Computational Statistics and Data Analysis 144 (2020) 106847

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Bayesian inference of a directional brain network model for intracranial EEG data

Tingting Zhang [a,*], Yinge Sun [a], Huazhang Li [a], Guofen Yan [b], Seiji Tanabe [c], Ruizhong Miao [a], Yaotian Wang [a], Brian S. Caffo [d], Mark S. Quigg [c]

[a] Department of Statistics, University of Virginia, United States of America
[b] Division of Biostatistics, University of Virginia, United States of America
[c] Department of Neurology, University of Virginia, United States of America
[d] Department of Biostatistics, Johns Hopkins University, United States of America

ARTICLE INFO

Article history:
Received 20 December 2018
Received in revised form 3 August 2019
Accepted 26 September 2019
Available online 15 November 2019

Keywords:
Ordinary differential equation
Bayesian methods
Dynamic system
Time series
Brain networks

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 another 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. The authors 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 meaningful network results, a cluster structure is assumed for the ODE model parameters that quantify directional connectivity among regions. The cluster structure is in line with the functional specialization of the human brain; the brain areas specialized in the same function tend to be in the same cluster. Two Bayesian methods are developed to estimate the model parameters of the proposed ODE model and to identify clusters of strongly connected brain regions. The proposed ODE model and Bayesian method are applied to iEEG data collected from a patient with medically intractable epilepsy and used to examine the patient's brain networks before the seizure onset.

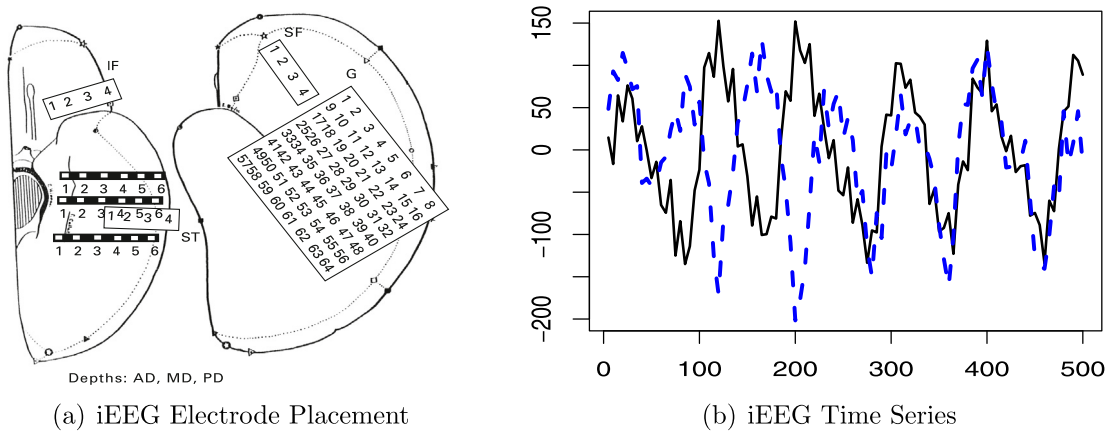
1. Introduction

The human brain is a network system, where brain regions, as network nodes, constantly interact with each other. The directional effect exerted by one brain region over another is referred to as directional connectivity and corresponds to a network edge in the brain network. Identifying connected brain regions and mapping the human brain network help us understand the mechanism of the brain as well as its normal and abnormal functions. In this article, we model the directional connectivity of the human brain and identify connected brain regions using intracranial electrocorticography (iEEG) data, multivariate time series measurements of many regions' neuronal activities.

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. Fig. 1(a) shows the placement of iEEG electrodes on the exposed brain of an epileptic

* Correspondence to: Halsey Hall 111, Charlottesville, VA 22904, United States of America.
E-mail address: tz3b@virginia.edu (T. Zhang).

https://doi.org/10.1016/j.csda.2019.106847
0167-9473/



(a) iEEG Electrode Placement (b) iEEG Time Series

Fig. 1. (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.

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. Fig. 1(b) illustrates two iEEG time series recorded in 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 a millisecond scale. As such, iEEG data have high spatial and temporal resolutions. Second, iEEG produces highly reliable and reproducible measurements of brain activity with a strong signal-to-noise ratio (SNR, Cervenka et al., 2013). As such, iEEG data are ideal for examining the brain network.

The human brain is also a continuous-time dynamic system, in which each brain region's activity depends on other regions' activities. It is biophysically natural to use ordinary differential equations (ODEs) to describe the dynamic mechanism of the brain and use model parameters to quantify directional connectivity among brain regions. The most popular ODE model for the brain's directional connectivity is the dynamic causal modeling (DCM, Daunizeau et al., 2011; David and Friston, 2003; David et al., 2006; Kiebel et al., 2006; Friston et al., 2003), which characterizes directional connectivity among only a few large regions (usually no more than 5) based on functional magnetic resonance imaging (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.

We develop a new high-dimensional ODE model for iEEG data to characterize directional connectivity among many regions (more than 50 at least). The new model has two advantages. First, the new model, as an extension of a biophysical model called damped harmonic oscillator (DHO, Serway and Jewett, 2003), characterizes the brain regions' oscillatory activity—periodic or rhythmic up-and-down temporal behavior of the neuronal activity. The oscillation is an important feature of any time series data measuring the neural activity of local areas of the brain (Fell and Axmacher, 2011; Fries, 2005; Schnitzler and Gross, 2005). Fig. 1(b) shows two brain regions' oscillatory activity recorded by iEEG. Second, to address the difficulty in specifying the complex interactive relationship among many regions, we use a linear approximation to model the complex mechanism of the high-dimensional brain system, an idea similar to the linear regression. As such, the new model combines the strengths of statistical modeling and scientific modeling: The model captures the major oscillatory feature of the brain through extending a physical oscillatory system to the high-dimension brain system and uses a statistical model formulation to provide the model flexibility. We refer to the new model as the oscillatory dynamic directional model (ODDM) (details of the model construction are provided in Section 2).

As a high-dimensional ODE model, the ODDM contains many free parameters for quantifying directional connectivity among many regions recorded by iEEG. We use a prior knowledge of brain networks to increase estimation efficiency of ODDM parameters. Specifically, we impose a cluster structure, also called the modularity, on the ODDM parameters. The cluster structure, consisting of several functionally independent subnetworks of lower dimensions, provides an intuitive interpretation of functional specialization of brain regions in different modules/clusters. Moreover, the cluster structure 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, the proposed model leads to scientifically meaningful network results. We refer to the ODDM with the cluster structure as the modular ODDM (MODDM).

