}_{ij}^{\gamma^W}$ and $\hat{P}_{ij}^{\gamma^B}$ from SWODDM. Nodes in the same color correspond to components in the same cluster. Grey arrows denote the directional connections in the same cluster and the green arrow denotes the directional connection between the two clusters. 4.2(e) Network edges with 5% FDR for $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma}$s from MODDM. 4.2(f) Posterior clustering probabilities $\hat{P}_{ij}^{m}$ for $i, j = 1, \ldots, 10$ from MODDM by the full Bayesian method.

terior between-cluster network edge selection probabilities $\hat{P}_{ij}^{\gamma^B}$. As a comparison, Figure 4.2(e) shows the network edges selected by MODDM. We can see that the MODDM isolates the component 2 from the other components in the first cluster, that is to say, MODDM fails to detect the connections between component 2 and the other components in the first cluster, and the connections from component 10 to component 2 denoted by the green arrow in Figure 4.2(d). Besides, the network result indicates that the failure of MODDM to capture the significant directional connections between clusters may lead to the misclassification of clusters. Figure

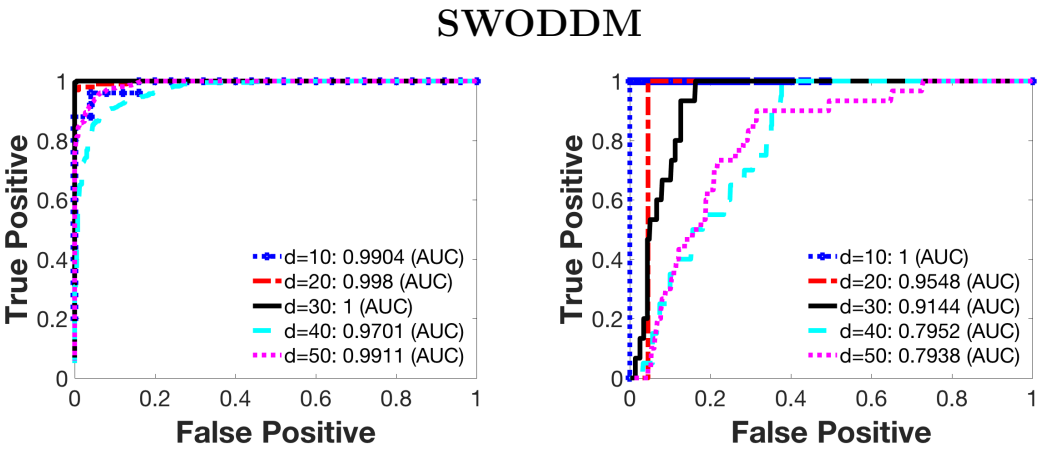
4.2(f) shows the posterior clustering probabilities $\hat{P}_{ij}^{m}$ estimated by MODDM using the full Bayesian method.

Comparison on Systems with Different Dimensions

We also compare the performance of the SWODDM and MODDM on five dynamic systems with different dimensions, from 10 to 50 with a step size of 10. Since the clustering in the network and directional connections between the clusters correspond to functional segregation and functional integration, the functional organization of human brain, we are particularly interested in the posterior clustering probabilities $\hat{P}_{ij}^{m}$ and posterior between-cluster network edge selection probabilities $\hat{P}_{ij}^{\gamma^B}$.

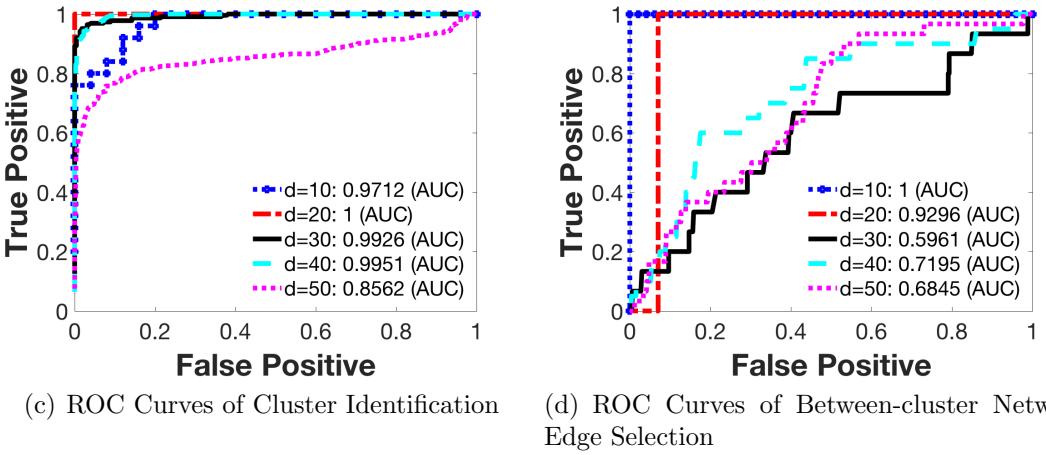
We generated five dynamic systems of different dimensions ($d = 10, 20, \ldots, 50$) from the assumed model SWODDM (2.6) and (2.1). For $d = 10, \ldots, 40$, each of the system consists 2 clusters of size $d/2$ and $d/2$ and for $d = 50$, the system consists 3 clusters of size 15, 15 and 20. The components within the same cluster are pairwise connected and some of components belonging to different clusters are also connected. For the system with $d = 10$, there is a directional connection from component 10 in the second cluster to component 2 in the first cluster. For the system with $d = 20$, there is a directional connection from component 5 in the first cluster to component 13 in the second cluster. For the system with $d = 30$, each of the components in the first cluster has a directional effect on component 23 in the second cluster, representing component 23 is affected by all components in the first cluster together. For the system with $d = 40$, each of the components in the first cluster has a directional effect on component 21 in the second cluster, representing component 21 is affected by all components in the first cluster together. And for the system with $d = 50$, component 36 in the third cluster has directional connections to all components in the first and the second clusters, representing component 36 exerts

effect on the other two clusters. We applied both the SWODDM and MODDM on the five simulated systems.



(a) ROC Curves of Cluster Identification

(b) ROC Curves of Between-cluster Network Edge Selection



(c) ROC Curves of Cluster Identification

(d) ROC Curves of Between-cluster Network Edge Selection

Figure 4.3: ROC curves of cluster identification and network edge selection between clusters for SWODDM and MODDM on systems of different dimensions.

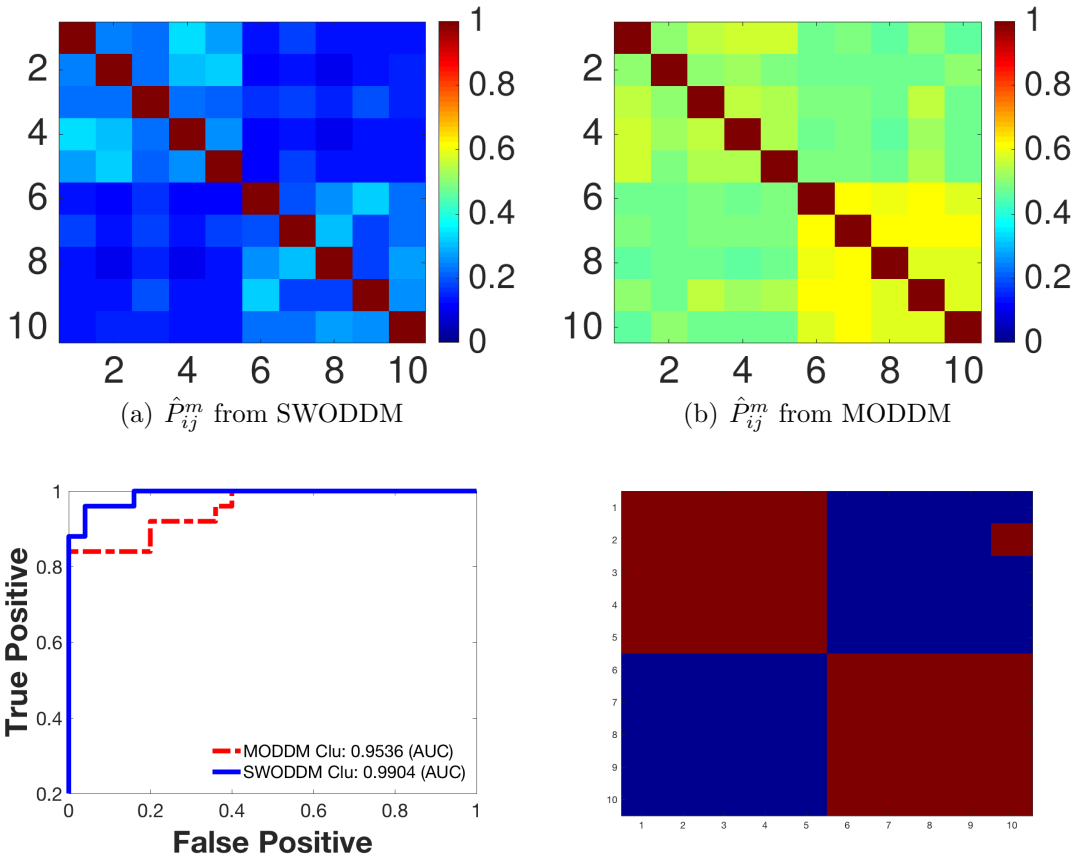
Figure 4.3(a) and 4.3(b) show the ROC curves of $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma^B}$s $(i, j = 1, \ldots, d)$ from the SWODDM on the five systems. The AUCs of posterior clustering probabilities $\hat{P}_{ij}^{m}$ are all above 0.97 for the five systems of different dimensions, which indicates the SWODDM is able to identify the clusters in low-dimensional systems as well as in high-dimensional systems. As the dimension of systems increases, the

AUC of the posterior between-cluster network edge selection probabilities $\hat{P}_{ij}^{\gamma^B}$ decreases, from 1 for $d = 10$ to around 0.8 for $d = 40, 50$. Figrue 4.3(c) and 4.3(d) show the ROC curves of $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma}$s $(i, j = 1, \ldots, d)$ from MODDM on the five systems. Figrue 4.3(c) indicates clustering in high-dimensional systems suffer more from between-cluster directional connections than in low-dimensional systems. And to make it comparable with posterior between-cluster network edge selection probabilities $\hat{P}_{ij}^{\gamma^B}$ in SWODDM, we use the posterior network edge selection probabilities $\hat{P}_{ij}^{\gamma}$ in MODDM to select the network edges between clusters as well. Originally, $\hat{P}_{ij}^{\gamma}$s are used to select the network edges within clusters identified by $\hat{P}_{ij}^{m}$s. Now we set two different thresholds, $\hbar^{\gamma^W}$ and $\hbar^{\gamma^B}$, for $\hat{P}_{ij}^{\gamma}$s to select within-cluster connections and between-cluster connections respectively. Specifically, if region i and region j belong to the same cluster and $\hat{P}_{ij}^{\gamma} > \hbar^{\gamma^W}$, the directional edge from region j to region i is selected, considered as a within-cluster connection. And if region i and region j do not belong to the same cluster and $\hat{P}_{ij}^{\gamma} > \hbar^{\gamma^B}$, the directional edge from region j to region i is also selected, considered as a between-cluster connection. The ROC curves in Figure 4.3(d) correspond to $\hat{P}_{ij}^{\gamma}$s constrained on region i and region j not in the same cluster, which is equivalent to $\hat{P}_{ij}^{\gamma^B}$s in SWODDM. Compared with Figure 4.3(b), it indicates that without modeling on the connections between clusters, the significant directional connections that reflect functional integration of human brain will not be detected, especially in high-dimensional systems. The AUC of $\hat{P}_{ij}^{\gamma}$s, where region i and region j are not in the same cluster, drops to below 0.6 for d above 30, which is pretty much like a random guess.

• More on comparison of SWODDM and MODDM in cluster identification

To further illustrate how MODDM achieves worse clustering performance than SWODDM due to lack of between-cluster connection modeling, we present several examples of simulations in different dimensions and structures to compare, i.e., one

in 10 dimension and three in 50 dimension considering the complexity of between-cluster connections in high dimensions.



(a) $\hat{P}_{ij}^m$ from SWODDM (b) $\hat{P}_{ij}^m$ from MODDM

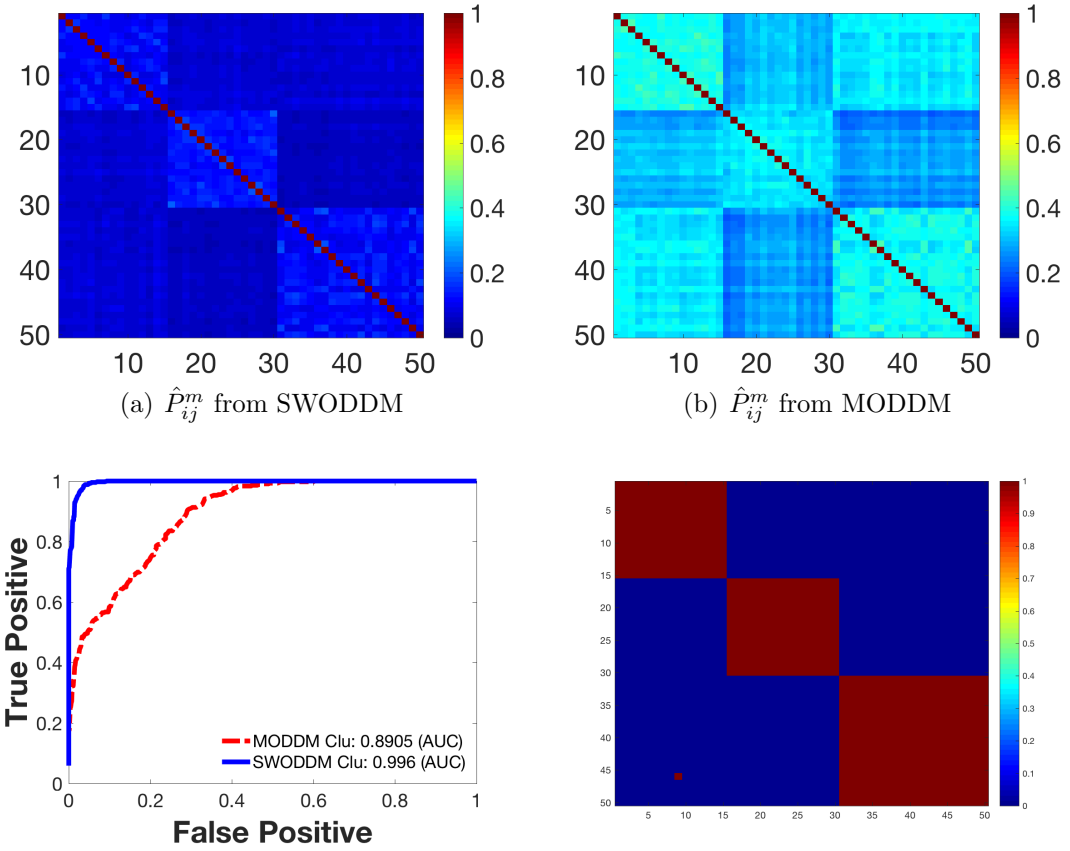
(c) ROC Curves of Cluster Identification (d) True Connection Structure

Figure 4.4: A 10-dimensional simulation example of 2 clusters with single connection between the 2 clusters. 4.4(a) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from SWODDM. 4.4(b) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from MODDM. 4.4(c) A comparison of ROC curves of cluster identification from SWODDM and MODDM. 4.4(d) The underlying connection structure with red grids corresponding to the existence of directional connection from component #{column index} to component #{row index} and blue ones corresponding to nonexistence.

The example in Figure 4.4 is the first 10-dimensional simulation example we presented in section 4.2, in which there are 2 clusters of size 5 and size 5 (components 1-5 as one cluster and 6-10 as the other), pairwise connections within the same cluster and a directional connection from component 10 to component 2 (as is shown

in Figure 4.4(d)). Compared with Figure 4.4(a), Figure 4.4(b) shows component 2 is identified as an isolated node in the network by MODDM since the MODDM posterior clustering probabilities involved with component 2, the second row and/or the second column, are low. We believe this is because MODDM assumes no connection between different clusters, which contradicts with the fact that there is a directional connection from the second cluster (component 10) to the first cluster (component 2). The existence of 10-to-2 connection sometimes makes MODDM cluster component 2 with components 6-10 while sometimes with components 1-5 in MCMC iterations, which leads to posterior clustering probabilities involved with component 2 are low overall. This is typical for low dimensional systems with simple between-cluster connection structure, i.e., single between-cluster connection.

