Bayesian Inference for High-Dimensional ODE Models with Applications to Brain Connectivity Studies

Banff Neuroimaging Data Analysis Workshop

Tingting Zhang
Department of Statistics
University of Virginia

Introduction

- The brain's functional organization is governed by two principles: functional specialization and functional integration (Friston, 2011).

- Functional specialization suggests that different brain areas are specialized for different functions.

- Functional integration refers to interactions among different specialized brain areas and how these interactions depend on different sensorimotor or cognitive information the brain is processing.

ODE Models

- It is biophysically natural to use ODEs to characterize the functional interactions among different regions.

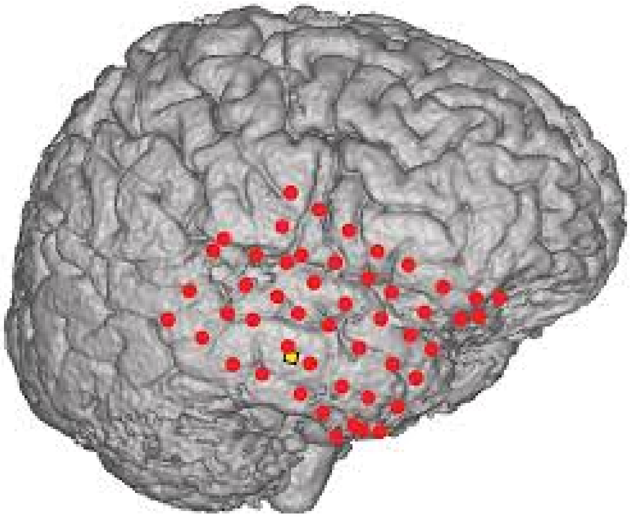
- Existing ODE models (Dynamic Causal Modeling, Daunizeau et al., 2011; David & Friston, 2003) for fMRI and EEG data.

1. Focus on connectivity among only a few regions.
2. The ODE formulation highly relies on the prior knowledge of the existence and strength of the connectivity between regions under study.

A High-dimensional ODE Model for ECoG Data

- ECoG, or intracranial EEG, is a form of electrophysiology whereby electrodes are placed directly (inside the skull and dura) on a living human cortex in the process of surgery for epilepsy care.

- ECoG's high temporal resolution and spatial localization make it an ideal dataset for building brain connectivity models.



Dynamic Directional Model (DDM)

- Neuronal Electrical State

$$\begin{aligned}\frac{dx_1(t)}{dt} &= A_{11}\, x_1(t)\cdot(1-u(t)) + \ldots + + A_{1d}\, x_d(t)\cdot(1-u(t)) \\ &+ B_{11}\, x_1(t)\cdot u(t) + \ldots + B_{1d}\, x_d(t)\cdot u(t) + C_1\cdot u(t) + D_1 \\ &\vdots \\ \frac{dx_d(t)}{dt} &= A_{d1}\, x_1(t)\cdot(1-u(t)) + \ldots + + A_{dd}\, x_d(t)\cdot(1-u(t)) \\ &+ B_{d1}\, x_1(t)\cdot u(t) + \ldots + B_{dd}\, x_d(t)\cdot u(t) + C_1\cdot u(t) + D_1\end{aligned}$$