The MODDM, like many other statistical models, is an approximation of the complex system under study, and thus, has a discrepancy from the underlying true mechanism of the brain. As the model uncertainty quantification is natural within the Bayesian framework, we develop Bayesian methods to estimate the model parameters of the MODDM while accounting for the discrepancy between the MODDM and the underlying true brain system. The quantification of the ODE model discrepancy has rarely been addressed in the literature. As such, this paper not only addresses a pressing need for statistical modeling of the biophysical mechanism of the brain but also introduces a new approach to inferring high-dimensional ODE models with many free parameters.

The rest of the manuscript is organized as follows. Section 2 introduces the MODDM. In Section 3, we present two Bayesian hierarchical methods to estimate the MODDM based on basis representation of brain regions' state functions and a Markov chain Monte Carlo (MCMC) algorithm to make posterior inference. We present simulation studies of the proposed Bayesian method in comparison with existing network methods in Section 4, and apply the MODDM to analyze a real iEEG study in Section 5. Section 6 discusses analysis results and future research directions.

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 $\mathbf{x}(t)$:

$$\mathbf{y}(t) = \mathbf{x}(t) + \boldsymbol{\epsilon}(t), \tag{1}$$

where $\boldsymbol{\epsilon}(t) = (\epsilon_1(t), \ldots, \epsilon_d(t))'$ is a d-dimensional vector of measurement errors with mean zeros. 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

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

where $z(t)$ is the state or the spatial location of a one-dimensional system (called oscillator) at time t, 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. Fig. 2(a) 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.

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 describe dynamic systems with oscillatory mechanisms. However, the DHO in these applications is for describing specific systems, but difficult to apply to other systems.

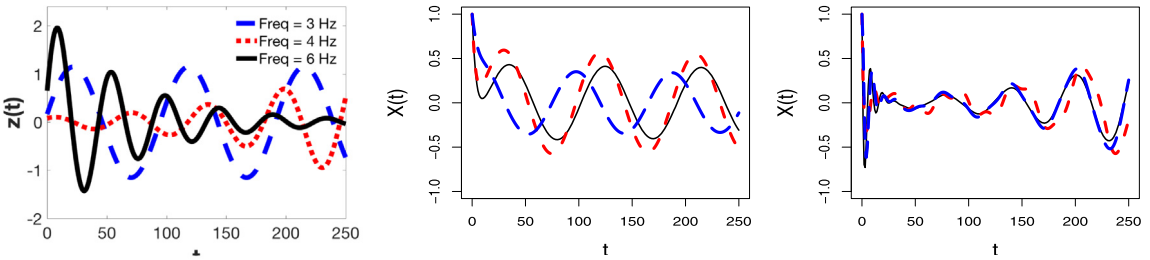
We extend the one-dimensional DHO to the high-dimensional brain system consisting of many interactive brain regions. Specifically, we model d brain regions under study as a set of interactive oscillators – each corresponding to one region – influenced by the effect exerted by others. For region i,

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

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

The functions $F_i(\mathbf{x}(t))$, $i = 1, \ldots, d$, represent the directional connectivity among d regions and are difficult to specify due to the limited understanding of the brain's biophysical mechanism. 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:

$$\frac{d^2x_i(t)}{dt^2} = D_i + \sum_{j\neq i}^{d} A_{ij} \cdot x_j(t) + A_{ii} \cdot x_i(t) + G_i\, \frac{dx_i(t)}{dt}, \tag{3}$$



(a) Simulated time series of DHO (b) $\mathbf{x}(t)$ in MODDM first cluster (c) $\mathbf{x}(t)$ in MODDM second cluster

Fig. 2. Simulated time series.

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 (3) brings three benefits in practice. First, the second-order ODEs have a physical foundation and are suitable for characterizing the brain's oscillatory activity. Second, this simple form provides a flexible way to model directional 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 (3) together with the observation model (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 directional 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., 2010, 2008, 2007; 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. Modular 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 the cluster structure, which consists of several functionally independent clusters. Connections among regions in the same cluster are dense. The cluster structure has been widely reported in the literature on brain networks (Milo et al., 2002, 2004; Newman, 2006; Sporns, 2011).

To characterize the cluster structure, 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 indicators γ_{ij}s, which take values 1 or 0, to distinguish significant directional connectivity from zero one. We propose the following ODE model, as an extension of the ODDM (3), for the sparse brain network in the cluster structure:

$$\frac{d^2x_i(t)}{dt^2} = \sum_{j=1}^{d} \delta(m_i, m_j) \cdot \gamma_{ij} \cdot A_{ij} \cdot x_j(t) + D_i + G_i \frac{dx_i(t)}{dt}, \tag{4}$$

where the delta function $\delta(m_i, m_j)$ equals 1 if $m_i = m_j$ and 0 otherwise. Model (4) implies that a directional effect from region j onto region i is nonzero, if and only if the two regions fall into the same module and the associated indicator γ_{ij} is nonzero.

Model (4) with the observation model (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. Fig. 2(b) shows $x(t)$s of three regions in one cluster and Fig. 2(c) shows $x(t)$s in the other cluster. The MODDM produces oscillatory state functions.

3. Bayesian estimation methods for MODDM

Under the MODDM, the focus is on identifying clusters and selecting network edges corresponding to nonzero γ_{ij}s. We develop two Bayesian methods to estimate these parameters.

Two types of approaches are proposed for estimating ODE models in the literature: discretization methods which numerically fit $\boldsymbol{x}(t)$ 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 et al., 2006; Huang and Wu, 2006; Li et al., 2005; Matteij and Molenaar, 2002; Xue et al., 2010) and basis-function-expansion approaches which represent $\boldsymbol{x}(t)$ with functional bases (Bhaumik and Ghosal, 2014; Brunel, 2008; Deuflhard and Bornemann, 2000; Poyton et al., 2006; Qi and Zhao, 2010; Ramsay and Silverman, 2005; Ramsay et al., 2007; Varah, 1982). 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 $\boldsymbol{x}(t)$ 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 $\{t_1 = 1, t_2, \ldots, t_q = T\}$ (where $L = q + 5 - 2$) of the interval $[1, T]$:

$$x_i(t) = \tilde{\eta}_i' \, \boldsymbol{b}(t), \tag{5}$$

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

As in Ramsay (2006), 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 (1) independent and identically distributed with a normal distribution with mean zero and unknown variance σ^2. As such,

$$y_i(t) \sim N(x_i(t), \sigma_i^2). \tag{6}$$

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 $\boldsymbol{x}(t)$.

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

$$p(\boldsymbol{\eta}|\Theta_I, \boldsymbol{\tau}) \propto \exp\left\{-\sum_{i=1}^{d} \frac{\mathrm{R}_i(\boldsymbol{\eta}, \Theta_I)}{2\tau_i}\right\}, \tag{7}$$

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

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

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 $d^2 x_i(t)/dt^2 = \tilde{\eta}_i' \, \boldsymbol{b}^{(2)}(t)$.