Next, we present 3 simulation examples in 50 dimensions. All of them consist of 3 pairwise connected clusters, 1-15, 16-30 and 31-50. The difference lies in the between-cluster connection structure. The example in Figure 4.5 has one between-cluster connection,from component 9 to component 46. The one in Figure 4.6 has two between-cluster connections, from component 10 to component 24 and from component 31 to component 19. And the one in Figure 4.7 has multiple between-cluster connections, of which sparsity is 0.02. It is clear to see that in all three cases, SWODDM identifies three clusters successfully while MODDM tends to cluster 2 clusters into one, unlike previous low-dimensional example which isolates components that receive effect from other clusters. Specifically, in the example of single between-cluster connection from the third cluster to the first cluster, MODDM identifies two clusters, the first and the third clusters as one large cluster and the second as the other. We believe the different clustering result of this example compared with the 10-dimensional example is because the individual underlying clusters in this example has a larger size than those in the 10-dimensional example, which



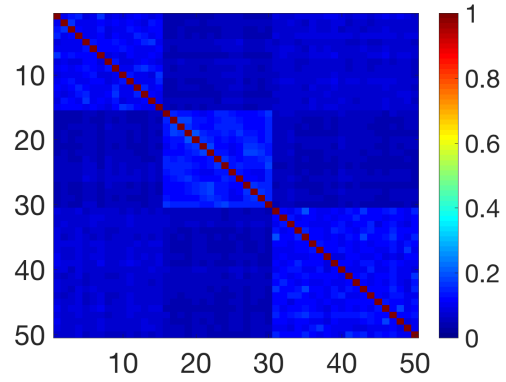
(a) $\hat{P}_{ij}^m$ from SWODDM

(b) $\hat{P}_{ij}^m$ from MODDM

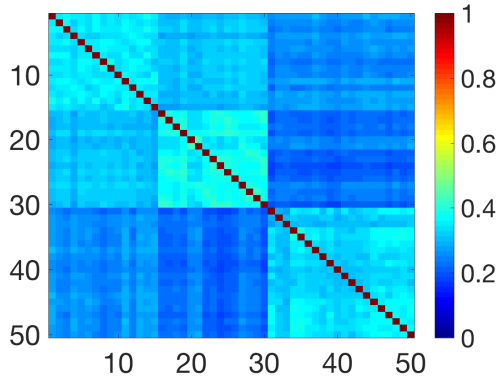
(c) ROC Curves of Cluster Identification

(d) True Connection Structure

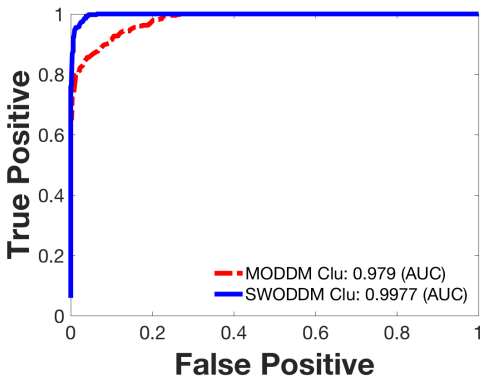
Figure 4.5: A 50-dimensional simulation example of 3 clusters with a single between-cluster connection. 4.5(a) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from SWODDM. 4.5(b) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from MODDM. 4.5(c) A comparison of ROC curves of cluster identification from SWODDM and MODDM. 4.5(d) The underlying connection structure with red grids corresponding to the existence of directional connection from component #{column index} to component #{row index} and blue ones corresponding to nonexistence.



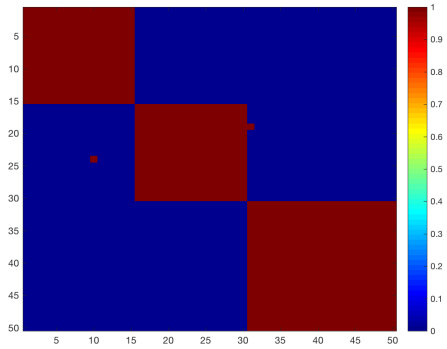
(a) $\hat{P}_{ij}^m$ from SWODDM



(b) $\hat{P}_{ij}^m$ from MODDM

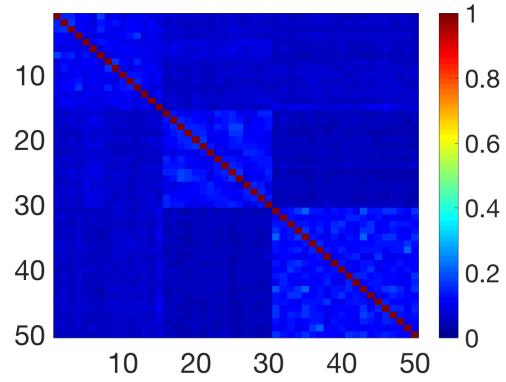


(c) ROC Curves of Cluster Identification

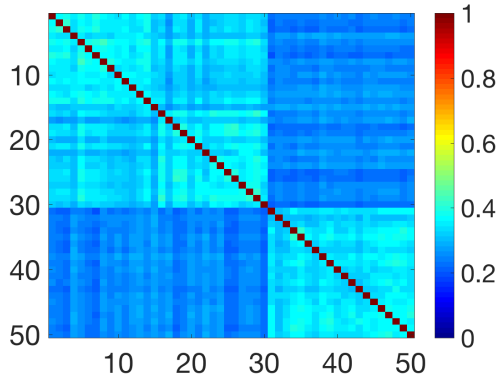


(d) True Connection Structure

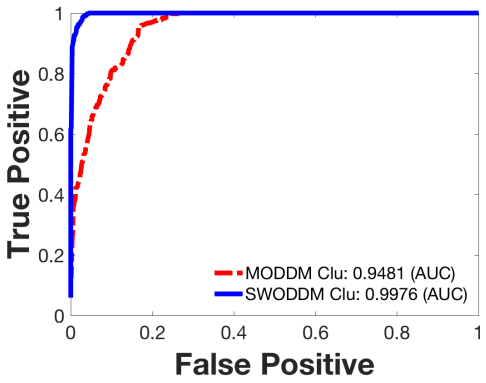
Figure 4.6: A 50-dimensional simulation example of 3 clusters with two between-cluster connections. 4.6(a) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from SWODDM. 4.6(b) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from MODDM. 4.6(c) A comparison of ROC curves of cluster identification from SWODDM and MODDM. 4.6(d) The underlying connection structure with red grids corresponding to the existence of directional connection from component #{column index} to component #{row index} and blue ones corresponding to nonexistence.



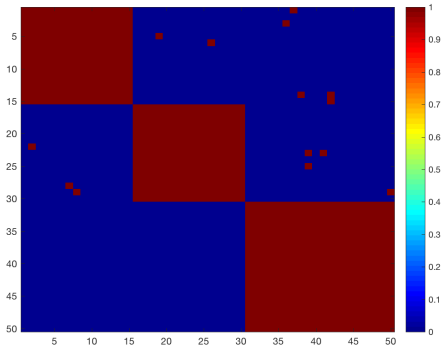
(a) $\hat{P}_{ij}^m$ from SWODDM



(b) $\hat{P}_{ij}^m$ from MODDM



(c) ROC Curves of Cluster Identification



(d) True Connection Structure

Figure 4.7: A 50-dimensional simulation example of 3 clusters with multiple between-cluster connections of sparsity 0.02. 4.7(a) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from SWODDM. 4.7(b) Posterior clustering probabilities $\hat{P}_{ij}^m$ for $i, j = 1, \ldots, 10$ from MODDM. 4.7(c) A comparison of ROC curves of cluster identification from SWODDM and MODDM. 4.7(d) The underlying connection structure with red grids corresponding to the existence of directional connection from component #{column index} to component #{row index} and blue ones corresponding to nonexistence.

leads to increased complexity of interactions within the clusters. In other words, it is hard for MODDM to distinguish the interactions within the same underlying cluster with the directional connections from other clusters (may or may not exist) in a high-dimensional and highly interactive system so that MODDM creates many false positive connections in MCMC iterations and has high posterior clustering probabilities for the two connected clusters. The latter two simulation examples are similar to the first one. But MODDM does not cluster the three underlying clusters into one big cluster although they are connected directly or indirectly. The clustering of MODDM should still depend on the strength of the connections and size ratio of the clusters. We have reason to believe clusters in similar size are tend to be clustered together by MODDM given a directional connection between them than clusters in different sizes.

Network Edge Selection of One Single Between-Cluster Connection on Systems with Different Dimensions

The specialized human brain functions can be integrated by only a few connections between segregated brain areas, especially **SINGLE** significant connection. Thus, we generated three dynamic systems of $d = 20, 35, 50$ with one single strong directional connection between clusters from the assumed model SWODDM (2.6) and (2.1). Specifically, for $d = 20$, the system consist of two clusters of size 10 and 10, where the components within the same cluster are pairwise connected and component 13 in the second cluster has a directional connection to component 5 in the first cluster. For $d = 35$, the system consist of two clusters of size 15 and 20, where the components within the same cluster are pairwise connected and component 23 in the second cluster has a directional connection to component 6 in the first cluster. For $d = 50$, the system consist of three clusters of size 15, 15 and 20, where the

components within the same cluster are pairwise connected and component 14 in the first cluster has a directional connection to component 23 in the second cluster. To illustrate the difficulty of selecting the single true between-cluster connection in high-dimensional systems, we made ROC plots of $\hat{P}_{ij}^{\gamma^B}$s, and network plots where clusters and within-cluster network edges are selected with a 5% FDR threshold but between-cluster network edges are selected by a threshold that just identifies the true connection, explained in detail below.

Figure 4.8 compare the performance of SWODDM and MODDM on the three dynamic systems with single between-cluster connections. Same as described in Section 4.2, $\hat{P}_{ij}^{\gamma}$s constrained on pairs of regions that are not in the same cluster are used to select the connections between clusters for MODDM. Figure 4.8(a), 4.8(b) and 4.8(c), where the blue/cyan curves consistently have larger AUCs than the red/pink curves, indicates that the SWODDM outperforms the MODDM for both low- and high-dimensional systems in terms of the cluster identification and the network edge selection between clusters. Especially for high-dimensional systems, the MODDM can only detect the single between-cluster connection with a high value of false positive rate, 0.8776, which is not acceptable considering the existence of a large amount of void between-cluster connections. Although the between-cluster network edge selection by SWODDM has a nearly perfect ROC curve, with AUC of 0.995, 0.9858 and 0.993 for $d = 20$, $d = 35$ and $d = 50$ respectively, it is hard to identify the single true connection between clusters by the FDR procedure described in Section 3.4.1. Actually, the single between-cluster connection was identified in none of the three systems ($d = 20, 35, 50$) at a 5% FDR threshold. To demonstrate why FDR procedure does not work for the between-cluster network edge selection here, especially in high-dimensional systems, we set the threshold for between-cluster network edge selection probabilities to $\hbar^{\gamma^B *}$ that can just identify the true single between-

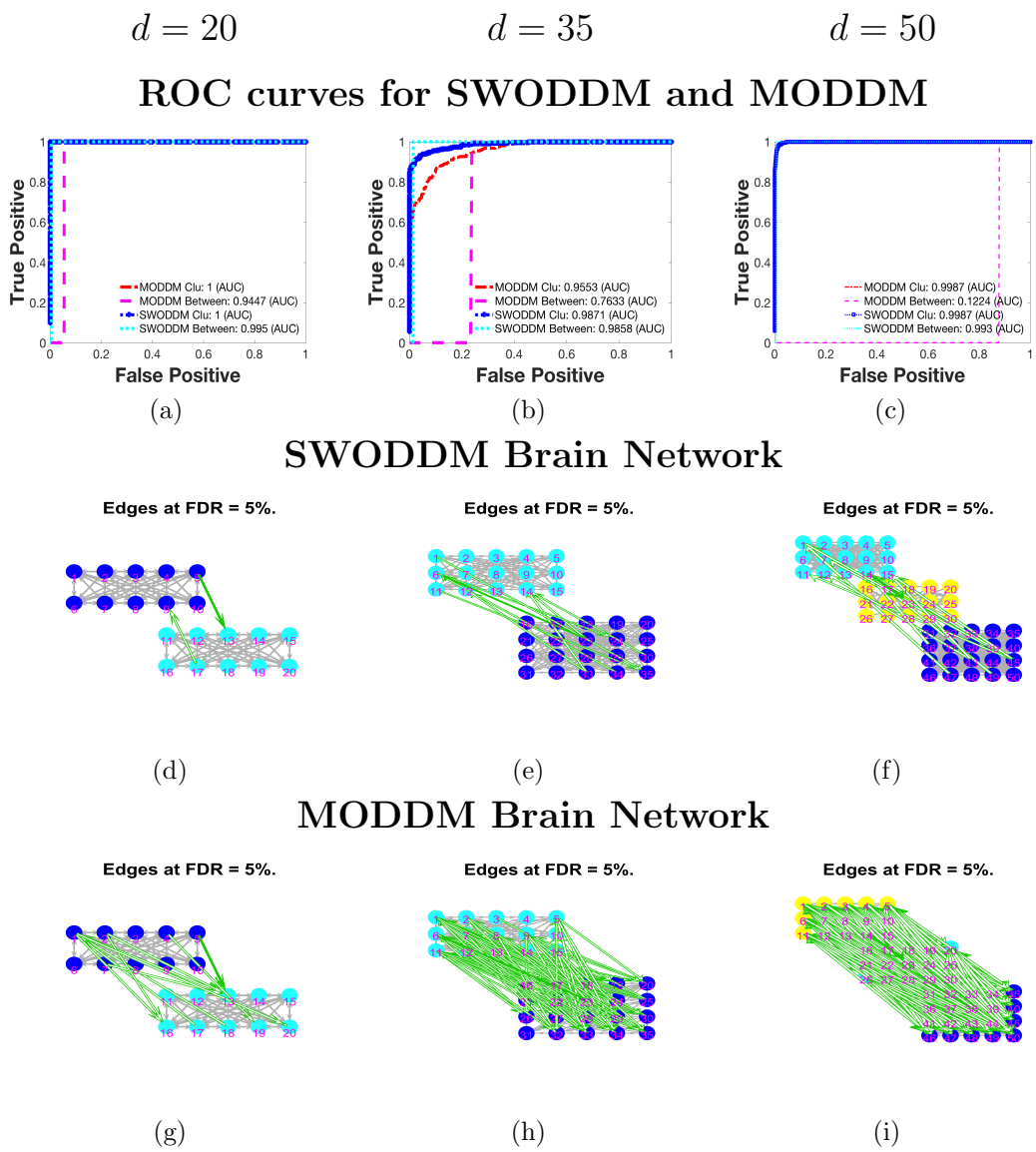


Figure 4.8: 4.8(a), 4.8(b) and 4.8(c) The ROC curves of cluster identification and network edge selection between clusters for SWODDM and MODDM on systems with $d = 10, 35$ and 50, respectively. 4.8(d), 4.8(e) and 4.8(f) The networks by SWODDM at 5% FDR thresholds for $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma^W}$ s and a deliberately chosen threshold for $\hat{P}_{ij}^{\gamma^B}$ s, for three systems with $d = 10, 35$ and 50, respectively. Solid green arrow corresponds to the true single connection between clusters and hallow green arrows correspond to the false connections between clusters. 4.8(g), 4.8(h) and 4.8(i) The networks of the three systems ($d = 10, 35, 50$) by MODDM at 5% FDR thresholds for $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma}$s to identify clusters and select within-cluster connections, and a deliberately chosen threshold for $\hat{P}_{ij}^{\gamma}$s select between-cluster connections.

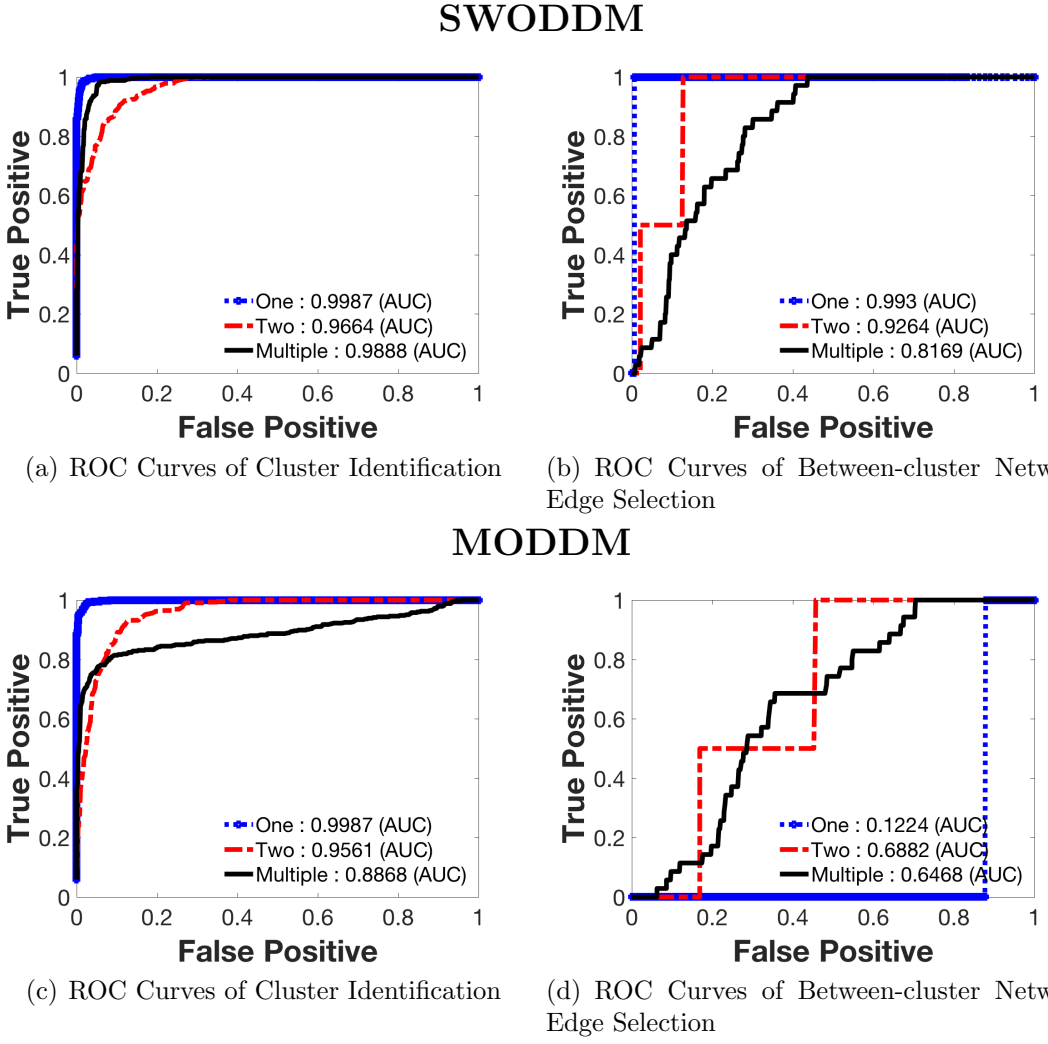
cluster connection from component j^* to component i^*, which means $\hat{P}^{\gamma^B}_{i^*j^*} > \hbar^{\gamma^B *}$ but $\hat{P}^{\gamma^B}_{ij} \leq \hbar^{\gamma^B *}$ for any (i,j)s satisfying $\hat{P}^{\gamma^B}_{ij} < \hat{P}^{\gamma^B}_{i^*j^*}$. In other words, a threshold slightly higher than $\hbar^{\gamma^B *}$ will not select the true connections from j^* to i^* and a threshold slightly lower than $\hbar^{\gamma^B *}$ will select more false positive between-cluster network edges. Figure 4.8(d), 4.8(e) and 4.8(f) show the networks of the three dynamic systems, in which the green arrows correspond to between-cluster network edges selected by SWODDM at threshold $\hbar^{\gamma^B *}$. And among all the green arrows, the solid green arrow correspond to the true single connection between clusters and the hollow green arrows correspond to the false connections between clusters. It is noted that the proportion of true connections among all selected connections are $1/2 = 0.5$ for $d = 20$, $1/9 = 0.1111$ for $d = 35$ and $1/12 = 0.0833$ for $d = 50$ and thus the FDRs, i.e., 1 minus those probabilities, are 0.5 for $d = 20$, 0.8889 for $d = 35$ and 0.9167 for $d = 50$, which makes it impossible to select the true single between-cluster connection by a conventional 5% FDR threshold. And as the dimension increases, the FDR dramatically increases, leading to the true connection more difficult to detect. However, the SWODDM still does much better than MODDM in detecting the single true between-cluster connection. As the networks by MODDM shown in Figure 4.8(g), 4.8(h) and 4.8(i) indicates, the proportion of true connections among all selected connections are $1/12 = 0.0833$ for $d = 20$, $1/144 = 0.0069$ for $d = 35$ and $1/1450 = 0.0007$ for $d = 50$. Correspondingly, the FDRs are 0.9167, 0.9931 and 0.9993 for systems with $d = 20$, $d = 35$ and $d = 50$ respectively. Therefore, the SWODDM reduces the FDR but the problem to detect single directional connection between clusters still remains.