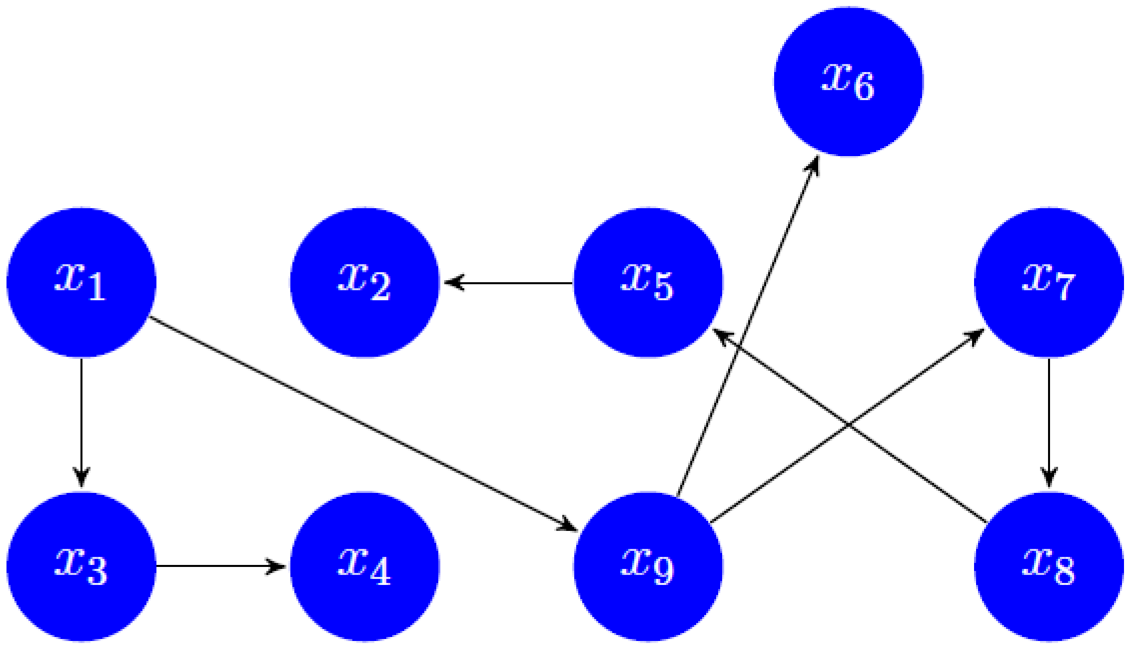
- Observation Model

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

Sparsity Assumption

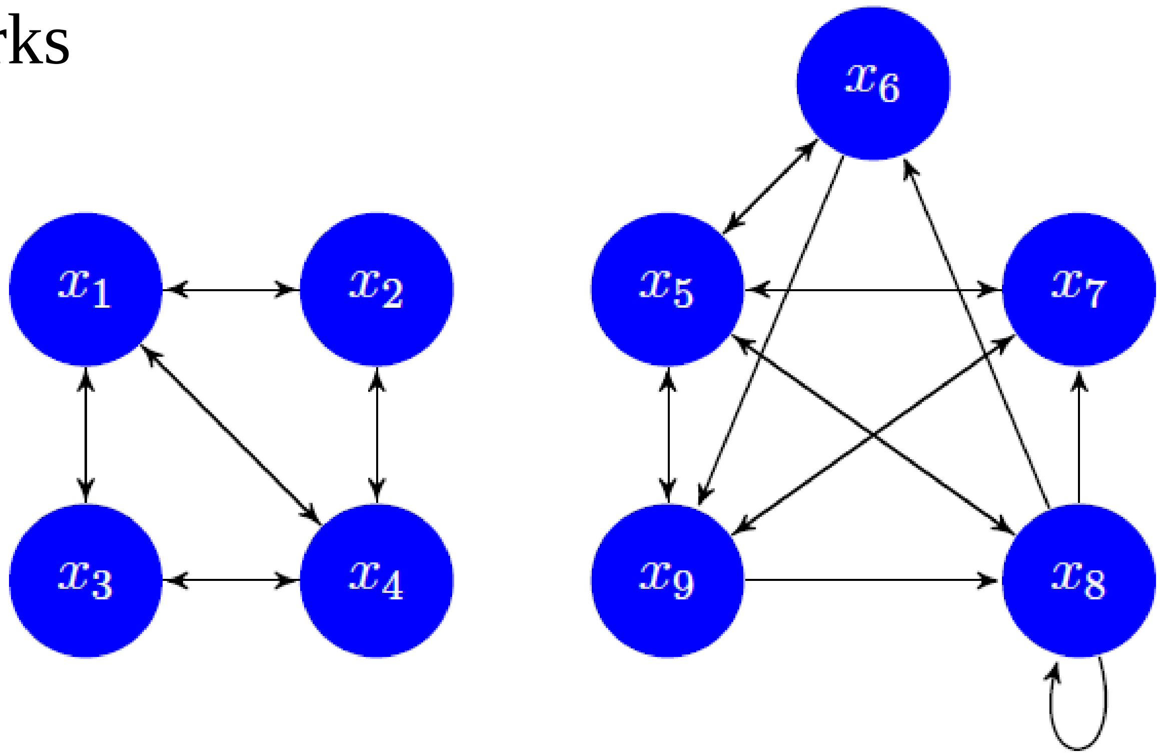
- Connections take up energy and space
- Economical Model
- So many coefficients in $\boldsymbol{A}$ and $\boldsymbol{B}$ are zeroes.