The distribution (7) 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. The probability (7) provides a generating model for the basis coefficients $\boldsymbol{\eta}$ and the state functions $\mathbf{x}(t)$ based on the given MODDM. The probability allows for the deviation of the state functions from the MODDM. Moreover, the formulation of the distribution (7), equivalent to a normal distribution for $\boldsymbol{\eta}$, as explained below, leads to normal posterior conditional distributions of model parameters $\boldsymbol{\theta} = \{\mathbf{A}, \mathbf{D}, \mathbf{G}\}$, which are easy to simulate.

With the linear basis representation for $\mathbf{x}(t)$ in (5), $\sum_{i=1}^{d} \mathrm{R}_i(\boldsymbol{\eta}, \Theta_I)/(2\tau_i)$ given Θ_I and $\boldsymbol{\tau}$ is quadratic of $\boldsymbol{\eta}$:

$$\sum_{i=1}^{d} \mathrm{R}_i(\boldsymbol{\eta}, \Theta_I)/(2\tau_i) = \boldsymbol{\eta}' \, \boldsymbol{\Omega}_{\Theta_I, \boldsymbol{\tau}} \, \boldsymbol{\eta} - 2\boldsymbol{\Lambda}'_{\Theta_I, \boldsymbol{\tau}} \, \boldsymbol{\eta} + \boldsymbol{\Xi}_{\Theta_I, \boldsymbol{\tau}}, \tag{8}$$

where $\Omega_{\Theta_l,\tau}$, $\Lambda_{\Theta_l,\tau}$, and $\Xi_{\Theta_l,\tau}$ are a $dL\times dL$ matrix, a $dL\times 1$ vector, and a scalar, respectively, and their formulas depending on Θ_l and τ are provided in Appendix. Thus, the probability model (7) is a multivariate normal distribution

$$\eta|\Theta_l, \tau \sim \text{MN}(\Omega^{-1}_{\Theta_l,\tau}\ \Lambda_{\Theta_l,\tau}, \Omega^{-1}_{\Theta_l,\tau}).$$

This means that the state functions follow a Gaussian process centered at the MODDM with variability controlled by τ.

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

$$p(\Theta_l|\tau) \propto \det(\Omega_{\Theta_l,\tau})^{-1/2} \cdot \exp\left\{\frac{1}{2}(\Lambda'_{\Theta_l,\tau}\ \Omega^{-1}_{\Theta_l,\tau}\Lambda_{\Theta_l,\tau} - \Xi_{\Theta_l,\tau})\right\} \cdot \exp\left\{-\mu\sum_{i,j=1}^{d}\delta(m_i, m_j)\right\} \tag{9}$$

$$\cdot p_0^{\sum_{i,j}\gamma_{ij}} \cdot (1-p_0)^{d^2-\sum_{i,j}\gamma_{ij}} \cdot \prod_{i,j=1}^{d}\phi\left(\frac{A_{ij}}{\xi_0}\right) \cdot \prod_{i=1}^{d}\phi\left(\frac{G_i}{\xi_0}\right) \cdot \prod_{i=1}^{d}\phi\left(\frac{D_i}{\xi_0}\right)$$

where $\phi(\cdot)$ is the standard normal density, ξ_0 is a large constant to give an almost flat prior for **A**, **G**, and **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 cluster structure and let $p_0 = 0.9$ to impose the prior belief that within-module 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 facilitate the cluster identification, while smaller p_0 leads to lower selection accuracy.

Prior for variances of model fitting errors. A commonly used non-informative prior for variance parameters τ (Gelman et al., 2014) is $p(\tau) \propto \prod_{i=1}^{d} 1/\tau_i$. However, this prior leads to an improper posterior. Thus, we propose the following prior for τ, which is close to the non-informative prior yet leads to a proper posterior:

$$p(\tau) \propto \prod_{i=1}^{d}(\frac{1}{\tau_i})^{\frac{3}{2}}. \tag{10}$$

Joint posterior distribution. Let $\mathbf{Y} = \{\mathbf{y}(t), t = 1, \ldots, T\}$ and $\sigma = \{\sigma_1^2, \ldots, \sigma_d^2\}$. The model (6) together with priors in (7), (9) and (10) defines a hierarchical Bayesian model for the MODDM. The joint posterior distribution is

$$p(\Theta_l, \tau, \eta, \sigma|\mathbf{Y}) \propto p(\mathbf{Y}|\eta, \sigma) \cdot p(\sigma) \cdot p(\eta|\Theta_l, \tau) \cdot p(\Theta_l|\tau) \cdot p(\tau), \tag{11}$$

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

3.1. The second Bayesian method

Standard approaches simulate from the posterior distribution $p(\Theta_l, \tau, \eta, \sigma|\mathbf{Y})$ 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_l only, while η contributes to most parameters in the Bayesian model (11). 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, 2006) are similar to those from the Bayesian model (11). Considering this, we propose to first fit $\mathbf{x}(t)$ based on the observed data $\mathbf{Y}$, and use a Bayesian method to estimate MODDM parameters based on the fitted $\mathbf{x}(t)$, i.e., the estimated η. We elaborate the details in the following.

We first estimate η by minimizing

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

where $x_i(t)$ is given by (5), $x_i^{(2)}(t) = \tilde{\eta}_i'\ \boldsymbol{b}^{(2)}(t)$, and the smoothing penalty parameter λ is chosen by the generalized cross-validation (Härdle, 1990).

Next, we treat estimated η as the observed data with the likelihood (7). As such, (7), (9) and (10) define the second hierarchical Bayesian model for the MODDM, and the ensuing joint posterior distribution of Θ_l is

$$p(\Theta_l, \tau|\eta) \propto p(\eta|\Theta_l, \tau) \cdot p(\Theta_l|\tau) \cdot p(\tau). \tag{12}$$

We refer to the Bayesian model with the posterior (11) as the full Bayesian method, and the model with the posterior (12) as the Bayesian smoothing method hereafter.

3.2. Posterior inference

We use a partially collapsed Gibbs Sampler (PCGS; van Dyk and Park, 2008) to sample from the posterior distributions $p(\boldsymbol{\Theta}_l, \boldsymbol{\tau}, \boldsymbol{\eta}, \boldsymbol{\sigma}|\mathbf{Y})$ and $p(\boldsymbol{\Theta}_l, \boldsymbol{\tau}|\boldsymbol{\eta})$. 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 Appendix.

Let $\varsigma^{(s)}$ be the sth MCMC simulation of the MODDM 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., 2014) to examine the convergence of the MCMC outputs.