Comparison on High-dimensional Systems with Different Structures of Between-cluster Connections

Although it is difficult to detect the single between-cluster connection, we see the potential of SWODDM to select between-cluster network edges at a high TPR and a low FPR as indicated by the ROC curves in Figure 4.3(b) and 4.8(a), 4.8(b), 4.8(c). And the cluster identification of SWODDM is better than MODDM across different dimensions of systems. Thus, we evaluate the cluster identification and between-cluster network edge selection by SWODDM and MODDM on high-dimensional systems with different structures of between-cluster connections, which may indicate different ways of functional integration.

We consider three 50-dimensional systems generated from the assumed model SWODDM (2.6) and (2.1), each of which consists three clusters of size 15, 15 and 20, and has pairwise connections within the same cluster. The first system has a single between-cluster connection from component 14 in the first cluster to component 23 in the second cluster. The second system has two between-cluster connections, one from component 20 in the second cluster to component 4 in the first cluster and the other from component 25 in the second cluster to component 2 in the first cluster. The third system has multiple between-cluster connections, from each of the components in the second and the third clusters to component 6 in the first cluster.

Figure 4.9 shows the ROC curves of cluster identification and between-cluster network edge selection on the three simulated systems by SWODDM and MODDM. The SWODDM are consistently better than MODDM in identifying the clusters and selecting the network edges between clusters for all three different between-cluster connection structures in terms of the AUCs. Specifically, comparing Figure 4.9(a)



(a) ROC Curves of Cluster Identification

(b) ROC Curves of Between-cluster Network Edge Selection

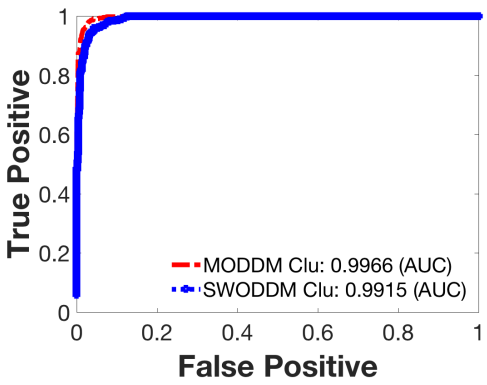
(c) ROC Curves of Cluster Identification

(d) ROC Curves of Between-cluster Network Edge Selection

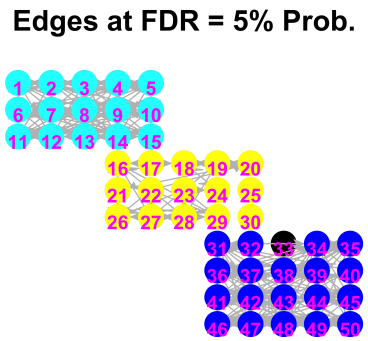
Figure 4.9: ROC curves of cluster identification and network edge selection between clusters for SWODDM and MODDM on high-dimensional systems of different between-cluster connection structures.

and Figure 4.9(c), the SWODDM is able to identify the clusters with a high AUC above 0.96 under all three different structures of connections between clusters while the MODDM has a poor performance of cluster identification when the connections between clusters get complicated e.g., a AUC of 0.8868 for the third system. That indicates when there are indeed various connections between clusters, the MODDM will not capture the true clusters well due to no modeling of the between-cluster connections. In other words, the clusters identified by MODDM are vulnerable to the existence of connections between clusters (which is common in brain networks) while clusters identified by SWODDM are not. Figure 4.9(b) shows, as the complexity of connections between clusters increases, the performance of SWODDM is worse, indicated by a higher false positive rate when all the true edges are detected for the third system (black line) compared with the first (blue line) and the second system (red line). But the overall performance of SWODDM on between-cluster network edge selection is good, with AUCs above 0.8 for all three different between-cluster connection structures, while the MODDM has AUCs below 0.7 or even worse than a random guess (see Figure 4.9(d)).

To demonstrate the performance of SWODDM on cluster identification for systems without connections between clusters, we generated a 50-dimensional system of three clusters of size 15, 15 and 20 from the MODDM (2.5) and (2.1). The components within the same cluster are pairwise connected and no connections are between the clusters. Figure 4.10(a) shows the ROC curves of cluster identification for both the SWODDM and MODDM, indicating the SWODDM is able to identify the true clusters as well as the MODDM. And Figure 4.10(b) shows the network identified by SWODDM at 5% FDR thresholds for $\hat{P}_{ij}^{m}$s, $\hat{P}_{ij}^{\gamma^W}$s and $\hat{P}_{ij}^{\gamma^B}$s. No between-cluster connections selected are expected, consistent with the true cluster structure of the simulated system.



(a) ROC Curves of Cluster Identification



(b) SWODDM Brain Network

Figure 4.10: 4.10(a) ROC curves of cluster identification for SWODDM and MODDM on a high-dimensional system without connection between clusters. 4.10(b) The network identified by SWODDM at 5% FDR thresholds for $\hat{P}_{ij}^{m}$s, $\hat{P}_{ij}^{\gamma^W}$s and $\hat{P}_{ij}^{\gamma^B}$s.

In summary, the SWODDM is able to identify the true clusters of high-dimensional systems no matter the structures of connections between clusters are like while the MODDM cannot. This is a great advantage of SWODDM over MODDM because the existence of connections between different clusters, representing functional integration of segregated areas, are common not only in brain networks but many other networks in life. In addition, the SWODDM is able to select connections between clusters with a high TRP and a low FPR.

Chapter 5

Real Data Study — An Epileptic iEEG Study

5.1 Data Description

The iEEG data under study were collected from a right-handed female adult with medically intractable epilepsy, prior to her brain surgery for seizure treatment. Use of de-identified data was approved by the University of Virginia IRB.

The iEEG recordings of the subject were obtained from 18 chronically implanted depth and 76 subdural electrodes with 10 mm inter-electrode spacing, over the patient's left hemisphere, as shown in Figure 1.2. The acronyms IF, SF, ST, AD, MD, PD, and G stand for inferior frontal, superior frontal, superior temporal, anterior depth, medial depth, posterior depth, and grid electrodes. The iEEG signals were sampled at 200 Hz filtered with a band pass range of 1-70Hz. An additional notch filter was used to filter out 60 Hz interference. The electrodes G07 and G08 were used as ground and reference electrodes, so they were excluded from the analysis. We included in the analysis additional two voltage time series collected from elec-

trodes placed on the chest. These electrodes pick up the signals generated by cardiac muscles, also known as electrocardiogram (EKG). As such, 94 time series were analyzed. For convenience, we associate the two EKG time series with the electrodes G07 and G08 when plotting brain networks.

The subject' brain activity was recorded continuously for over 9 days. In total, iEEG captured four seizures from a single seizure onset zone (SOZ, the brain region where seizures start) G44, which is over supersylvian frontoparietal cortex. Because epileptic activity originating from a single SOZ is the simplest case to study, and G44 was a major SOZ, we evaluated the epileptic patient's brain networks before seizures onset at G44.

5.2 Data Analysis

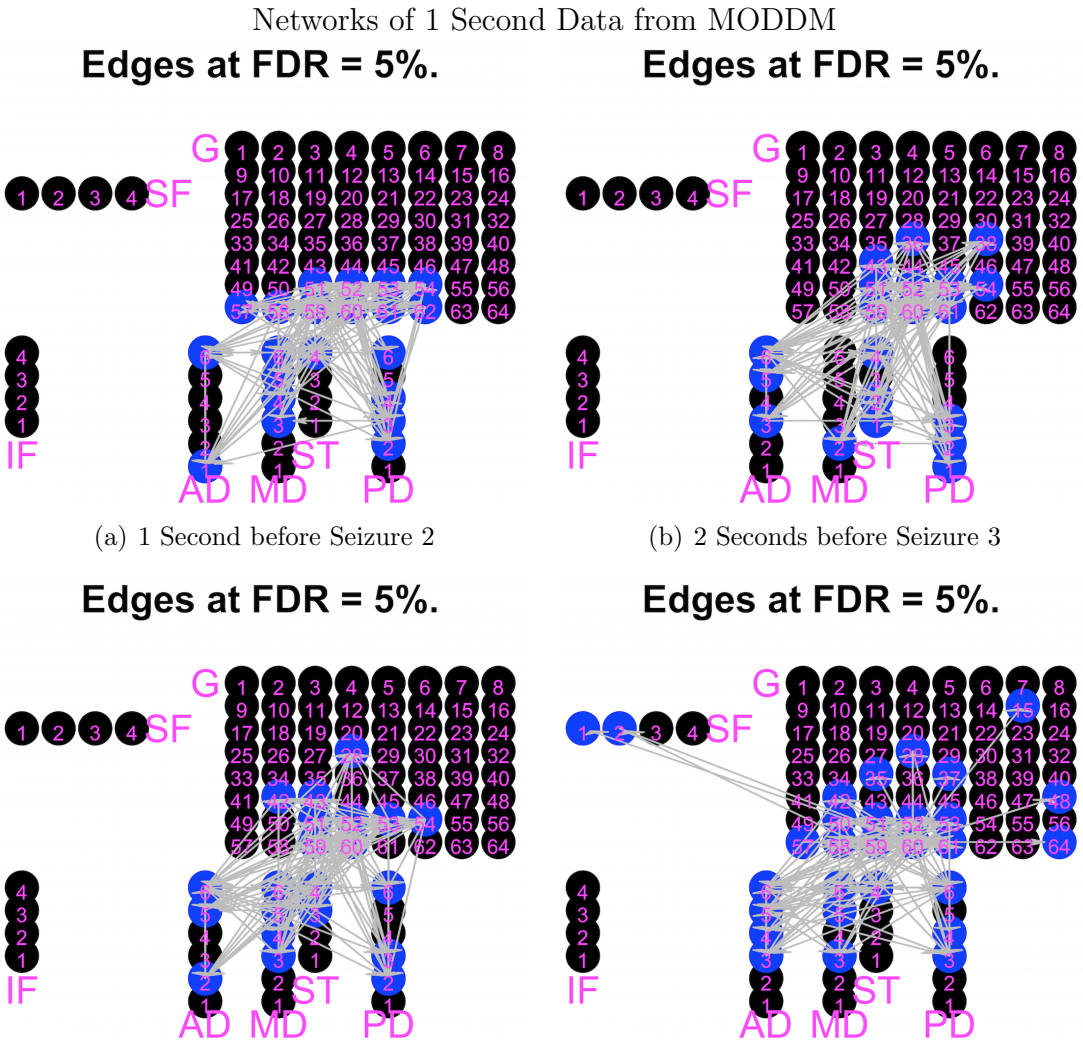
We focus on iEEG data 20 seconds around seizure onset times (10 seconds before and 10 seconds after seizure onset times) at the region G44. Following (Burns *et al.*, 2014), we divided iEEG time series into segments of 1-second length and applied both the MODDM and SWODDM to each segment independently. We selected the directional network edges using the FDR of 5% so that the ensuring networks have sparse edges and are scientifically interpretable. To compare the clustering results of MODDM and SWODDM, we present the summary results of the MODDM and the SWODDM respectively.

Figure 5.1 and Figure 5.2 show brain networks around seizure onsets of the 2nd, 3rd, 4th and 6th recorded seizures from MODDM and SWODDM respectively. Nodes in black correspond to brain regions that are disconnected from other regions, and nodes in blue correspond to brain regions in the same cluster. Directional network edges in grey indicate directional connections among regions in the

same cluster. Directional network edges in green indicate directional connections among regions in different clusters. Based on the analysis results of 1-second segments around seizure onsets, we found that brain regions within the temporal lobe, including electrodes G50-G54, G58-G62, AD03-AD06, MD03-MD06, ST03-ST04, and PD03-PD06, are constantly connected with each other.

The brain network results based on 1-second segments have large variability, most likely due to the limited data (200 time points for each segment) and the enormous model parameters to be estimated. To get stable network results, we combine data information across seizures by taking the average of posterior probabilities across 4 seizures and selecting network edges by using the FDR of 5% for average posterior probabilities. As such, we obtained 20 average brain networks within -10 to 10 second window centered at seizure onset (one average for each 1-second segment). As an illustration, Figures 5.3 and 5.4 show the average networks around seizure onset from MODDM and SWODDM respectively.

Based on average network results, we found that (1) brain regions in the temporal lobe had the strongest connections with each other, and (2) the SOZ G44 was disconnected from the rest of the regions, including its neighboring regions, around the seizure onset times except 1 second before the seizure onset time (from SWODDM only). The result of isolated SOZ from the rest of regions around seizure onset is in line with the prior network research by (Burns *et al.*, 2014; Nissen *et al.*, 2016; Park and Friston, 2013). And Figure 5.5 shows the change of SOZ G44 connection with the rest of the regions along with time around the seizure onset zone. 0 indicates G44 is disconnected with the rest of the regions while 1 indicates G44 is connected with the rest of the regions. It is noted that based on the average network results from SWODDM, G44 was disconnected from the rest of the regions until 1 second before the seizure onset and connected with its neighbours at the seizure onset and



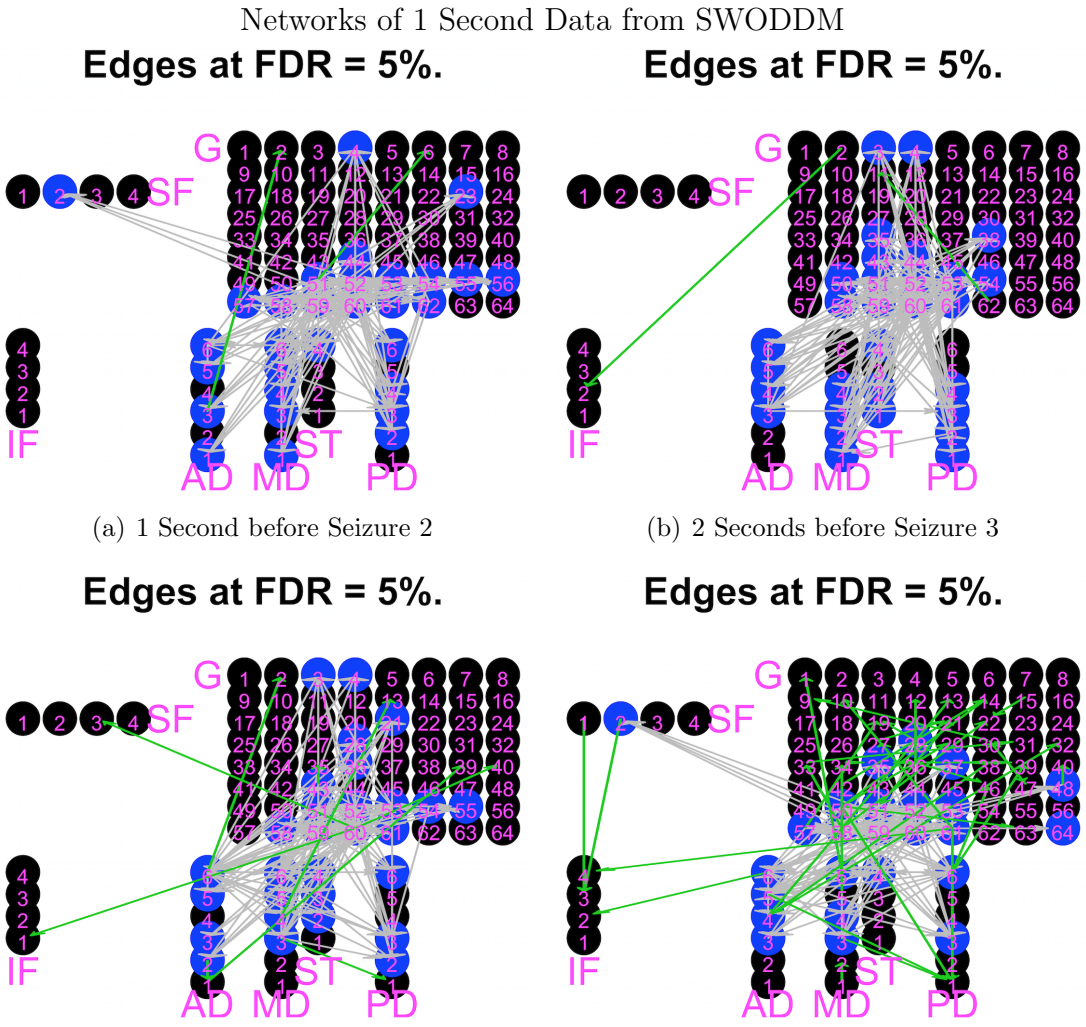
(a) 1 Second before Seizure 2

(b) 2 Seconds before Seizure 3

(c) 1 Second after Seizure 4

(d) 2 Seconds after Seizure 6

Figure 5.1: 5.1(a) The brain network at 1 second before the 2nd seizure onset. 5.1(b) The brain network at 2 seconds before the 3rd seizure onset. 5.1(c) The brain network at 1 second after the 4th seizure onset. 5.1(d) The brain network at 2 seconds after the 6th seizure onset. Nodes in black correspond to the regions that are isolated from the rest regions and form a cluster having only one region. Nodes in blue correspond to the regions in the same cluster. The directional network edges in grey are for the pairs of regions in the same cluster with their network edges selected with the FDR of 5%.