Two posterior quantities of particular interest are the posterior clustering probability and the 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 $\hat{P}^m_{ij} = \frac{1}{S}\sum_{s=1}^{S} \delta(m_i^{(s)}, m_j^{(s)})$, $i, j = 1, \ldots, d$. The latter is the posterior probability that the directional connection from region i to region j is nonzero and is estimated by $\hat{P}^\gamma_{ij} = \frac{1}{S}\sum_{s=1}^{S} \delta(m_i^{(s)}, m_j^{(s)}) \cdot \gamma_{ij}^{(s)}$, $i, j = 1, \ldots, d$.

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

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 $\hat{P}^m_{ij}$s and $\hat{P}^\gamma_{ij}$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 $\{\hat{P}^{m_0}_{ij}, i, j = 1, \ldots, d\}$ and $\{\hat{P}^{\gamma_0}_{ij}, i, j = 1, \ldots, d\}$ based on $\mathbf{Y}^0$ give the empirical null distributions of $\hat{P}^m_{ij}$ and $\hat{P}^\gamma_{ij}$, respectively. Based on these null distributions, we evaluate the significance levels of $\hat{P}^m_{ij}$ and $\hat{P}^\gamma_{ij}$ for every pair of regions i and j and determine the thresholds corresponding to 5% FDR using the method by Benjamini and Hochberg (1995) and Efron and Tibshirani (2002).

3.3. 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., 2011; Suchard et al., 2010). It generally takes 15 min for 5000 iterations of the proposed MCMC algorithm to finish analyzing time series of 50 regions.

4. Simulation study

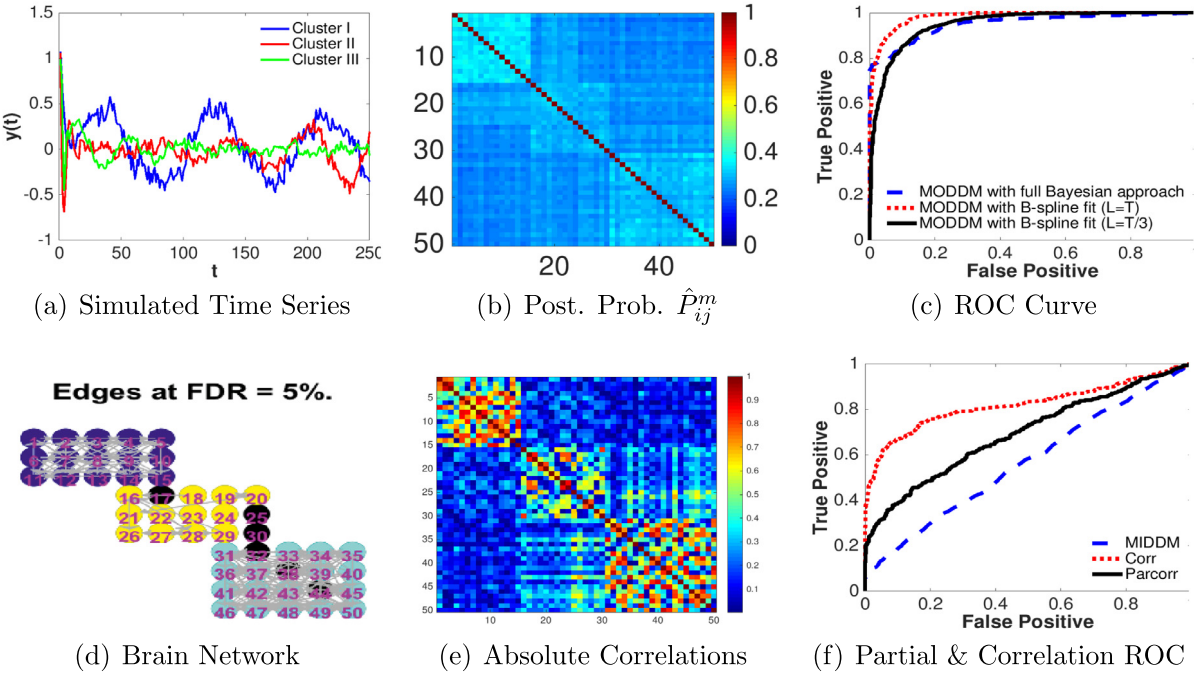
We consider a 50-dimensional dynamic system ($d = 50$) that has 3 clusters of size 15, 15, and 20. We first generated $x_i(t)$, $i = 1, \ldots, 50$, from the MODDM (4). For simplicity, we let the components within the same cluster be all pairwise connected and let $D_i = 0$ and $x_i(0) = 1$ for all i. 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 equaling 0.5 and median pairwise spatial correlations between regions equaling 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 $\text{var}(x_i(t))/\text{var}(\epsilon_i(t))$ – 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 $\epsilon_i(t)$.

We independently simulated 100 multivariate time series from the same model. For each simulated data set, we first standardized every time series $y_i(t)$, $i = 1, \ldots, 50$, to mean zero and norm 1 and applied the developed Bayesian approach to the data. We calculated the true positive rates (TPR) and false positive rates (FPR) of the proposed method using different thresholds for $\hat{P}^m_{ij} \cdot \hat{P}^\gamma_{ij}$. TPR is the percentage of the network edges with $\hat{P}^m_{ij} \cdot \hat{P}^\gamma_{ij}$ above the threshold among all true network edges; FPR is the percentage of the network edges with $\hat{P}^m_{ij} \cdot \hat{P}^\gamma_{ij}$ above the threshold among all void network edges. Table 1 presents the summaries of areas under the ROC curve (AUC) for 100 simulations. The result indicates that the proposed method is robust to violation of model assumptions and can consistently detect connected regions with high accuracy.

We present analysis results of one simulated $\mathbf{Y}$ with an AUC 0.9508 (close to but slightly lower than the median 0.9531) in Fig. 3. Fig. 3(a) shows three representative time series $y_i(t)$, each from one unique cluster. Fig. 3(b) shows the posterior clustering probabilities $\hat{P}^m_{ij}$ estimated by the Bayesian smoothing method. From the figure, we see that $\hat{P}^m_{ij}$s for

Table 1
Summaries of AUC for 100 replicates.