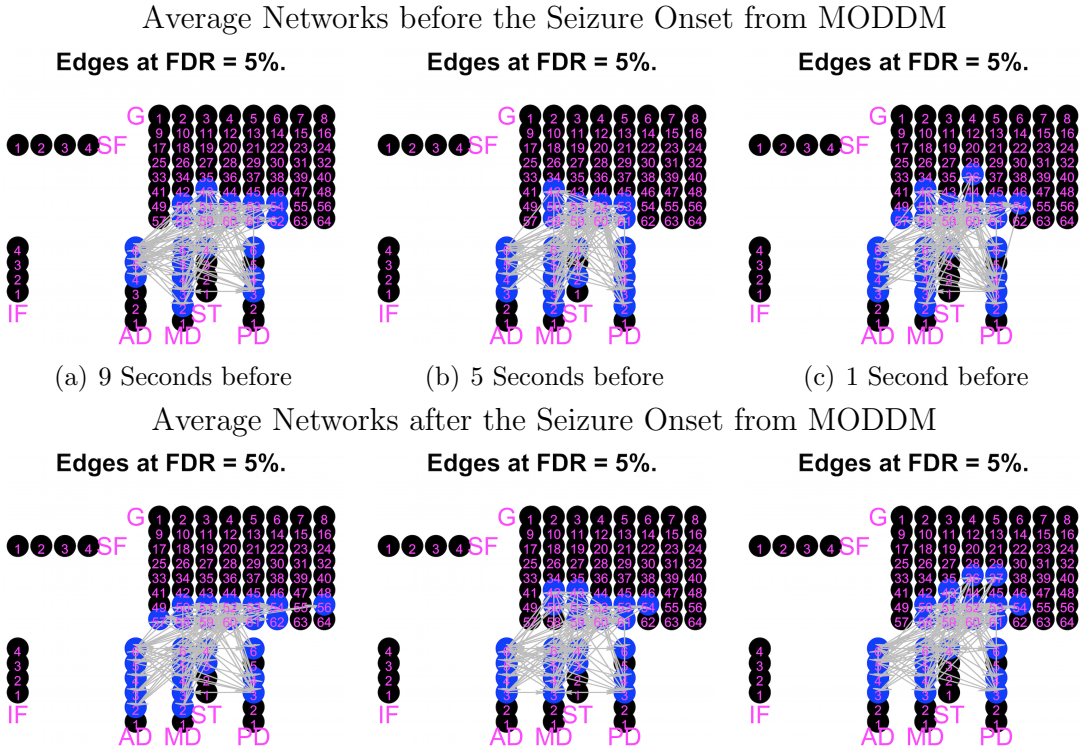
(a) 1 Second before Seizure 2

(b) 2 Seconds before Seizure 3

(c) 1 Second after Seizure 4

(d) 2 Seconds after Seizure 6

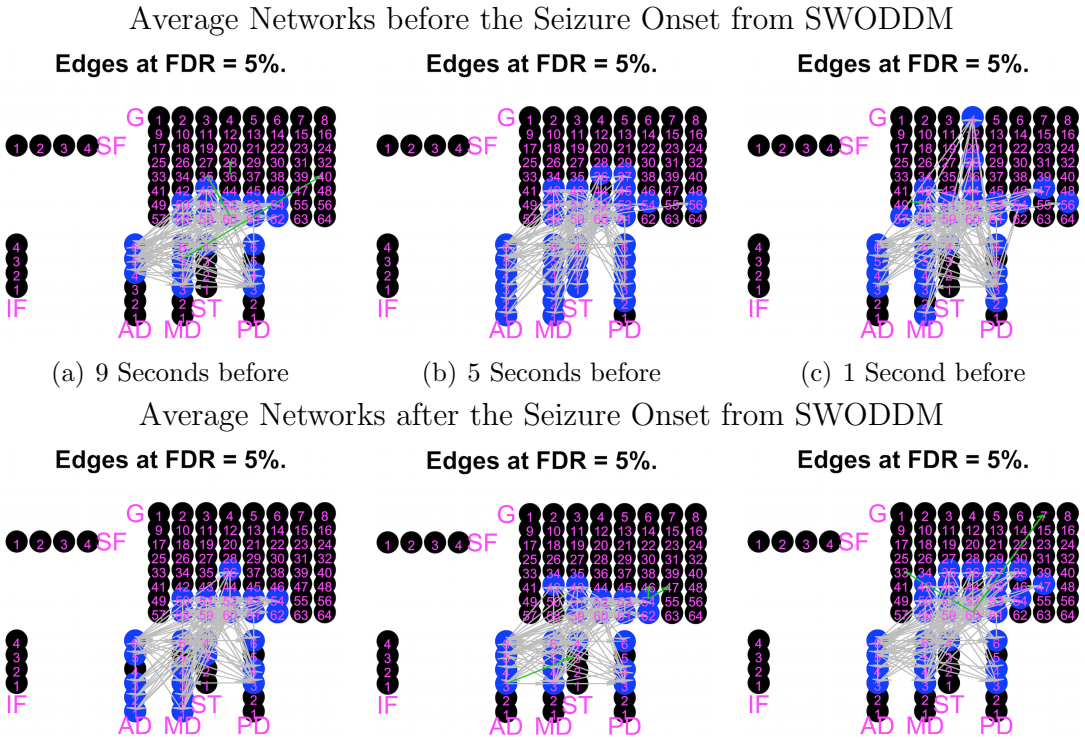
Figure 5.2: 5.2(a) The brain network at 1 second before the 2nd seizure onset. 5.2(b) The brain network at 2 seconds before the 3rd seizure onset. 5.2(c) The brain network at 1 second after the 4th seizure onset. 5.1(d) The brain network at 2 seconds after the 6th seizure onset. Nodes in black correspond to the regions that are isolated from the rest regions and form a cluster having only one region. Nodes in blue correspond to the regions in the same cluster. The directional network edges in grey are for the pairs of regions in the same cluster with their network edges selected with the FDR of 5%. The directional network edges in green are for the pairs of regions between different clusters with their network edges selected with the FDR of 5%.



(a) 9 Seconds before (b) 5 Seconds before (c) 1 Second before

(d) 1 Second after (e) 5 Seconds after (f) 9 Seconds after

Figure 5.3: 5.3(a), 5.3(b) and 5.3(c): The average networks across 4 seizures at 9 seconds, 5 seconds and 1 second before seizure onset, respectively. 5.3(e), 5.3(f) and 5.3(c): The average networks across 4 seizures at 1 second, 5 seconds and 9 seconds before seizure onset, respectively. Nodes in black correspond to the regions that are isolated from the rest regions and form a cluster having only one region. Nodes in blue correspond to the regions in the same cluster. The directional network edges in grey are for the pairs of regions in the same cluster with their network edges selected with the FDR of 5%.



(a) 9 Seconds before (b) 5 Seconds before (c) 1 Second before

(d) 1 Second after (e) 5 Seconds after (f) 9 Seconds after

Figure 5.4: 5.4(a), 5.4(b) and 5.4(c): The average networks across 4 seizures at 9 seconds, 5 seconds and 1 second before seizure onset, respectively. 5.4(d), 5.4(e) and 5.4(f): The average networks across 4 seizures at 1 second, 5 seconds and 9 seconds after seizure onset, respectively. Nodes in black correspond to the regions that are isolated from the rest regions and form a cluster having only one region. Nodes in blue correspond to the regions in the same cluster. The directional network edges in grey are for the pairs of regions in the same cluster with their network edges selected with the FDR of 5%. The directional network edges in green are for the pairs of regions between different clusters with their network edges selected with the FDR of 5%.

then disconnected again after the seizure onset. We can conclude that SWODDM is able to identify the connection change of SOZ G44 at the seizure onset time while MODDM cannot (20 consecutive average network results from MODDM and SWODDM are attached in Appendix 7.3). We believe that the unique connectivity property of the SOZ detected by the proposed method will be helpful for clinicians to locate the SOZ in practice, which will be the focus of the future research.

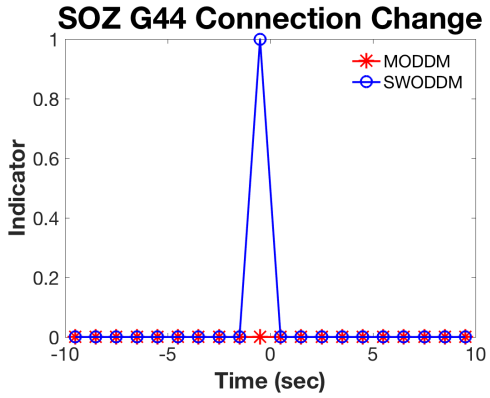
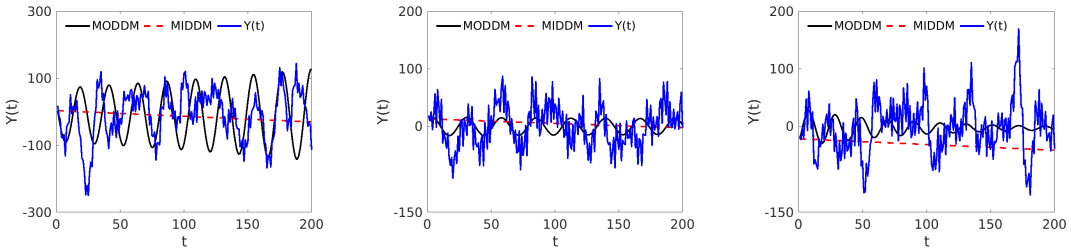


Figure 5.5: SOZ G44 connection status change along with time. 0 indicates G44 is disconnected with the rest of the regions while 1 indicates G44 is connected with the rest of the regions. 0s is the seizure onset time.

To demonstrate that the proposed ODDM for iEEG data can better characterize the brain's oscillatory activity than the existing first-order ODE model MIDDM (Zhang *et al.*, 2017), we used the posterior estimates (posterior medians) of the model parameters to regenerate the state functions, which was the same approach used in the simulation study. Figure 5.6 shows the regenerated $\hat{\mathbf{x}}(t)$ by the MIDDM and MODDM, in comparison to the observed data $\mathbf{Y}(t)$. The state functions from the MIDDM are linear over time, thus, having zero oscillatory frequencies. In contrast, the MODDM provides a substantially better fit to the oscillatory patterns of $\mathbf{Y}(t)$.



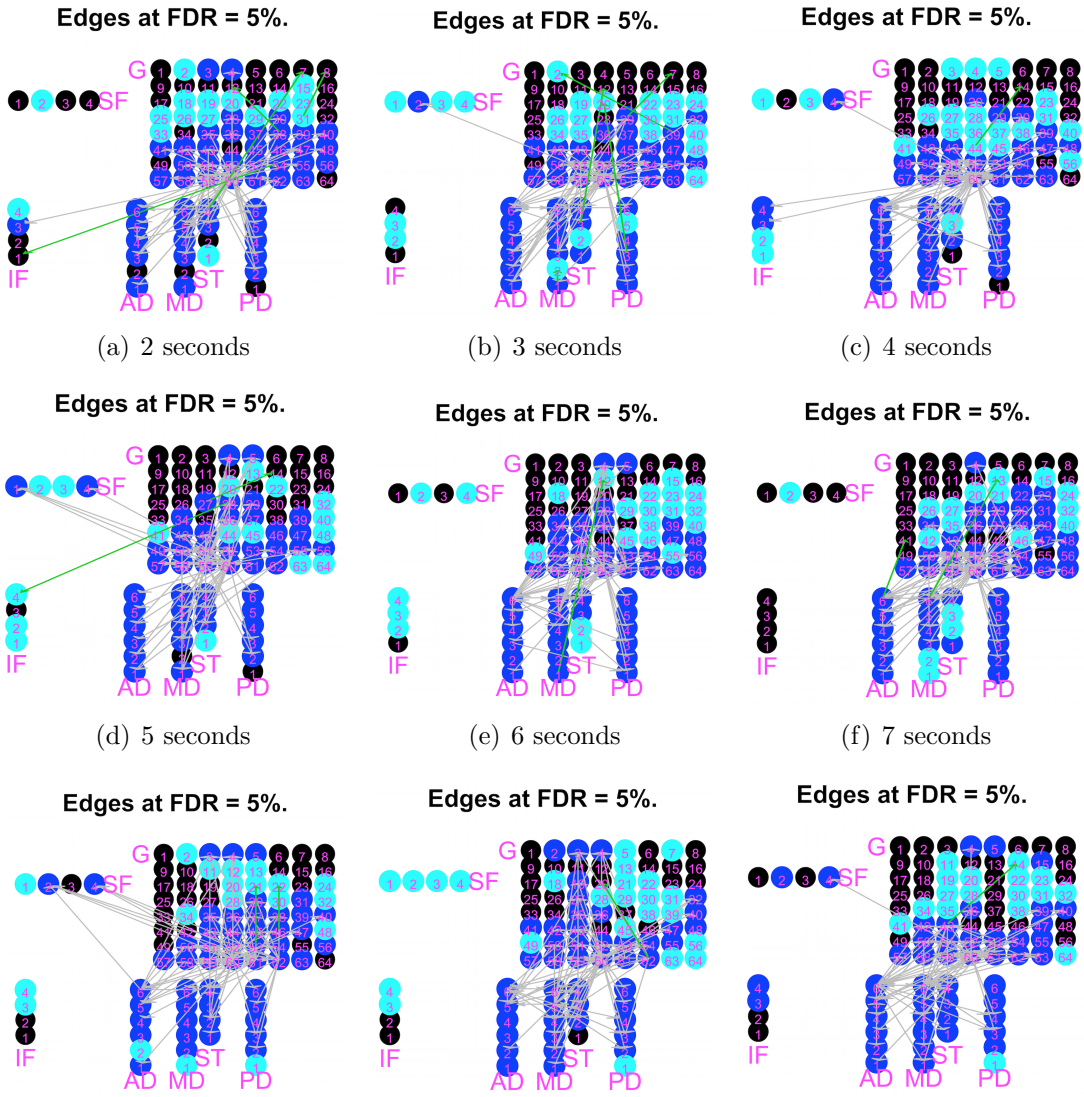
(a) Regenerated $\hat{\mathbf{x}}(t)$ of region G1 (b) Regenerated $\hat{\mathbf{x}}(t)$ of region G42 (c) Regenerated $\hat{\mathbf{x}}(t)$ of region G44

Figure 5.6: Regenerated $\hat{\mathbf{x}}(t)$ using the posterior estimates of MODDM and MIDDM parameters for regions G1, G42 and G44.

A Little More on the Cluster Identification

We proposed the thresholding method to identify clusters for both the MODDM and SWODDM in Section 3.4, i.e., region i and region j are classified into same the cluster when $\hat{P}_{ij}^m$ $(= \hat{P}_{ji}^m)$ is larger than the predetermined threshold, which is proved to be successful in simulation studies. But it is not hard to see the thresholding is a rigorous condition that might go wrong under certain circumstances.

Specifically, suppose region i and region j are in two different clusters, the thresholding method will group the two clusters which region i and region j are in into one when the posterior clustering probabilities $\hat{P}_{ij}^m$ is overestimated, say larger than the predetermined threshold. That means a single error in estimation of the posterior clustering probabilities P_{ij}^m can distort the network results. Thus, we also applied the community (cluster) extraction method proposed by (Zhao *et al.*, 2011) on posterior clustering probabilities to identify the clusters. The idea of this community extraction method to identify clusters is based on maximizing the community extraction criterion, which incorporates two principles of clustering, dense connections within the same cluster and sparse connections between different clusters. Basically, it calculates the extraction criterion on a given adjacency matrix which represents

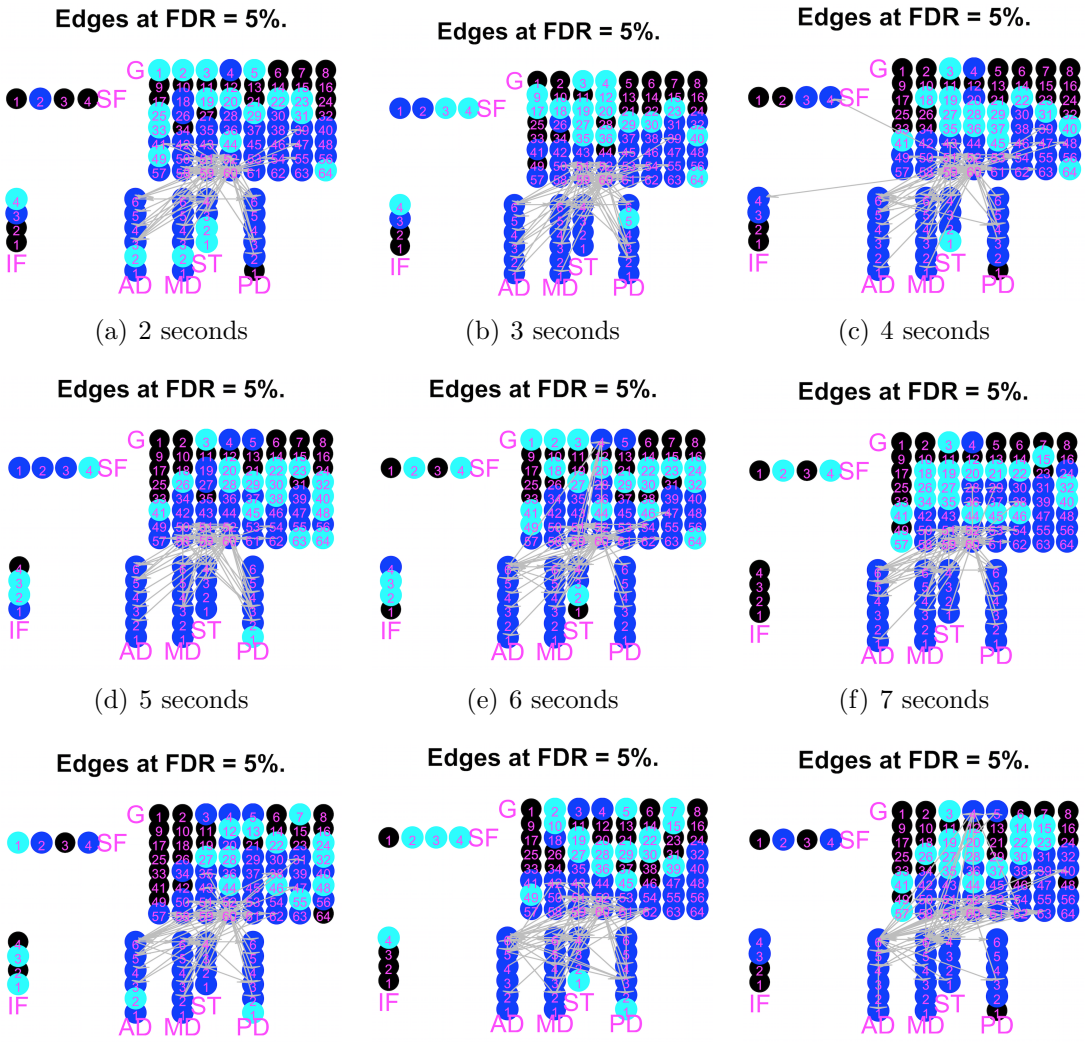


(a) 2 seconds (b) 3 seconds (c) 4 seconds

(d) 5 seconds (e) 6 seconds (f) 7 seconds

(g) 8 seconds (h) 9 seconds (i) 10 seconds

Figure 5.7: The SWODDM brain networks across 9 consecutive one-second segments before the 2nd seizure onset with the clusters identified by the community extraction method. Nodes in blue/cyan correspond to the regions in two different clusters, and nodes in black correspond to the isolated regions.



(a) 2 seconds (b) 3 seconds (c) 4 seconds

(d) 5 seconds (e) 6 seconds (f) 7 seconds

(g) 8 seconds (h) 9 seconds (i) 10 seconds

Figure 5.8: The MODDM brain networks across 9 consecutive one-second segments before the 2nd seizure onset with the clusters identified by the community extraction method. Nodes in blue/cyan correspond to the regions in two different clusters, and nodes in black correspond to the isolated regions.

the effective connectivity among all components and then find the cluster scheme that optimizes the extraction criterion. The adjacency matrix here could be posterior network edge selection probabilities from the MODDM and SWODDM. The problem of overestimation on posterior clustering probabilities of a few edges will be mediated by the optimzation algorithm.

SWODDM is a better candidate than MODDM to identify the clusters by the community extraction method because it characterizes both the within-cluster and between-cluster connections, consistent with the optimization criterion in the community extraction method. Thus we use the posterior network edge selection probabilities from SWODDM, i.e., $\hat{P}_{ij}^{\gamma^W} + \hat{P}_{ij}^{\gamma^B}$, as the adjacency matrix for the community extraction method. Figure 5.7 shows the brain networks of 9 consecutive one-second segments before the seizure onset of the 2nd recorded seizure (2-10 seconds before the seizure onset respectively), identified by the community extraction method (for clusters), the 5% FDR thresholding (for within-cluster connections, $\hat{P}_{ij}^{\gamma^W}$s) and a thresholding with a high value (for between-cluster connections, $\hat{P}_{ij}^{\gamma^B}$s). Nodes in blue/cyan correspond to the regions in two different clusters, and nodes in black correspond to the isolated regions. Grey arrows correspond to connections within the same cluster and green arrows correspond to connections between different clusters.

Based on the network results from SWODDM, we found that G44 is always isolated from its surrounding brain regions for 9 consecutive seconds (except at 8 seconds before the seizure onset), either in a different cluster with its surrounding brain regions or completely isolated from the other regions. It is consistent with previous findings from the MODDM network results but more robust. As a comparison, Figure 5.8 shows the brain networks of the same 9 segments from MODDM. The clusters are identified by the community extraction method with an adjacency matrix of posterior network edge selection probabilities, i.e., $\hat{P}_{ij}^{\gamma}$, and the within-cluster

connections are selected with a 5% FDR threshold for $\hat{P}_{ij}^{\gamma}$s. The network results are similar with SWODDM (G44 isolated from the other surrounding brain regions most of the time) but less consistent across 9 consecutive seconds — G44 connected with its surrounding brain regions in 3 one-second segments (4, 5, 9 seconds before the seizure onset respectively). That indicates the between-cluster connections not captured by MODDM may disturb its clustering/network results.

Chapter 6

Discussion

We have developed new ODE models, ODDM, for directional connectivity among many brain regions recorded by iEEG. The ODDM incorporates a physical mechanism (i.e., damped harmonic oscillator) to characterize the brain's oscillation, and uses a linear form to approximate the underlying directional interactions among regions. As such, the ODDM combines the strengths of scientific modeling and statistical modeling. We have shown through both simulation study and real data analysis that the new model ODDM outperforms the existing ODE model by providing a substantially better fit to multivariate oscillatory iEEG data and detecting connected regions with much higher accuracy. We applied the developed model and Bayesian method to an epileptic patient's iEEG data and examined the patient's brain network. The analysis results revealed that the SOZ tends to be disconnected from other regions in the brain network around the time of seizure onset. This unique connectivity property of the SOZ can be used to identify the SOZ among many regions recorded by iEEG. Our method has a great potential to enhance understanding of epileptic brain networks, increase the accuracy in SOZ localization, and ultimately improve epilepsy diagnosis and treatment.

Despite that ODDM provides a much better fit to the oscillatory time series data than existing network models/methods, there remains data variation that the ODDM cannot explain. There are possibly two reasons. First, the ODDM is a linear approximation of the brain system whose dynamic mechanism is highly complex and mostly unknown. In the literature, the existing ODE models that can fit the brain data well all deal with low-dimensional data and usually use more ODEs than the number of time series to fit the data. We here deal with a significantly more challenging problem: building an ODE model that is in the same high dimension as the number of time series/regions to explain all the regions' activities. Consequently, there is a considerable discrepancy between the ODDM and the true underlying system. Second, like linear regressions, the use of linear expansion in the ODDM to approximate the underlying system is effective only for a short period. We, therefore, applied the ODDM to short data segments independently. With limited data information, the model parameter estimates also have large variances. As shown in Simulation Study, even if the assumed ODDM is a true model, the estimated model parameters cannot reproduce exactly the underlying state functions. Because of these reasons, the traditional evaluation of the model fitting to the data may not be appropriate for the problem under study. Nonetheless, we have shown through both simulation and real data analysis that the ODDM has much better efficiency in detecting connected brain regions than the existing ODE model for iEEG data. Research is greatly needed to develop highly efficiency statistical models and methods to provide a better fit to high-dimensional brain data.

The ODDM combines the strength of physical modeling and statistical modeling, expanded from one-dimensional oscillatory physical system to the high-dimensional brain system through a linear approximation. The oscillatory brain activities of each brain region are characterized by a one-dimensional DHO, which is a widely

used physical model for describing one-dimensional oscillatoryphysical system. It is based on a deterministic ODE model, in which no model errors are assumed to account for the discrepancy between the brain's underlying mechanism and the DHO. This is mainly because the physical models are believed to be the golden truth and there is no randomness in the golden truth. So to maintain the characteristics of physical modeling, we assumed no model errors in ODDM. However, we estimate the model parameters within a Bayesian framework so as to allow variabilities in the parameters, which means we can still account for the model uncertainty even if we didn't assume model errors. In addition, the DHO with an error term is no longer a ODE model but a stochastic differential equation (SDE) model, which is a stochastic model instead of a deterministic one. In this setting, the external force or say directional effect from other brain regions to one brain region will be considered as a random force/effect and the underlying neuronal state at each brain region is a random process. Considering the current understanding of nervous system, we stick with ODE to describe the brain's mechanism and characterize the directional connectivities among different brain regions.

Furthermore, the MODDM assumes no connection between cluster and is focused on within-cluster connections only. Thus, it is easier to detect dense connections within the same cluster, and the MODDM is suitable for iEEG data that are usually collected from spatially close regions. We extend the MODDM to SWODDM to accommodate long-range, sparse connections between clusters by introducing additional indicators for between-cluster connections. Under this model, each network edge can be either within-cluster or between-cluster connection. These two types of connections may have different densities and play different functional roles in the brain network. The within-cluster connections are those between the regions with a similar function while the between-cluster connections ensure integration among re-

gions specializing in different functions. We have demonstrated the advantage of the SWODDM over MODDM by simulation studies on detecting the clusters in various systems with different structures of between-cluster connections. The SWODDM is able to detect the true clusters in high-dimensional systems regardless the existence and the complexity of between-cluster connections. In addition, the SWODDM is able to select significant between-cluster connections with a high TPR and a low FPR. However, the FDR of between-cluster connections is hard to control due to the sparsity of between-cluster connections. Even for a small FPR, the number of false positives could be large because of a large amount of negatives. From this perspective, the between-cluster connection selection is essentially a data imbalance problem. How to borrow information from other sources, both internal and external, to regulate the estimation of between-cluster connectivities is an important topic in the future research.

A part of this work has been published in *Computational Statistics & Data Analysis* (Zhang *et al.*, 2020).

Chapter 7

Appendix

7.1 Appendix A: Technical Details for PCGS Algorithm to Sample from the Posterior Distribution.

A.I Derive the joint posterior distribution $p(\boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}|\boldsymbol{\eta})$

In the following, we use $p(\theta|-)$ to denote the full posterior conditional distribution of θ. Based on the formulation of the joint distribution (3.12), given the rest of the parameters, $\{A_{ij}, D_i, G_i, j = 1, \ldots, d\}$ are independent for $i = 1, \ldots, d$, so we will first derive the posterior conditional distribution of $\{A_{ij}, D_i, G_i, j = 1, \ldots, d\}$.

We use $\mathbf{M}[, \mathbf{I}]$ to denote the submatrix consisting of columns indexed by $\mathbf{I}$ of $\mathbf{M}$, and $\mathbf{M}[\mathbf{I},]$ to denote the submatrix consisting of rows indexed by $\mathbf{I}$ of $\mathbf{M}$. Let $\mathcal{G}_i = \{j, \delta(m_i, m_j) \cdot \gamma_{ij} \neq 0 \text{ and } j = 1, \ldots, d\}$ for MODDM or $\mathcal{G}_i = \{j, \delta(m_i, m_j) \cdot \gamma_{ij}^W + (1 - \delta(m_i, m_j)) \cdot \gamma_{ij}^B \neq 0 \text{ and } j = 1, \ldots, d\}$ for SWODDM. Define a $d \times d$ diagonal matrix $\mathbf{I}_i$ where diagonal entries corresponding to $\mathcal{G}_i$ equal 1, and the rest diagonal entries equal 0. Let $\mathbf{Z}_i(t) = \mathbf{I}_i \ \mathbf{x}(t)$, and $\Lambda_i(t) = (\mathbf{Z}_i(t)', 1, dx_i(t)/dt)' =$

$\left(\mathbf{Z}_i(t)', 1, \tilde{\eta}_i' \; \boldsymbol{b}^{(1)}(t)\right)'$, so $\mathbf{Z}_i(t)$ and $\Lambda_i(t)$ are vectors whose elements are functions of time t. Also, we let $\boldsymbol{\theta}_i = (\mathbf{A}[i, \mathcal{G}_i], D_i, G_i)'$. Since

$$p\left(\mathbf{A}[i,], D_i, G_i|-\right) \propto \exp\left\{-\frac{1}{2\tau_i} \int_0^T \left(\Lambda_i\left(t\right)\boldsymbol{\theta}_i - \frac{d^2 x_i\left(t\right)}{dt^2}\right)^2 dt\right\} \cdot \prod_{j=1}^{d} \phi\left(\frac{A_{ij}}{\xi_0}\right) \cdot \phi\left(\frac{D_i}{\xi_0}\right) \cdot \phi\left(\frac{G_i}{\xi_0}\right),$$

where $d^2 x_i\left(t\right)/dt^2 = \tilde{\eta}_i' \boldsymbol{b}^{(2)}(t)$. After integrating out A_{ij} corresponding to zero indicator values in the above equation, we have

$$\begin{aligned} p\left(\boldsymbol{\theta}_i|-\right) \;\; \propto \;\; & \exp\left\{-\frac{1}{2}\boldsymbol{\theta}_i' \left(\frac{1}{\tau_i}\int_0^T \Lambda_i\left(t\right)\Lambda_i'\left(t\right)dt + \frac{1}{\xi_0^2}\mathbb{I}\right)\boldsymbol{\theta}_i + \frac{1}{\tau_i}\int_0^T \frac{d^2 x_i\left(t\right)}{dt^2}\Lambda_i'\left(t\right)dt \; \boldsymbol{\theta}_i\right\} \cdot \\ & \cdot \exp\left\{-\frac{1}{2\tau_i}\int_0^T \left(\frac{d^2 x_i\left(t\right)}{dt^2}\right)^2 dt\right\}, \end{aligned} \tag{7.1}$$

where $\mathbb{I}$ denotes an identity matrix.