Min	Q1	Median	Q3	Max	Standard deviation
0.9244	0.9453	0.9531	0.9595	0.9658	0.01



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

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

Fig. 3. 3(a) Simulated three time series from three different clusters. 3(b) Posterior clustering probabilities $\hat{P}_{ij}^{m}$ for $i, j = 1, \ldots, 50$ of the Bayesian smoothing method. 3(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 $L = [T/3]$. 3(d) Network edges with 5% FDR for $\hat{P}_{ij}^{m}$ and $\hat{P}_{ij}^{\gamma}$. Nodes in the same color of either dark blue, light blue, and yellow correspond to components in the same cluster identified by the Bayesian smoothing method. Nodes in black correspond to regions identified to be disconnected from other regions. 3(e) The matrix of the absolute correlations of the simulated data. 3(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. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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. Fig. 3(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. Fig. 3(d) shows the directional network edges selected using 5% FDR for $\hat{P}_{ij}^{m}$s and $\hat{P}_{ij}^{\gamma}$s estimated by the Bayesian smoothing method. The nodes in the same color of either dark blue, light blue, and yellow correspond to components identified to be in the same cluster. The nodes in black correspond to components identified to be disconnected from other regions. 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 Fig. 3(f), the MIDDM gave mediocre results with an AUC of 0.57, indicating that the first-order ODEs are unable to capture connectivity among regions with oscillatory activities. Furthermore, we compared with network methods based on correlations and partial correlations. Fig. 3(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. Fig. 3(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 0.69). 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.

5. 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 Fig. 1(a). 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 1–70 Hz. 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 electrodes 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

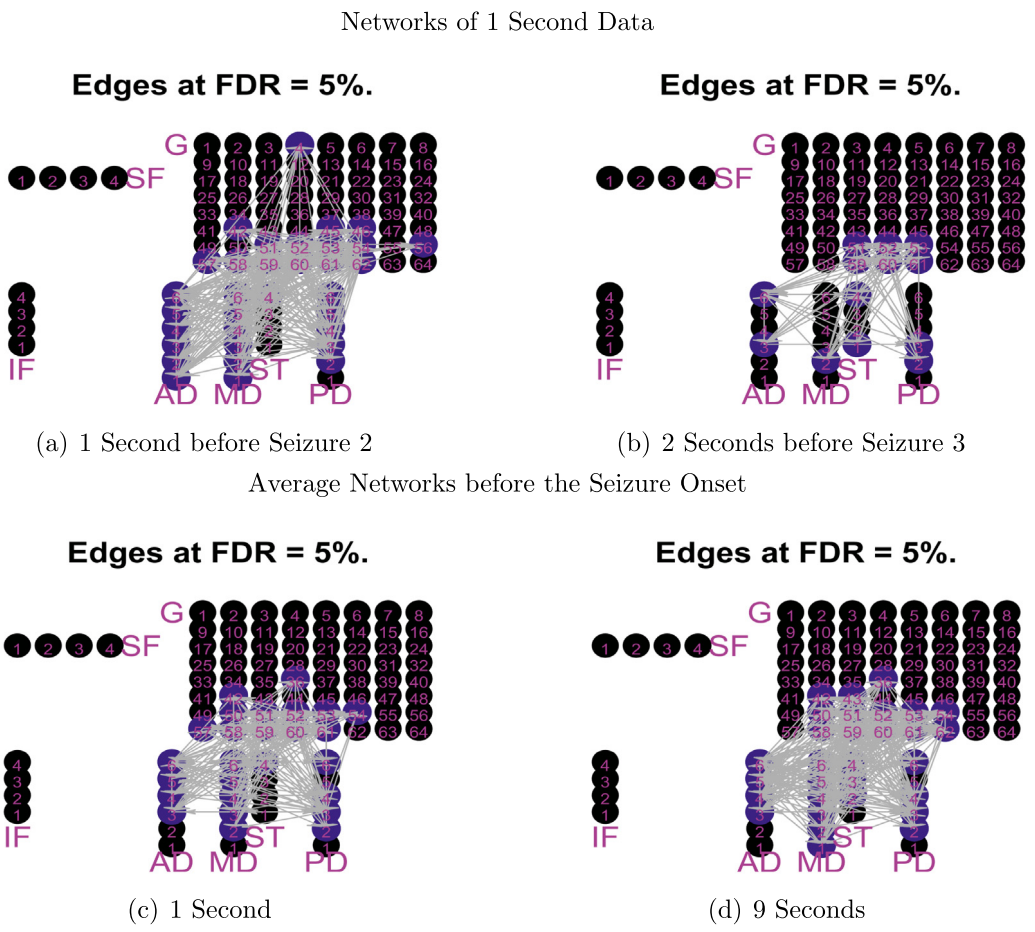
We focus on iEEG data 10 s before seizure onset times at the region G44. Following Burns et al. (2014), we divided iEEG time series into segments of 1-second length and applied the MODDM to each segment independently. We selected the directional network edges using the FDR of 5% so that the ensuring networks have sparse edges and are scientifically interpretable.

Figs. 4(a) and 4(b) show brain networks before seizure onsets of the 2nd and 3rd recorded seizures, 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 gray indicate directional connections among regions in the same cluster. 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. Moreover, the SOZ G44 tended to be disconnected from other regions around seizure onset times.

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 s of seizure onsets (one average for each 1-second). As an illustration, Figs. 4(c) and 4(d) show the average networks at 1 s and 9 s before seizure onset, 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. 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) and Warren et al. (2010). 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.

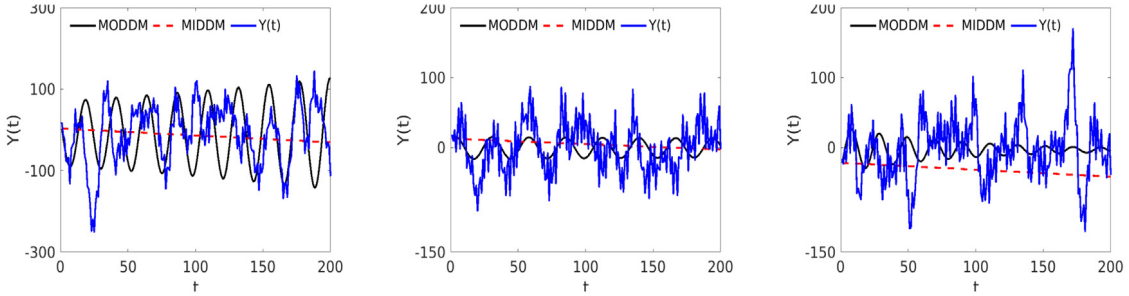
To demonstrate that the proposed MODDM for iEEG data can better characterize the brain's oscillatory activity than the existing first-order ODE model MIDDM (Zhang et al., 2017), we used the posterior estimates (posterior medians) of the model parameters to regenerate the state functions, which was the same approach used in the simulation study. Fig. 5 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) 1 Second before Seizure 2 (b) 2 Seconds before Seizure 3

(c) 1 Second (d) 9 Seconds

Fig. 4. 4(a) The brain network at 1 s before the 2nd seizure onset. 4(b) The brain network at 2 s before the 3rd seizure onset. 4(c) and 4(d) The average networks across 4 seizures at 1 s and 9 s 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 gray are for the pairs of regions in the same cluster with their network edges selected with the FDR of 5%.