Let $\mathbb{M}_i = \frac{1}{\tau_i}\int_0^T \Lambda_i'\left(t\right)\Lambda_i\left(t\right)dt + \frac{1}{\xi_0^2}\mathbb{I}$ and $\mathbb{V}_i = \frac{1}{\tau_i}\int_0^T \frac{d^2 x_i(t)}{dt^2} \cdot \Lambda_i'\left(t\right)dt$. Based on the equation (7.1), the posterior joint distribution after integrating out $\mathbf{A}$, $\mathbf{D}$, and $\mathbf{G}$ is

$$\begin{aligned} p\left(\boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}|\boldsymbol{\eta}\right) \;\; \propto \;\; & \prod_{i=1}^{d} \det\left(\mathbb{M}_i\right)^{-1/2} \cdot \exp\left\{\sum_{i=1}^{d} \frac{\mathbb{V}_i' \mathbb{M}_i^{-1} \mathbb{V}_i}{2}\right\} \cdot \exp\left\{-\sum_{i=1}^{d}\int_0^T \frac{1}{2\tau_i}\left(\frac{d^2 x_i\left(t\right)}{dt^2}\right)^2 dt\right\} \\ & \cdot \exp\left\{-\mu \sum_{i,j=1}^{d} \delta\left(m_i, m_j\right)\right\} \\ & \cdot p_0^{\sum_{i,j}\gamma_{ij}} \cdot \left(1-p_0\right)^{d^2 - \sum_{i,j}\gamma_{ij}} \;\; \text{(MODDM)}, \end{aligned}$$

$$\begin{aligned} p\left(\boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}|\boldsymbol{\eta}\right) \;\; \propto \;\; & \prod_{i=1}^{d} \det\left(\mathbb{M}_i\right)^{-1/2} \cdot \exp\left\{\sum_{i=1}^{d} \frac{\mathbb{V}_i' \mathbb{M}_i^{-1} \mathbb{V}_i}{2}\right\} \cdot \exp\left\{-\sum_{i=1}^{d}\int_0^T \frac{1}{2\tau_i}\left(\frac{d^2 x_i\left(t\right)}{dt^2}\right)^2 dt\right\} \\ & \cdot \exp\left\{-\mu \sum_{i,j=1}^{d} \delta\left(m_i, m_j\right)\right\} \\ & \cdot p_w^{\sum_{i,j}\gamma_{ij}^W} \cdot \left(1-p_w\right)^{d^2 - \sum_{i,j}\gamma_{ij}^W} \cdot p_b^{\sum_{i,j}\gamma_{ij}^B} \cdot \left(1-p_b\right)^{d^2 - \sum_{i,j}\gamma_{ij}^B} \;\; \text{(SWODDM)}. \end{aligned}$$

From the above equation, we have $p(\boldsymbol{m}, \boldsymbol{\gamma}|\boldsymbol{\tau}, \boldsymbol{\eta}) \propto \mathbb{J}(\boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta})$, where

$$\begin{aligned}\mathbb{J}(\boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta}) &= \prod_{i=1}^{d} \det(\mathbb{M}_i)^{-1/2} \cdot \exp\left\{\sum_{i=1}^{d} \frac{\mathbb{V}_i' \mathbb{M}_i^{-1} \mathbb{V}_i}{2} - \mu \sum_{i,j=1}^{d} \delta(m_i, m_j)\right\} \\ &\quad \cdot p_0^{\sum_{i,j} \gamma_{ij}} \cdot (1-p_0)^{d^2 - \sum_{i,j} \gamma_{ij}}\end{aligned}$$

for MODDM, and

$$\begin{aligned}\mathbb{J}(\boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta}) &= \prod_{i=1}^{d} \det(\mathbb{M}_i)^{-1/2} \cdot \exp\left\{\sum_{i=1}^{d} \frac{\mathbb{V}_i' \mathbb{M}_i^{-1} \mathbb{V}_i}{2} - \mu \sum_{i,j=1}^{d} \delta(m_i, m_j)\right\} \\ &\quad \cdot p_w^{\sum_{i,j} \gamma_{ij}^W} \cdot (1-p_w)^{d^2 - \sum_{i,j} \gamma_{ij}^W} \cdot p_b^{\sum_{i,j} \gamma_{ij}^B} \cdot (1-p_b)^{d^2 - \sum_{i,j} \gamma_{ij}^B}\end{aligned}$$

for SWODDM ($\boldsymbol{\gamma} = \{\boldsymbol{\gamma}^W, \boldsymbol{\gamma}^B\}$).

A.II Sequentially simulate m_i from $p(m_i|\boldsymbol{m}_{-i}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta})$ for $i = 1, \ldots, d$.

Let $\boldsymbol{V}_{-i}$ be the set of distinct values in $\boldsymbol{m}_{-i}$, and v_{-i} be any positive integer smaller than $d+1$ and not belonging to $\boldsymbol{V}_{-i}$. Then the posterior conditional distribution of m_i is discrete and has a support of $\{\boldsymbol{V}_{-i}, v_{-i}\}$. In addition, for each $z \in \{\boldsymbol{V}_{-i}, v_{-i}\}$,

$$p(m_i = z|\boldsymbol{m}_{-i}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta}) \propto \mathbb{J}(m_i = z, \boldsymbol{m}_{-i}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta}).$$

A.III Sequentially simulate γ_{ij}s from $p(\gamma_{ij}|\boldsymbol{m}, \boldsymbol{\gamma}_{-ij}, \boldsymbol{\tau}, \boldsymbol{\eta})$ for $i, j = 1, \ldots, d$ (MODDM), or Sequentially simulate γ_{ij}^Ws from $p\left(\gamma_{ij}^W|\boldsymbol{m}, \boldsymbol{\gamma}_{-ij}^W, \boldsymbol{\gamma}^B, \boldsymbol{\tau}, \boldsymbol{\eta}\right)$ and γ_{ij}^Bs from $p\left(\gamma_{ij}^B|\boldsymbol{m}, \boldsymbol{\gamma}^W, \boldsymbol{\gamma}_{-ij}^B, \boldsymbol{\tau}, \boldsymbol{\eta}\right)$ for $i, j = 1, \ldots, d$ (SWODDM).

For MODDM, given parameter values $\boldsymbol{m}, \boldsymbol{\gamma}_{-ij}, \boldsymbol{\tau}$ and $\boldsymbol{\eta}$, γ_{ij} for $i, j = 1, \ldots, d$ follows a Bernoulli distribution with probability

$$\frac{\mathbb{J}(\boldsymbol{m}, \gamma_{ij} = 1, \boldsymbol{\gamma}_{-ij}, \boldsymbol{\tau}, \boldsymbol{\eta})}{\mathbb{J}(\boldsymbol{m}, \gamma_{ij} = 1, \boldsymbol{\gamma}_{-ij}, \boldsymbol{\tau}, \boldsymbol{\eta}) + \mathbb{J}(\boldsymbol{m}, \gamma_{ij} = 0, \boldsymbol{\gamma}_{-ij}, \boldsymbol{\tau}, \boldsymbol{\eta})}.$$

That the above probability equals p_0 if $m_i \neq m_j$. This is because if $m_i \neq m_j$, $\delta(m_i, m_j) = 0$, and the value of γ_{ij} in the model (2.5) does not affect the model fitting. Thus, in this case, the posterior distribution of γ_{ij} is not affected by the data and is the same as the prior distribution.

Similarly for SWODDM, given parameter values $\boldsymbol{m}, \boldsymbol{\gamma}^W_{-ij}, \boldsymbol{\gamma}^B, \boldsymbol{\tau}$ and $\boldsymbol{\eta}$, γ^W_{ij} for $i, j = 1, \ldots, d$ follows a Bernoulli distribution with probability

$$\frac{\mathbb{J}\left(\boldsymbol{m}, \gamma^W_{ij} = 1, \boldsymbol{\gamma}^W_{-ij}, \boldsymbol{\gamma}^B, \boldsymbol{\tau}, \boldsymbol{\eta}\right)}{\mathbb{J}\left(\boldsymbol{m}, \gamma^W_{ij} = 1, \boldsymbol{\gamma}^W_{-ij}, \boldsymbol{\gamma}^B, \boldsymbol{\tau}, \boldsymbol{\eta}\right) + \mathbb{J}\left(\boldsymbol{m}, \gamma^W_{ij} = 0, \boldsymbol{\gamma}^W_{-ij}, \boldsymbol{\gamma}^B, \boldsymbol{\tau}, \boldsymbol{\eta}\right)},$$

which equals p_w if $\delta(m_i, m_j) = 0$; and given parameter values $\boldsymbol{m}, \boldsymbol{\gamma}^W, \boldsymbol{\gamma}^B_{-ij}, \boldsymbol{\tau}$ and $\boldsymbol{\eta}$, γ^B_{ij} for $i, j = 1, \ldots, d$ follows a Bernoulli distribution with probability

$$\frac{\mathbb{J}\left(\boldsymbol{m}, \boldsymbol{\gamma}^W, \gamma^B_{ij} = 1, \boldsymbol{\gamma}^B_{-ij}, \boldsymbol{\tau}, \boldsymbol{\eta}\right)}{\mathbb{J}\left(\boldsymbol{m}, \boldsymbol{\gamma}^W, \gamma^B_{ij} = 1, \boldsymbol{\gamma}^B_{-ij}, \boldsymbol{\tau}, \boldsymbol{\eta}\right) + \mathbb{J}\left(\boldsymbol{m}, \boldsymbol{\gamma}^W, \gamma^B_{ij} = 0, \boldsymbol{\gamma}^B_{-ij}, \boldsymbol{\tau}, \boldsymbol{\eta}\right)},$$

which equals p_b if $\delta(m_i, m_j) = 0$.

A.IV Simulate $\boldsymbol{\theta}$ from $p\left(\boldsymbol{\theta} | \boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta}\right)$

Based on the posterior conditional distribution of $\boldsymbol{\theta}_i$ (7.1),

$$A_{ij} | \delta\left(m_i, m_j\right) \cdot \gamma_{ij} = 0 \overset{i.i.d.}{\sim} \mathrm{N}\left(0, \xi^2\right) \text{ (MODDM)}$$

or

$$A_{ij} | \delta(m_i, m_j) \cdot \gamma^W_{ij} + (1 - \delta(m_i, m_j)) \cdot \gamma^B_{ij} = 0 \overset{i.i.d.}{\sim} \mathrm{N}\left(0, \xi^2\right) \text{ (SWODDM) for } i, j = 1, \ldots, d,$$

and

$$\boldsymbol{\theta}_i | \boldsymbol{m}, \boldsymbol{\gamma}, \boldsymbol{\tau}, \boldsymbol{\eta} \overset{ind}{\sim} \mathrm{MN}\left(\mathbb{M}_i^{-1} \mathbb{V}_i, \mathbb{M}_i^{-1}\right) \text{ (Both) for } i = 1, \ldots, d.$$

A.V Simulate $\boldsymbol{\tau}$ from $p\left(\boldsymbol{\tau}|\boldsymbol{\Theta}_I, \boldsymbol{\eta}\right)$

From the joint posterior distribution (3.12), we have

$$\tau_i|\boldsymbol{\Theta}_I, \boldsymbol{\eta} \overset{ind}{\sim} \text{Inv-Gamma}\left(\frac{1}{2}, \frac{\mathrm{R}_i(\boldsymbol{\eta}, \boldsymbol{\Theta}_I)}{2}\right) \quad \text{for} \quad i = 1, \ldots, d.$$

7.2 Appendix B. Proof of the Normal Distribution of the Data $p\left(\eta|\Theta_I, \tau\right)$

Based on the model on basis coefficients $\boldsymbol{\eta}$, we have

$$\begin{aligned} p\left(\boldsymbol{\eta}|\boldsymbol{\Theta}_I, \boldsymbol{\tau}\right) &\propto \exp\left\{-\sum_{i=1}^{d} \frac{\mathrm{R}_i\left(\boldsymbol{\eta}, \boldsymbol{\Theta}_I\right)}{2\tau_i}\right\} \propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{\eta}'\ \boldsymbol{\Omega}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\ \boldsymbol{\eta} - 2\boldsymbol{\Lambda}'_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\ \boldsymbol{\eta} + \boldsymbol{\Xi}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\right)\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{\eta} - \boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\right)' \boldsymbol{\Omega}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}} \left(\boldsymbol{\eta} - \boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\right)\right\}, \end{aligned} \tag{7.2}$$

Thus, from (7.2),

$$\boldsymbol{\eta}|\boldsymbol{\Theta}_I, \boldsymbol{\tau} \sim \text{MN}\left(\boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}, \boldsymbol{\Omega}^{-1}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}\right). \tag{7.3}$$

Notations of $\boldsymbol{\Omega}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}, \boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}$, and $\boldsymbol{\Xi}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}$ are introduced in equation (3.4), and we here derive their formulas conditioning on $\boldsymbol{\Theta}_I$ and $\boldsymbol{\tau}$ in the following.

Define vectors with $d \cdot L$ elements:

$$\begin{aligned} \boldsymbol{\Delta}_i\left(t\right) &= \Big(A_{i1} \cdot \delta\left(m_i, m_1\right) \cdot \gamma_{i1} \cdot b_1\left(t\right), \ldots, A_{i1} \cdot \delta\left(m_i, m_1\right) \cdot \gamma_{i1} \cdot b_L\left(t\right), \\ &\quad A_{i2} \cdot \delta\left(m_i, m_2\right) \cdot \gamma_{i2} \cdot b_1\left(t\right), \ldots, A_{id} \cdot \delta\left(m_i, m_d\right) \cdot \gamma_{id} \cdot b_L\left(t\right)\Big), \qquad (7.4) \\ \boldsymbol{E}_i\left(t\right) &= \left(\mathbf{0}_L, \ldots, \left(\frac{d^2\boldsymbol{b}\left(t\right)}{dt^2}\right)', \ldots, \mathbf{0}_L\right), \quad \boldsymbol{J}_i\left(t\right) = \left(\mathbf{0}_L, \ldots, \left(\frac{d\boldsymbol{b}\left(t\right)}{dt}\right)', \ldots, \mathbf{0}_L\right), \end{aligned}$$

where $\mathbf{0}_L$ is a zero vector with L elements, and the $(i-1)\cdot L+1$th to $i\cdot L$th elements of $\boldsymbol{E}_i(t)$ and $\boldsymbol{J}_i(t)$ are nonzero. With the basis representation (3.1), the MODDM (2.5) can be rewritten as $\boldsymbol{E}_i(t)\ \boldsymbol{\eta}-\boldsymbol{\Delta}_i(t)\ \boldsymbol{\eta}-D_i-G_i\cdot\boldsymbol{J}_i(t)\ \boldsymbol{\eta}=0$. Let $\mathbf{S}_i(t)=\boldsymbol{E}_i(t)-\boldsymbol{\Delta}_i(t)-G_i\cdot\boldsymbol{J}_i(t)$. We have

$$\sum_{i=1}^{d}\mathrm{R}_i(\boldsymbol{\eta},\boldsymbol{\Theta}_I)/\tau_i=\sum_{i=1}^{d}\frac{1}{\tau_i}\left(\boldsymbol{\eta}'\int_0^T\mathbf{S}_i'(t)\,\mathbf{S}_i(t)\,dt\ \boldsymbol{\eta}-2\int_0^T D_i\cdot\mathbf{S}_i(t)\,dt\ \boldsymbol{\eta}+\int_0^T D_i^2dt\right).$$

Comparing the above to equation (3.4), we have

$$\begin{aligned}\boldsymbol{\Omega}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}} &= \sum_{i=1}^{d}\frac{1}{\tau_i}\int_0^T\mathbf{S}_i'(t)\,\mathbf{S}_i(t)\,dt,\quad \boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}=\sum_{i=1}^{d}\frac{1}{\tau_i}\int_0^T D_i\cdot\mathbf{S}_i'(t)\,dt,\\ \boldsymbol{\Xi}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}} &= \sum_{i=1}^{d}\frac{1}{\tau_i}\int_0^T D_i^2dt.\end{aligned}$$

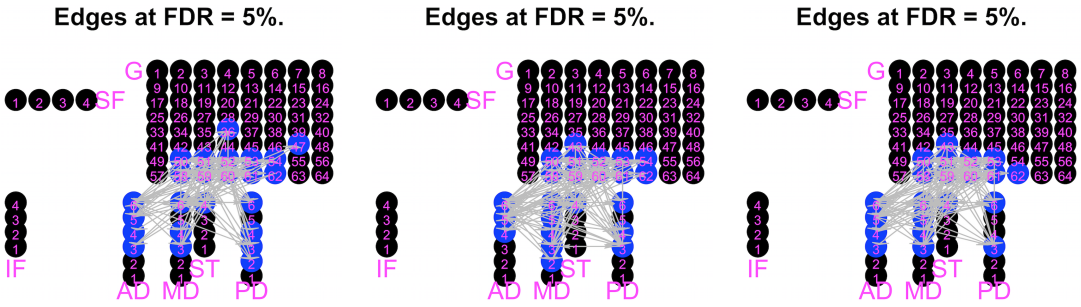
Simply redefine $\boldsymbol{\Delta}_i(t)$ in (7.4) by:

$$\begin{aligned}\boldsymbol{\Delta}_i(t) = \Big(&A_{i1}\cdot\left[\delta(m_i,m_1)\cdot\gamma_{i1}^W+(1-\delta(m_i,m_1))\cdot\gamma_{i1}^B\right]\cdot b_1(t),\ldots,A_{i1}\cdot\left[\delta(m_i,m_1)\cdot\gamma_{i1}^W+(1-\delta(m_i,m_1))\cdot\gamma_{i1}^B\right]\cdot b_L(t),\\ &A_{i2}\cdot\left[\delta(m_i,m_2)\cdot\gamma_{i2}^W+(1-\delta(m_i,m_2))\cdot\gamma_{i2}^B\right]\cdot b_1(t),\ldots,A_{id}\cdot\left[\delta(m_i,m_d)\cdot\gamma_{id}^W+(1-\delta(m_i,m_d))\cdot\gamma_{id}^B\right]\cdot b_L(t)\Big),\end{aligned}$$

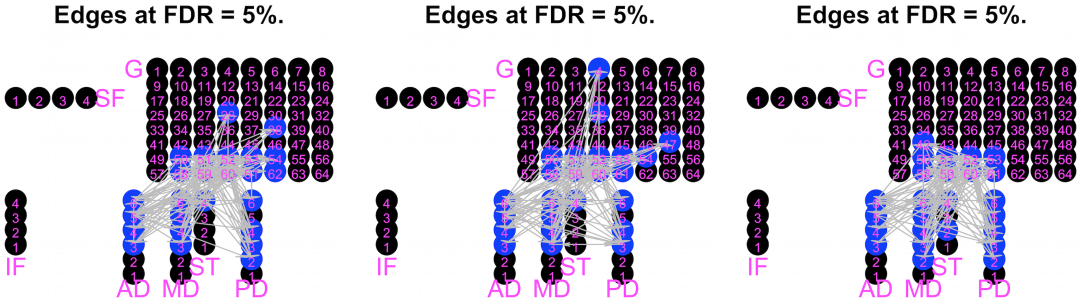
$\boldsymbol{\Omega}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}$, $\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}$, and $\boldsymbol{\Xi}_{\boldsymbol{\Theta}_I,\boldsymbol{\tau}}$ for SWODDM (2.6) will have the same representations as those for MODDM (2.5), and thus (7.3) holds for both MODDM and SWODDM.

7.3 Appendix C: Additional Average Network Results of 20 Consecutive 1-second Segments around Seizure Onset Times.

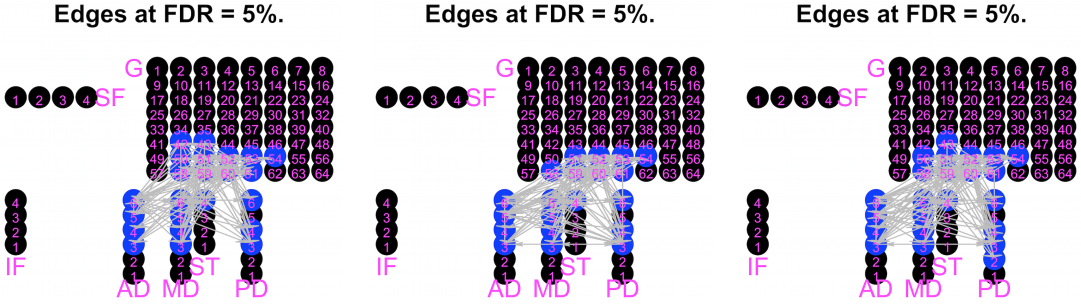
Average Networks around the Seizure Onset from MODDM



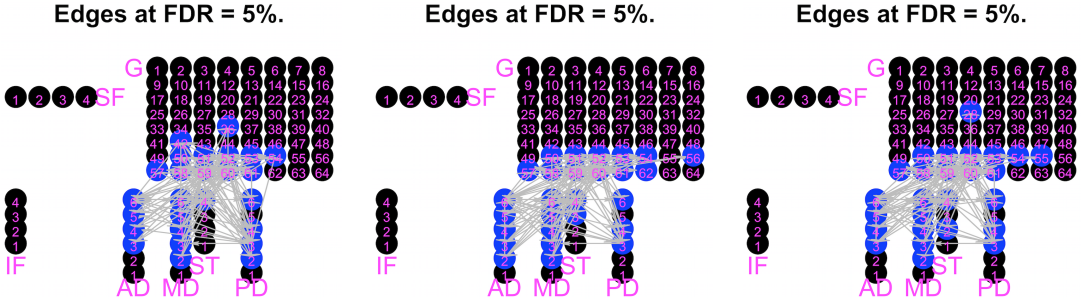
(a) 10 Seconds before (b) 9 Seconds before (c) 8 Seconds before



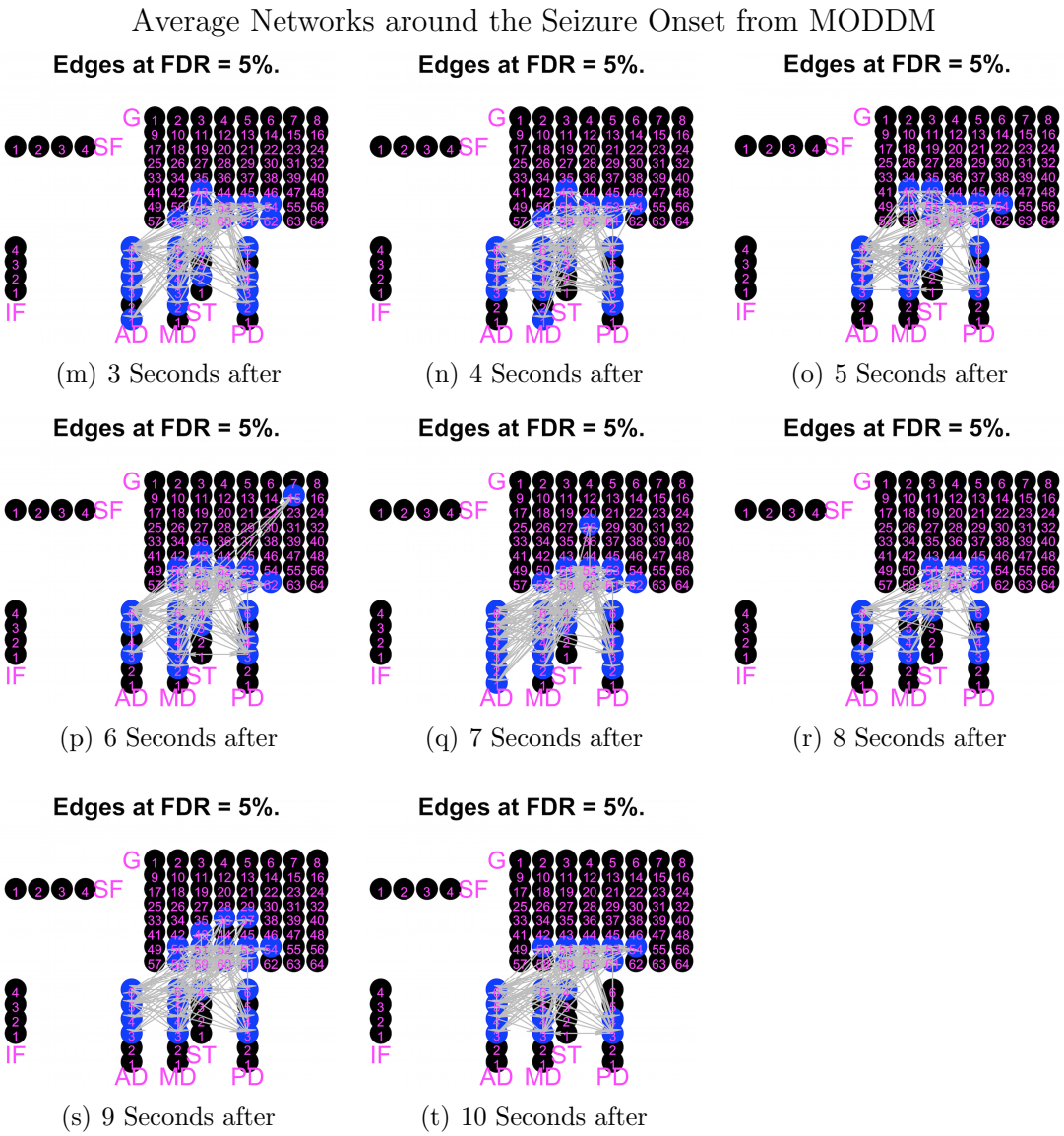
(d) 7 Seconds before (e) 6 Seconds before (f) 5 Seconds before



(g) 4 Seconds before (h) 3 Seconds before (i) 2 Seconds before



(j) 1 Second before (k) 1 Second after (l) 2 Seconds after



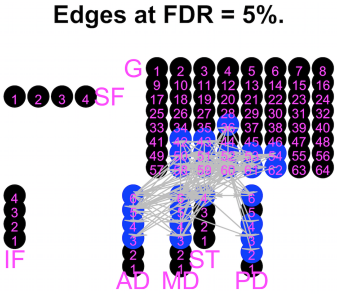
(m) 3 Seconds after (n) 4 Seconds after (o) 5 Seconds after

(p) 6 Seconds after (q) 7 Seconds after (r) 8 Seconds after

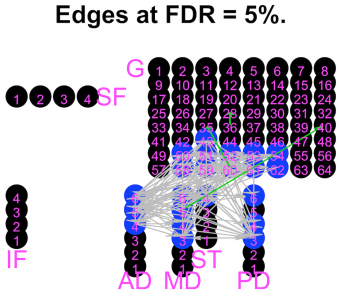
(s) 9 Seconds after (t) 10 Seconds after

Figure 7.13: The average networks across 4 seizures from -10 seconds to 10 seconds centered at the seizure onset time. Nodes in black correspond to the regions that are isolated from the rest regions and form a cluster having only one region. Nodes in blue correspond to the regions in the same cluster. The directional network edges in grey are for the pairs of regions in the same cluster with their network edges selected with the FDR of 5%.

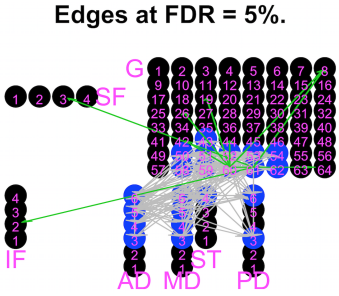
Average Networks around the Seizure Onset from SWODDM



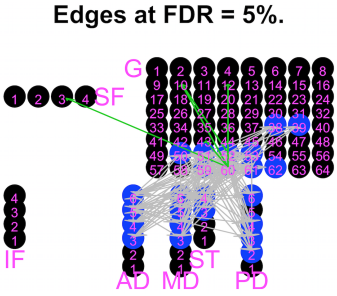
(a) 10 Seconds before



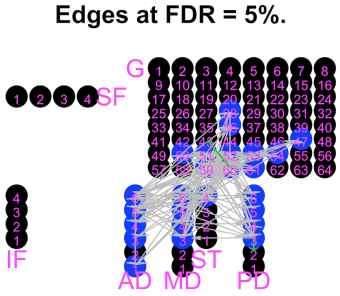
(b) 9 Seconds before



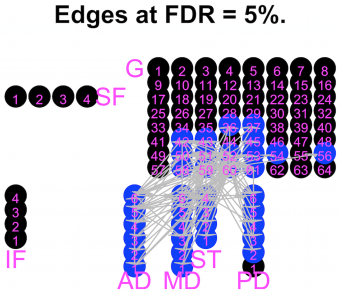
(c) 8 Seconds before



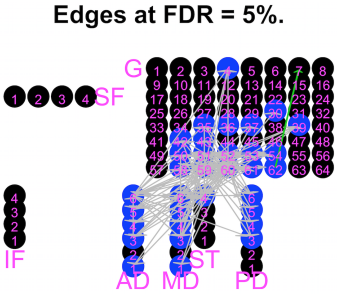
(d) 7 Seconds before



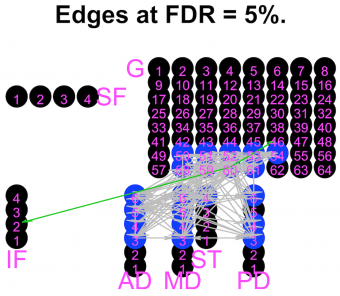
(e) 6 Seconds before



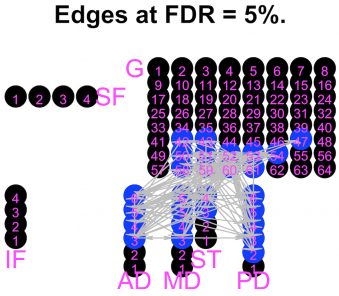
(f) 5 Seconds before



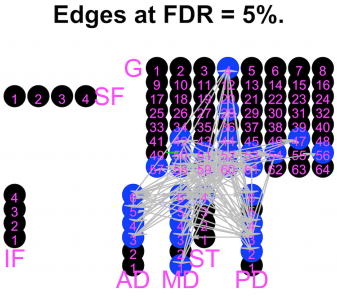
(g) 4 Seconds before



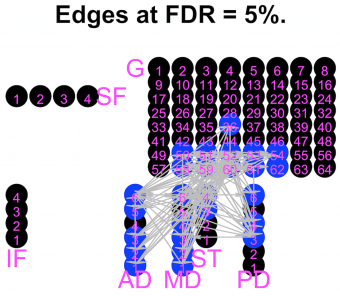
(h) 3 Seconds before



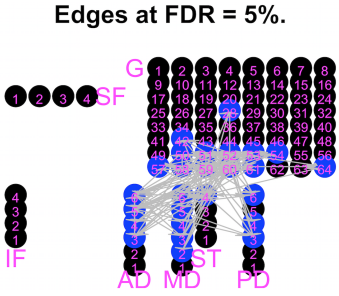
(i) 2 Seconds before



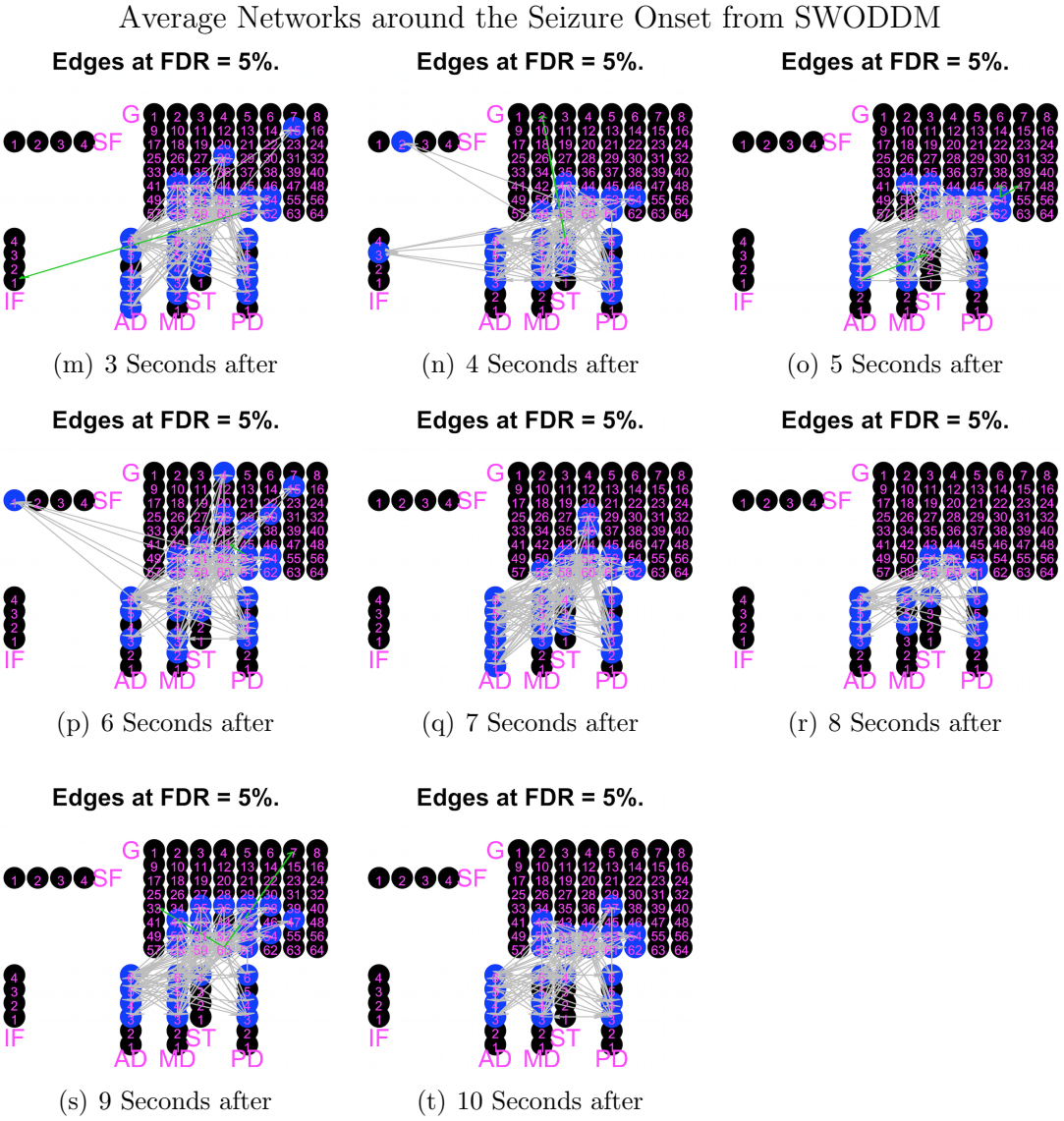
(j) 1 Second before



(k) 1 Second after



(l) 2 Seconds after



(m) 3 Seconds after (n) 4 Seconds after (o) 5 Seconds after

(p) 6 Seconds after (q) 7 Seconds after (r) 8 Seconds after

(s) 9 Seconds after (t) 10 Seconds after

Figure 7.13: The average networks across 4 seizures from -10 seconds to 10 seconds centered at the seizure onset time. Nodes in black correspond to the regions that are isolated from the rest regions and form a cluster having only one region. Nodes in blue correspond to the regions in the same cluster. The directional network edges in grey are for the pairs of regions in the same cluster with their network edges selected with the FDR of 5%. The directional network edges in green are for the pairs of regions between different clusters with their network edges selected with the FDR of 5%.

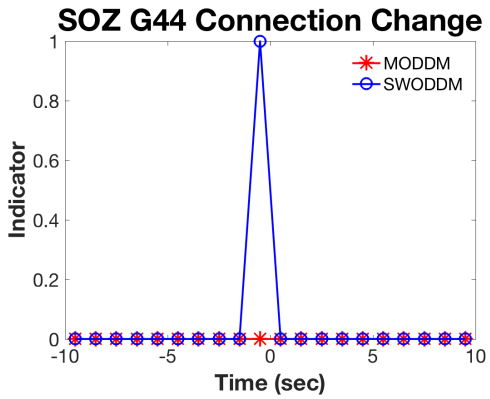


Figure 7.14: SOZ G44 connection status change along with time. 0 indicates G44 is disconnected with the rest of the regions while 1 indicates G44 is connected with the rest of the regions. 0s is the seizure onset time.

Bibliography

Anderson, J. (2005). Learning in sparsely connected and sparsely coded system. *Ersatz Brain Project working note*.

Bard, Y. (1974). Nonlinear parameter estimation.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, **57**(1), 289–300.

Bhaumik, P. and Ghosal, S. (2014). Bayesian estimation in differential equation models. *arXiv preprint arXiv:1403.0609*.

Biegler, L., Damiano, J., and Blau, G. (1986). Nonlinear parameter estimation: a case study comparison. *AIChE Journal*, **32**(1), 29–45.

Brunel, N. J. *et al.* (2008). Parameter estimation of ode's via nonparametric estimators. *Electronic Journal of Statistics*, **2**, 1242–1267.