(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

Fig. 5. Regenerated $\hat{\mathbf{x}}(t)$ using the posterior estimates of MODDM and MIDDM parameters for regions G1, G42 and G44.

6. Discussion

We have developed a new ODE model, MODDM, for directional connectivity among many brain regions recorded by iEEG. The MODDM 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 MODDM combines the strengths of scientific modeling and statistical modeling. We have shown through both simulation

study and real data analysis that the new model 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.

The use of the first-order Taylor expansion to approximate a complex function F_i in Eq. (2) is similar to the linear regression. In many complicated high-dimensional situations, the true underlying model is almost impossible to specify. The linear regression does not explain all possible variation of the response variable, but acts as a useful working model for explaining the relationship between the response and predictors, identifying useful predictors, and predicting the response variable. For the same reasons, our ODE model as well as many other linear ODE models in the literature (Chen and Wu, 2008; Xue et al., 2010; Lu et al., 2011; Wu et al., 2014a; Zhang et al., 2015, 2017) uses the first-order Taylor expansion to approximate the underlying complex dynamic system.

Despite that MODDM provides a much better fit to the oscillatory time series data than existing network models/methods, there remains data variation that the MODDM cannot explain. There are possibly two reasons. First, the MODDM is an 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 MODDM and the true underlying system. Second, like linear regressions, the use of linear expansion in the MODDM to approximate the underlying system is effective only for a short period. We, therefore, applied the MODDM 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 MODDM 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 MODDM has much better efficiency in detecting connected brain regions than the existing ODE model for iEEG data. Research is greatly needed to develop highly efficient statistical models and methods to provide a better fit to high-dimensional brain data.

The MODDM assumes no connection between modules and is focused on within-module connections only. The within-module connections tend to be short-range, strong, and dense, while between-module connections are long-range and sparse (Park and Friston, 2013). Thus, it is easier to detect within-module connections, and the MODDM is suitable for iEEG data that are usually collected from spatially close regions. Nevertheless, we can extend the MODDM to accommodate long-range connections by introducing additional indicators for between-module connections. Under this new model, each network edge can be either within-module or between-module connection. These two types of connections may have different densities and play different functional roles in the brain network. The within-module connections are those between the regions with a similar function while the between-module connections ensure integration among regions specializing in different functions. Differentiating, modeling, and estimating within-module and between-module directional connectivity will be an important topic in the future research.

Acknowledgments

Dr. Zhang's research is supported by NSF, United States of America-1758095 and the Quantitative Collaborative at the University of Virginia, United States of America.

Appendix A. Technical details for PCGS algorithm to sample from the posterior distribution

A.1. 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 (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 $\mathcal{G}_i = \{j, \delta(m_i, m_j) \cdot \gamma_{ij} \neq 0 \text{ and } j = 1, \ldots, d\}$. 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 $\Lambda_i(t) = \left(\mathbf{Z}_i(t)', 1, dx_i(t)/dt\right)' = \left(\mathbf{Z}_i(t)', 1, \tilde{\eta}_i'\ \boldsymbol{b}^{(1)}(t)\right)'$, 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

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

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

$$p(\boldsymbol{\theta}_i|-) \propto \exp\left\{-\frac{1}{2}\boldsymbol{\theta}_i'\left(\frac{1}{\tau_i}\int_0^T \Lambda_i(t)\,\Lambda_i'(t)\,dt + \frac{1}{\xi_0^2}\mathbb{I}\right)\boldsymbol{\theta}_i + \frac{1}{\tau_i}\int_0^T \frac{d^2x_i(t)}{dt^2}\Lambda_i'(t)\,dt\;\boldsymbol{\theta}_i\right\}$$
$$\cdot\ \exp\left\{-\frac{1}{2\tau_i}\int_0^T\left(\frac{d^2x_i(t)}{dt^2}\right)^2 dt\right\}, \tag{13}$$

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

Let $\mathbb{M}_i = \frac{1}{\tau_i}\int_0^T \Lambda_i'(t)\,\Lambda_i(t)\,dt + \frac{1}{\xi_0^2}\mathbb{I}$ and $\mathbb{V}_i = \frac{1}{\tau_i}\int_0^T \frac{d^2x_i(t)}{dt^2}\cdot\Lambda_i'(t)\,dt$. Based on Eq. (13), the posterior joint distribution after integrating out **A**, **D**, and **G** is

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

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

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

A.2. 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.3. Sequentially simulate γ_{ij}s from $p\left(\gamma_{ij}|\boldsymbol{m},\boldsymbol{\gamma}_{-ij},\boldsymbol{\tau},\boldsymbol{\eta}\right)$ for $i,j = 1,\ldots,d$

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

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

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 (4) 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.

A.4. Simulate $\boldsymbol{\theta}$ from $p(\boldsymbol{\theta}|\boldsymbol{m},\boldsymbol{\gamma},\boldsymbol{\tau},\boldsymbol{\eta})$

Based on the posterior conditional distribution of $\boldsymbol{\theta}_i$ (13),

$$A_{ij}|\delta\left(m_i,m_j\right)\cdot\gamma_{ij} = 0 \overset{i.i.d.}{\sim} \mathrm{N}\left(0,\xi^2\right) \text{ and}$$
$$\boldsymbol{\theta}_i|\boldsymbol{m},\boldsymbol{\gamma},\boldsymbol{\tau},\boldsymbol{\eta} \overset{ind}{\sim} \mathrm{MN}\left(\mathbb{M}_i^{-1}\mathbb{V}_i,\mathbb{M}_i^{-1}\right) \quad \text{for} \quad i = 1,\ldots,d.$$

A.5. Simulate $\boldsymbol{\tau}$ from $p(\boldsymbol{\tau}|\Theta_I,\boldsymbol{\eta})$

From the joint posterior distribution (12), we have

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

Appendix B. Proof of the normal distribution of the data $p(\eta|\Theta_l, \tau)$

Based on the model (8) for basis coefficients η, we have

$$p(\eta|\Theta_l, \tau) \propto \exp\left\{-\sum_{i=1}^{d}\frac{\mathrm{R}_i(\eta, \Theta_l)}{2\tau_i}\right\} \propto \exp\left\{-\frac{1}{2}\left(\eta' \Omega_{\Theta_l,\tau}\ \eta - 2\Lambda'_{\Theta_l,\tau}\ \eta + \Xi_{\Theta_l,\tau}\right)\right\}$$
$$\propto \exp\left\{-\frac{1}{2}\left(\eta - \Omega^{-1}_{\Theta_l,\tau}\Lambda_{\Theta_l,\tau}\right)' \Omega_{\Theta_l,\tau}\left(\eta - \Omega^{-1}_{\Theta_l,\tau}\Lambda_{\Theta_l,\tau}\right)\right\}. \tag{14}$$

Thus, from (14), $\eta|\Theta_l, \tau \sim \mathrm{MN}\left(\Omega^{-1}_{\Theta_l,\tau}\Lambda_{\Theta_l,\tau}, \Omega^{-1}_{\Theta_l,\tau}\right)$.