Büchel, C. and Friston, K. J. (1997). Effective connectivity and neuroimaging. *SPMcourse, short course notes*.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoreti-

cal analysis of structural and functional systems. *Nature Reviews Neuroscience*, **10**(3), 186.

Burle, B., Spieser, L., Roger, C., Casini, L., Hasbroucq, T., and Vidal, F. (2015). Spatial and temporal resolutions of eeg: Is it really black and white? a scalp current density view. *International Journal of Psychophysiology*, **97**(3), 210–220.

Burns, S. P., Santaniello, S., Yaffe, R. B., Jouny, C. C., Crone, N. E., Bergey, G. K., Anderson, W. S., and Sarma, S. V. (2014). Network dynamics of the brain and influence of the epileptic seizure onset zone. *Proceedings of the National Academy of Sciences*, **111**(49), E5321–E5330.

Caffo, B., Peng, R., Dominici, F., Louis, T., and Zeger, S. (2010). Parallel mcmc for analyzing distributed lag models with systematic missing data for an application in environmental epidemiology. *Handbook of Markov Chain Monte Carlo: Methods and Applications*.

Campbell, D. A. (2007). *Bayesian collocation tempering and generalized profiling for estimation of parameters from differential equation models*. Ph.D. thesis, McGill University Canada.

Cao, J., Huang, J. Z., and Wu, H. (2012). Penalized nonlinear least squares estimation of time-varying parameters in ordinary differential equations. *Journal of computational and graphical statistics*, **21**(1), 42–56.

Cervenka, M. C., Franaszczuk, P. J., Crone, N. E., Hong, B., Caffo, B. S., Bhatt, P., Lenz, F. A., and Boatman-Reich, D. (2013). Reliability of early cortical auditory gamma-band responses. *Clinical Neurophysiology*, **124**(1), 70–82.

Chen, J. and Wu, H. (2008). Efficient local estimation for time-varying coefficients

in deterministic dynamic models with applications to hiv-1 dynamics. *Journal of the American Statistical Association*, **103**(481), 369–384.

Daunizeau, J., David, O., and Stephan, K. E. (2011). Dynamic causal modelling: a critical review of the biophysical and statistical foundations. *Neuroimage*, **58**(2), 312–322.

David, O. and Friston, K. J. (2003). A neural mass model for meg/eeg:: coupling and neuronal dynamics. *NeuroImage*, **20**(3), 1743–1755.

David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., and Friston, K. J. (2006). Dynamic causal modeling of evoked responses in eeg and meg. *NeuroImage*, **30**(4), 1255–1272.

Dehaene-Lambertz, G., Dehaene, S., Anton, J.-L., Campagne, A., Ciuciu, P., Dehaene, G. P., Denghien, I., Jobert, A., LeBihan, D., Sigman, M., *et al.* (2006). Functional segregation of cortical language areas by sentence repetition. *Human brain mapping*, **27**(5), 360–371.

Deuflhard, P. and Bornemann, F. (2012). *Scientific computing with ordinary differential equations*, volume 42. Springer Science & Business Media.

Efron, B. and Tibshirani, R. (2002). Empirical bayes methods and false discovery rates for microarrays. *Genetic epidemiology*, **23**(1), 70–86.

Fell, J. and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature reviews neuroscience*, **12**(2), 105.

Fitzpatrick, R. (2013). *Oscillations and waves: an introduction*. CRC Press.

Fogassi, L. and Luppino, G. (2005). Motor functions of the parietal lobe. *Current opinion in neurobiology*, **15**(6), 626–631.

Földiák, P. and Young, M. P. (1995). Sparse coding in the primate cortex. In 895-898, editor, *The Handbook of Brain Theory and Neural Networks*. The MIT Press.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, **9**(10), 474–480.

Friston, K., Frith, C., Liddle, P., and Frackowiak, R. (1993a). Functional connectivity: the principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow & Metabolism*, **13**(1), 5–14.

Friston, K., Frith, C., and Frackowiak, R. (1993b). Time-dependent changes in effective connectivity measured with pet. *Human Brain Mapping*, **1**(1), 69–79.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, **2**(1-2), 56–78.

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain connectivity*, **1**(1), 13–36.

Friston, K. J., Mechelli, A., Turner, R., and Price, C. J. (2000). Nonlinear responses in fmri: the balloon model, volterra kernels, and other hemodynamics. *NeuroImage*, **12**(4), 466–477.

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, **19**(4), 1273–1302.

Gelman, A., Bois, F., and Jiang, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association*, **91**(436), 1400–1412.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Gerstein, G. L. and Perkel, D. H. (1969). Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science*, **164**(3881), 828–830.

Girolami, M. (2008). Bayesian inference for differential equations. *Theoretical Computer Science*, **408**(1), 4–16.

Graner, F. and Glazier, J. A. (1992). Simulation of biological cell sorting using a two-dimensional extended potts model. *Physical review letters*, **69**(13), 2013.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438.

Granger, C. W. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and control*, **2**, 329–352.

Greenblatt, R. E., Pflieger, M., and Ossadtchi, A. (2012). Connectivity measures applied to human brain electrophysiological data. *Journal of neuroscience methods*, **207**(1), 1–16.

Grefkes, C. and Fink, G. R. (2011). Reorganization of cerebral networks after stroke: new insights from neuroimaging with connectivity approaches. *Brain*, **134**(5), 1264–1276.

Härdle, W. (1990). *Applied nonparametric regression*. Number 19. Cambridge university press.

He, B., Yang, L., Wilke, C., and Yuan, H. (2011). Electrophysiological imaging of brain activity and connectivity-challenges and opportunities. *IEEE transactions on biomedical engineering*, **58**(7), 1918–1931.

Hemker, P. (1972). Numerical methods for differential equations in system simulation and in parameter estimation. *Analysis and Simulation of biochemical systems*, **28**, 59–80.

Huang, Y. and Wu, H. (2006). A bayesian approach for estimating antiviral efficacy in hiv dynamic models. *Journal of Applied Statistics*, **33**(2), 155–174.

Huang, Y., Liu, D., and Wu, H. (2006). Hierarchical bayesian methods for estimation of parameters in a longitudinal hiv dynamic system. *Biometrics*, **62**(2), 413–423.

Huettel, S. A., Song, A. W., McCarthy, G., *et al.* (2004). *Functional magnetic resonance imaging*, volume 1. Sinauer Associates Sunderland, MA.

Kiebel, S. J., David, O., and Friston, K. J. (2006). Dynamic causal modelling of evoked responses in eeg/meg with lead field parameterization. *NeuroImage*, **30**(4), 1273–1284.

Kramer, M. A., Kolaczyk, E. D., and Kirsch, H. E. (2008). Emergent network topology at seizure onset in humans. *Epilepsy research*, **79**(2-3), 173–186.

Kramer, M. A., Eden, U. T., Kolaczyk, E. D., Zepeda, R., Eskandar, E. N., and Cash, S. S. (2010). Coalescence and fragmentation of cortical networks during focal seizures. *Journal of Neuroscience*, **30**(30), 10076–10085.

Kramer, M. A., Truccolo, W., Eden, U. T., Lepage, K. Q., Hochberg, L. R., Eskandar, E. N., Madsen, J. R., Lee, J. W., Maheshwari, A., Halgren, E., *et al.* (2012).

Human seizures self-terminate across spatial scales via a critical transition. *Proceedings of the National Academy of Sciences*, **109**(51), 21116–21121.

Li, Z., Osborne, M. R., and Prvan, T. (2005). Parameter estimation of ordinary differential equations. *IMA Journal of Numerical Analysis*, **25**(2), 264–285.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fmri signal. *Nature*, **412**(6843), 150.

Lu, T., Liang, H., Li, H., and Wu, H. (2011). High-dimensional odes coupled with mixed-effects modeling techniques for dynamic gene regulatory network identification. *Journal of the American Statistical Association*, **106**(496), 1242–1258.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science*, **309**(5744), 2228–2232.

Mattheij, R. and Molenaar, J. (1996). *Ordinary differential equations in theory and practice*, volume 43. SIAM.

Micheloyannis, S. (2012). Graph-based network analysis in schizophrenia. *World journal of psychiatry*, **2**(1), 1.

Milgram, S. (1967). The small world problem. *Psychology today*, **2**(1), 60–67.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, **298**(5594), 824–827.

Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M., and Alon, U. (2004). Superfamilies of evolved and designed networks. *Science*, **303**(5663), 1538–1542.

Molnar-Szakacs, I., Iacoboni, M., Koski, L., and Mazziotta, J. C. (2004). Functional segregation within pars opercularis of the inferior frontal gyrus: evidence from fmri studies of imitation and action observation. *Cerebral Cortex*, **15**(7), 986–994.

Mormann, F., Kreuz, T., Rieke, C., Andrzejak, R. G., Kraskov, A., David, P., Elger, C. E., and Lehnertz, K. (2005). On the predictability of epileptic seizures. *Clinical neurophysiology*, **116**(3), 569–587.

Moutoussis, K. and Zeki, S. (1997). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **264**(1387), 1407–1414.

Netoff, T. I. and Schiff, S. J. (2002). Decreased neuronal synchronization during experimental seizures. *Journal of Neuroscience*, **22**(16), 7297–7307.

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the national academy of sciences*, **103**(23), 8577–8582.

Nissen, I. A., van Klink, N. E., Zijlmans, M., Stam, C. J., and Hillebrand, A. (2016). Brain areas with epileptic high frequency oscillations are functionally isolated in meg virtual electrode networks. *Clinical Neurophysiology*, **127**(7), 2581–2591.

Olshausen, B. and Field, D. (2004). Sparse coding of sensor inputs. *Current Opinions in Neurobiology*, **14**, 481–487.

Park, H.-J. and Friston, K. (2013). Structural and functional brain networks: from connections to cognition. *Science*, **342**(6158), 1238411.

Pearl, J. (2000). *Causality: models, reasoning and inference*, volume 29. Springer.

Pereda, E., Quiroga, R. Q., and Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Progress in neurobiology*, **77**(1-2), 1–37.

Potts, R. B. (1952). Some generalized order-disorder transformations. In *Mathematical proceedings of the cambridge philosophical society*, volume 48, pages 106–109. Cambridge University Press.

Poyton, A., Varziri, M. S., McAuley, K. B., McLellan, P. J., and Ramsay, J. O. (2006). Parameter estimation in continuous-time dynamic models using principal differential analysis. *Computers & chemical engineering*, **30**(4), 698–708.

Protzner, A. B. and McIntosh, A. R. (2006). Testing effective connectivity changes with structural equation modeling: what does a bad model tell us? *Human brain mapping*, **27**(12), 935–947.

Qi, X., Zhao, H., *et al.* (2010). Asymptotic efficiency and finite-sample properties of the generalized profiling estimation of parameters in ordinary differential equations. *The Annals of Statistics*, **38**(1), 435–481.

Rajapakse, J. C. and Zhou, J. (2007). Learning effective brain connectivity with dynamic bayesian networks. *Neuroimage*, **37**(3), 749–760.

Ramsay, J. O. (2004). Functional data analysis. *Encyclopedia of Statistical Sciences*, **4**.

Ramsay, J. O., Hooker, G., Campbell, D., and Cao, J. (2007). Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **69**(5), 741–796.

Rizzolatti, G., Fogassi, L., and Gallese, V. (1997). Parietal cortex: from sight to action. *Current opinion in neurobiology*, **7**(4), 562–567.

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality and fmri. *Neuroimage*, **25**(1), 230–242.

Roebroeck, A., Formisano, E., and Goebel, R. (2011). The identification of interacting networks in the brain using fmri: model selection, causality and deconvolution. *Neuroimage*, **58**(2), 296–302.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of computational and graphical statistics*, **11**(4), 735–757.

Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with eeg/meg. *Computers in biology and medicine*, **41**(12), 1110–1117.

Salmelin, R., Hámáaláinen, M., Kajola, M., and Hari, R. (1995). Functional segregation of movement-related rhythmic activity in the human brain. *Neuroimage*, **2**(4), 237–243.

Schiff, S. J., Sauer, T., Kumar, R., and Weinstein, S. L. (2005). Neuronal spatiotemporal pattern discrimination: the dynamical evolution of seizures. *Neuroimage*, **28**(4), 1043–1055.

Schindler, K., Leung, H., Elger, C. E., and Lehnertz, K. (2006). Assessing seizure dynamics by analysing the correlation structure of multichannel intracranial eeg. *Brain*, **130**(1), 65–77.

Schindler, K., Amor, F., Gast, H., Müller, M., Stibal, A., Mariani, L., and Rum-

mel, C. (2010). Peri-ictal correlation dynamics of high-frequency (80–200 hz) intracranial eeg. *Epilepsy research*, **89**(1), 72–81.

Schindler, K. A., Bialonski, S., Horstmann, M.-T., Elger, C. E., and Lehnertz, K. (2008). Evolving functional network properties and synchronizability during human epileptic seizures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **18**(3), 033119.

Schnitzler, A. and Gross, J. (2005). Normal and pathological oscillatory communication in the brain. *Nature reviews neuroscience*, **6**(4), 285.

Schuster, P. (1983). Biophysics, ed. by walter hoppe ltd.

Serway, R. A. and Jewett, J. W. (2003). Oscillations and mechanical waves. In *Physics for Scientists and Engineers*, pages 433–543. Brooks/Cole.

Sporns, O. (2011). Networks of the brain mit press.

Sporns, O., Chialvo, D. R., Kaiser, M., and Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in cognitive sciences*, **8**(9), 418–425.

Squire, L. R. and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, **253**(5026), 1380–1386.

Suchard, M. A., Wang, Q., Chan, C., Frelinger, J., Cron, A., and West, M. (2010). Understanding gpu programming for statistical computation: Studies in massively parallel massive mixtures. *Journal of computational and graphical statistics*, **19**(2), 419–438.

Tononi, G., Sporns, O., and Edelman, G. M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, **91**(11), 5033–5037.

Vaina, L. M. (1994). Functional segregation of color and motion processing in the human visual cortex: clinical evidence. *Cerebral Cortex*, **4**(5), 555–572.

Valdes-Sosa, P. A., Roebroeck, A., Daunizeau, J., and Friston, K. (2011). Effective connectivity: influence, causality and biophysical modeling. *Neuroimage*, **58**(2), 339–361.

Van Dyk, D. A. and Park, T. (2008). Partially collapsed gibbs samplers: Theory and methods. *Journal of the American Statistical Association*, **103**(482), 790–796.

Varah, J. M. (1982). A spline least squares method for numerical parameter estimation in differential equations. *SIAM Journal on Scientific and Statistical Computing*, **3**(1), 28–46.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *nature*, **393**(6684), 440.

Welvaert, M. and Rosseel, Y. (2013). On the definition of signal-to-noise ratio and contrast-to-noise ratio for fmri data. *PloS one*, **8**(11), e77089.

Wendling, F., Bellanger, J.-J., Badier, J.-M., and Coatrieux, J.-L. (1996). Extraction of spatio-temporal signatures from depth eeg seizure signals based on objective matching in warped vectorial observations. *IEEE Transactions on biomedical engineering*, **43**(10), 990–1000.

Wilson, J. R. and Corlett, N. (2005). *Evaluation of human work*. CRC press.

Wu, H., Lu, T., Xue, H., and Liang, H. (2014a). Sparse additive ordinary differential equations for dynamic gene regulatory network modeling. *Journal of the American Statistical Association*, **109**(506), 700–716.

Wu, L. and Gotman, J. (1998). Segmentation and classification of eeg during epileptic seizures. *Clinical Neurophysiology*, **106**(4), 344–356.

Wu, S., Xue, H., Wu, Y., and Wu, H. (2014b). Variable selection for sparse high-dimensional nonlinear regression models by combining nonnegative garrote and sure independence screening. *Statistica Sinica*, **24**(3), 1365.

Wu, T., Wang, L., Hallett, M., Chen, Y., Li, K., and Chan, P. (2011). Effective connectivity of brain networks during self-initiated movement in parkinson's disease. *Neuroimage*, **55**(1), 204–215.

Xue, H., Miao, H., and Wu, H. (2010). Sieve estimation of constant and time-varying coefficients in nonlinear ordinary differential equation models by considering both numerical error and measurement error. *Annals of statistics*, **38**(4), 2351.

Yu, C., Zhou, Y., Liu, Y., Jiang, T., Dong, H., Zhang, Y., and Walter, M. (2011). Functional segregation of the human cingulate cortex is confirmed by functional connectivity based neuroanatomical parcellation. *Neuroimage*, **54**(4), 2571–2581.

Yu, S., Huang, D., Singer, W., and Nikolić, D. (2008). A small world of neuronal synchrony. *Cerebral cortex*, **18**(12), 2891–2901.

Zhang, T., Wu, J., Li, F., Caffo, B., and Boatman-Reich, D. (2015). A dynamic directional model for effective brain connectivity using electrocorticographic (ecog) time series. *Journal of the American Statistical Association*, **110**(509), 93–106.

Zhang, T., Yin, Q., Caffo, B., Sun, Y., Boatman-Reich, D., *et al.* (2017). Bayesian inference of high-dimensional, cluster-structured ordinary differential equation models with applications to brain connectivity studies. *The Annals of Applied Statistics*, **11**(2), 868–897.

Zhang, T., Sun, Y., Li, H., Yan, G., Tanabe, S., Miao, R., Wang, Y., Caffo, B. S., and Quigg, M. S. (2020). Bayesian inference of a directional brain network model for intracranial eeg data. *Computational Statistics & Data Analysis*, **144**, 106847.