Notations of $\Omega_{\Theta_l,\tau}$, $\Lambda_{\Theta_l,\tau}$, and $\Xi_{\Theta_l,\tau}$ are introduced in Eq. (8), and we here derive their formulas conditioning on Θ_l and τ in the following.

Define vectors with $d \cdot L$ elements:

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

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 (5), the MODDM (4) can be rewritten as $\boldsymbol{E}_i(t)\ \eta - \Delta_i(t)\ \eta - D_i - G_i\cdot\boldsymbol{J}_i(t)\ \eta = 0$. Let $\mathbf{S}_i(t) = \boldsymbol{E}_i(t) - \Delta_i(t) - G_i\cdot\boldsymbol{J}_i(t)$. We have

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

Comparing the above to Eq. (8), we have

$$\Omega_{\Theta_l,\tau} = \sum_{i=1}^{d}\frac{1}{\tau_i}\int_0^T \mathbf{S}_i'(t)\,\mathbf{S}_i(t)\,dt, \quad \Lambda_{\Theta_l,\tau} = \sum_{i=1}^{d}\frac{1}{\tau_i}\int_0^T D_i\cdot\mathbf{S}_i'(t)\,dt, \quad \text{and}$$
$$\Xi_{\Theta_l,\tau} = \sum_{i=1}^{d}\frac{1}{\tau_i}\int_0^T D_i^2 dt.$$

References

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

Bard, Y., 1974. Nonlinear Parameter Estimation. Academic Press, New York.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 289–300.

Bhaumik, P., Ghosal, S., 2014. Bayesian estimation in differential equation models. Preprint, 1.

Biegler, L., Damiano, J., Blau, G., 1986. Nonlinear parameter estimation: a case study comparison. AIChE J. 32, 29–45.

Brunel, N., 2008. Parameter estimation of odes via nonparametric estimators. Electron. J. Stat. 2, 1242–1267.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (3), 186–198.

Burns, S., Santaniello, S., Yaffe, R., Jouny, C., Crone, N., Bergey, G., Anderson, W., Sarma, S., 2014. Network dynamics of the brain and influence of the epileptic seizure onset zone. Proc. Natl. Acad. Sci. USA 111 (49), 5321–5330.

Caffo, B., Peng, R., Dominici, F., Louis, T., Zeger, S., 2011. Parallel mcmc for analyzing distributed lag models with systematic missing data for an application in environmental epidemiology. In: Brooks, S., Gelman, A., Jones, G., Meng, X. (Eds.), The Handbook of Markov Chain Monte Carlo. CRC Press, pp. 493–511.

Campbell, D., 2007. Bayesian Collocation Tempering and Generalized Profiling for Estimation of Parameters from Differential Equation Models. McGill University, Montreal.

Cao, J., Huang, J., Wu, H., 2012. Penalized nonlinear least squares estimation of time-varying parameters in ordinary differential equations. J. Comput. Graph. Statist. 21, 42–56.

Cervenka, M., Franaszczuk, P., Crone, N., Hong, B., Caffo, B., Bhatt, P., Lenz, F., Boatman-Reich, D., 2013. Reliability of early cortical auditory gamma-band responses. Clin. Neurophysiol. 124 (1), 70–82.

Chen, J., Wu, H., 2008. Efficient local estimation for time-varying coefficients in deterministic dynamic models with applications to hiv-1 dynamics. J. Amer. Statist. Assoc. 103, 369–384.

Daunizeau, J., David, O., Stephan, K., 2011. Dynamic causal modelling: A critical review of the biophysical and statistical foundations. NeuroImage 58, 312–322.

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

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

Deuflhard, P., Bornemann, F., 2000. Scientific Computing with Ordinary Differential Equations. Springer, New York.

Efron, B., Tibshirani, R., 2002. Empirical bayes methods and false discovery rates for microarrays. Genet. Epidemiol. 23 (1), 70–86.
Fell, J., Axmacher, N., 2011. The role of phase synchronization in memory processes. Nat. Rev. Neurosci. 12 (2), 105–118.
Fitzpatrick, R., 2013. Oscillations and Waves: An Introduction. CRC Press.
Földiák, P., Young, M.P., 1995. Sparse coding in the primate cortex. In: The Handbook of Brain Theory and Neural Networks. The MIT Press, 895-898, editor.
Fries, P., 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn. Sci. 9 (10), 474–480.
Friston, K., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19, 1273–1302.
Gelman, A., Bois, F., Jiang, J., 1996. Physiological pharamacokinetic analysis using population modeling and informative prior distributions. J. Amer. Statist. Assoc. 91, 1400–1412.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2014. Bayesian Data Analysis, Vol. 2. Chapman & Hall/CRC, Boca Raton, FL, USA.
Girolami, M., 2008. Bayesian inference for differential equations. Theoret. Comput. Sci. 408, 4–16.
Härdle, W., 1990. Applied Nonparametric Regression, Number 19. Cambridge university press.
Hemker, P., 1972. Numerical methods for differential equations in system simulations and in parameter estimation. Anal. Simul. Biochem. Syst. 59–80.
Huang, Y., Liu, D., Wu, H., 2006. Hierarchical bayesian methods for estimation of parameters in a longitudinal hiv dynamic system. Biometrics 62, 413–423.
Huang, Y., Wu, H., 2006. A bayesian approach for estimating antiviral efficacy in hiv dynamic models. J. Appl. Stat. 33, 155–174.
Kiebel, S., David, O., Friston, K., 2006. Dynamic causal modelling of evoked responses in eeg/meg with lead-field parameterization. NeuroImage 30, 1273–1284.
Kramer, M., Eden, U., Kolaczyk, E., Zepeda, R., Eskandar, E., Cash, S., 2010. Coalescence and fragmentation of cortical networks during focal seizures. J. Neurosci. 30 (30), 10076–10085.
Kramer, M., Kolaczyk, E., Kirsch, H., 2008. Emergent network topology at seizure onset in humans. Epilepsy Res. 79 (2–3), 173–186.
Kramer, M., Truccolo, W., Eden, U., Lepage, K., Hochberg, L., Eskandar, E., Madsen, J., Lee, J., Maheshwari, A., Halgren, E., Chu, C., Cash, S., 2012. Human seizures self-terminate across spatial scales via a critical transition. Proc. Natl. Acad. Sci. USA 109 (51), 21116–21121.
Li, Z., Osborne, M., Pravan, T., 2005. Parameter estimation of ordinary differential equations. IMA J. Numer. Anal. 25, 264–285.
Lu, T., Liang, H., Li, H., Wu, H., 2011. High dimensional odes coupled with mixed-effects modeling techniques for dynamic gene regulatory network identification. J. Amer. Statist. Assoc. 106, 1242–1258.
Matteij, R., Molenaar, J., 2002. Ordinary Differential Equations in Theory and Practice. SIAM, Philadelphia.
Micheloyannis, S., 2012. Graph-based network analysis in schizophrenia. World J. Psychiatry 2 (1), 1–12.
Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M., Alon, U., 2004. Superfamilies of evolved and designed networks. Science 303 (5663), 1538–1542.
Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U., 2002. Network motifs: Simple building blocks of complex networks. Science 298 (5594), 824–827.
Mormann, F., Kreuz, T., Rieke, C., Andrzejak, R., Kraskov, A., David, P., Elger, C., Lehnertz, K., 2005. On the predictability of epileptic seizures. Clin. Neurophysiol. 116 (3), 569–587.
Netoff, T., Schiff, S., 2002. Decreased neuronal synchronization during experimental seizures. J. Neurosci. 22 (16), 7297–7307.
Newman, M., 2006. Modularity and community structure in networks. Proc. Natl. Acad. Sci. USA 103 (23), 8577–8696.
Nissen, I.A., van Klink, N.E., Zijlmans, M., Stam, C.J., Hillebrand, A., 2016. Brain areas with epileptic high frequency oscillations are functionally isolated in meg virtual electrode networks. Clin. Neurophysiol. 127 (7), 2581–2591.
Olshausen, B., Field, D., 2004. Sparse coding of sensor inputs. Curr. Opin. Neurobiol. 14, 481–487.
Park, H.-J., Friston, K., 2013. Structural and functional brain networks: From connections to cognition. Science 342 (6158).
Poyton, A., Varziri, M., McAuley, K., McLellan, P., Ramsay, J., 2006. Parameter estimation in continuous dynamic models using principal differential analysis. Comput. Chem. Eng. 30, 698–708.
Qi, X., Zhao, H., 2010. Asymptotic efficiency and finite-sample properties of the generalized profiling estimation of the parameters in ordinary differential equations. Ann. Statist. 38, 435–481.
Ramsay, J.O., 2006. Functional Data Analysis. Wiley Online Library.
Ramsay, J., Hooker, G., Campbell, D., Cao, J., 2007. Parameter estimation for differential equations: A generalized smoothing approach (with discussion). J. R. Stat. Soc. B 69, 741–796.
Ramsay, J., Silverman, B., 2005. Functional Data Analysis. Springer, New York.
Ruppert, D., 2002. Selecting the number of knots for penalized splines. J. Comput. Graph. Stat. 11 (4), 735–757.
Schiff, S., Sauer, T., Kumar, R., Weinstein, S., 2005. Neuronal spatiotemporal pattern discrimination: The dynamical evolution of seizures. Neuroimage 28 (4), 1043–1055.
Schindler, K., Amor, F., Gast, H., Müller, M., Stibal, A., Mariani, L., Rummel, C., 2010. Peri-ictal correlation dynamics of high-frequency (80-200 hz) intracranial eeg. Epilepsy Res. 89 (1), 72–81.
Schindler, K., Bialonski, S., Horstmann, M., Elger, C., Lehnertz, K., 2008. Evolving functional network properties and synchronizability during human epileptic seizures. Chaos 18 (3), 033119.
Schindler, K., Leung, H., Elger, C., Lehnertz, K., 2007. Assessing seizure dynamics by analysing the correlation structure of multichannel intracranial eeg. Brain 130 (Pt 1), 65–77.
Schnitzler, A., Gross, J., 2005. Normal and pathological oscillatory communication in the brain. Nat. Rev. Neurosci. 6 (4), 285–296.
Schuster, P., 1983. Biophysics, ed. by Walter Hoppe Ltd. Springer-Verlag, Berlin.
Serway, R., Jewett, J., 2003. Oscillations and mechanical waves. In: Physics for Scientists and Engineers. Brooks/Cole, pp. 433–543.
Sporns, O., 2011. The MIT Press, Cambridge, Massachusetts, p. 71.
Suchard, M., Wang, Q., Chan, C., Frelinger, J., Cron, A., West, M., 2010. Understanding gpu programming for statistical computation: Studies in massively parallel massive mixtures. J. Comput. Graph. Statist. 19, 419–438.
van Dyk, D., Park, T., 2008. Partially collapsed gibbs samplers: Theory and methods. J. Amer. Statist. Assoc. 103, 790–796.
Varah, J., 1982. A spline least squares method for numerical parameter estimation in differential equations. SIAM J. Sci. Comput. 3, 28–46.
Warren, C.P., Hu, S., Stead, M., Brinkmann, B.H., Bower, M.R., Worrell, G.A., 2010. Synchrony in normal and focal epileptic brain: the seizure onset zone is functionally disconnected. J. Neurophysiol. 104 (6), 3530–3539.
Wendling, F., Bellanger, J., Badier, J., Coatrieux, J., 1996. Extraction of spatio-temporal signatures from depth eeg seizure signals based on objective matching in warped vectorial observations. IEEE Trans. Biomed. Eng. 43 (10), 990–1000.
Wu, L., Gotman, J., 1998. Segmentation and classification of eeg during epileptic seizures. Electroencephalogr. Clin. Neurophysiol. 106 (4), 344–356.
Wu, H., Lu, T., Xue, H., Liang, H., 2014a. Sparse additive ordinary differential equations for dynamic gene regulatory network modeling. J. Amer. Statist. Assoc. 109, 700–716.
Wu, S., Xue, H., Wu, Y., Wu, H., 2014b. Variable selection for sparse high-dimensional nonlinear regression models by combining nonnegative garrote and sure independence screening. Statist. Sinica 24 (3), 1365.

Xue, H., Miao, H., Wu, H., 2010. Sieve estimation of constantand time- varying coefficients in nonlinear ordinary differential equation models by considering both numerical error and measurement error. Ann. Statist. 38, 2351–2387.

Zhang, T., Wu, J., Li, F., Caffo, B., Boatman-Reich, D., 2015. A dynamic directional model for effective brain connectivity using electrocorticographic (ecog) time series. J. Amer. Statist. Assoc. 110, 93–106.

Zhang, T., Yin, Q., Caffo, B., Sun, Y., Boatman-Reich, D., 2017. Bayesian inference of high-dimensional, cluster-structured ordinary differential equation models with applications to brain connectivity studies. Ann. Appl. Stat. 11 (2), 868–897.