Different Sparse Network Structures



Different Sparse Network Structures

- The community/cluster structure (modularity) in brain networks



Modular and Indicator Based DDM (MIDDM)

- DDM

$$\frac{dx_{i_1}(t)}{dt} = \sum_{i_2=1}^{d} A_{i_1 i_2} \cdot x_{i_2}(t) \cdot (1 - u(t)) + \sum_{i_2=1}^{d} B_{i_1 i_2} \cdot x_{i_2}(t) \cdot u(t) + C_{i_1} \cdot u(t) + D_{i_1}.$$

- MIDDM

$$\frac{dx_{i_1}(t)}{dt} = \sum_{i_2=1}^{d} \delta(m_{i_1}, m_{i_2}) \cdot \gamma^{A}_{i_1 i_2} \cdot A_{i_1 i_2} \cdot x_{i_2}(t) \cdot (1 - u(t)) + \sum_{i_2=1}^{d} \delta(m_{i_1}, m_{i_2}) \cdot \gamma^{B}_{i_1 i_2} \cdot B_{i_1 i_2} \cdot x_{i_2}(t) \cdot u(t) + C_{i_1} \cdot u(t) + D_{i_1};$$

- The MIDDM, assuming different properties for connectivity within and between modules, is hierarchical, in contrast to typical single-layer ODE models.

- The proposed new ODE model, motivated by statistical considerations, is considered an approximation rather than a principle for the underlying mechanism. It is important to account for model uncertainty when estimating the ODE model.

- We propose a Bayesian approach for two reasons.

- It is natural to characterize this multilevel structure within a unified Bayesian framework, as simultaneous variable selection and clustering in multiple regression were often addressed in Bayesian texts, such as Tadesse et al. (2005); Kim et al. (2006) and Dunson et al. (2008).

- Second, the ODE model uncertainty can be naturally quantified and incorporated into parameter estimation within a Bayesian framework.

- Kennedy and O’Hagan (2001) have developed a Bayesian framework to quantify various sources of uncertainty in approximating systems with complex mathematical models.

- Chkrebtii et al. (2015) and Conrad et al. (2015) developed approaches within this framework to quantify discretization uncertainty of ODE models.

Bayesian Hierarchical Model for Making Inferences of MIDDM

- Neuronal State Model

$$\begin{aligned}\frac{dx_{i_1}(t)}{dt} &= \sum_{i_2=1}^{d} \delta(m_{i_1}, m_{i_2}) \cdot \gamma_{i_1 i_2}^{A} \cdot A_{i_1 i_2} \cdot x_{i_2}(t) \cdot (1 - u(t)) \\ &+ \sum_{i_2=1}^{d} \delta(m_{i_1}, m_{i_2}) \cdot \gamma_{i_1 i_2}^{B} \cdot B_{i_1 i_2} \cdot x_{i_2}(t) \cdot u(t) + C_{i_1} \cdot u(t) + D_{i_1};\end{aligned}$$

- Observation Model

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

Differential Equation Model Estimation

1. Discretization methods using numerical approximation (Biegler et al., 1986; Campbell, 2007; Gelman et al., 1996).

2. Basis function expansion (Deuflhard & Bornemann, 2000; Poyton et al., 2006; Ramsay & Silverman, 2005; Ramsay et al., 2007; Varah, 1982).

Bayesian MIDDM

- Represent **x**(t) by a set of spline bases:

$$\mathbf{x}(t) = \mathbf{\Gamma}\, \boldsymbol{\phi}(t),$$

$\boldsymbol{\phi}(t) = (\phi_1(t), \ldots, \phi_p(t))'$ is a vector of basis functions

- Model for the observed data

$$Y_i | \mathbf{\Gamma}[i,] \overset{\text{ind}}{\sim} \text{MN}(\mathbf{\Phi}\, (\mathbf{\Gamma}[i,])', \sigma_i^2 \mathbf{I}_T) \quad \text{for } i = 1, 2 \ldots, d$$

Prior specification for the basis coefficients

$$\boldsymbol{\Theta}_I = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \boldsymbol{m}, \boldsymbol{\gamma}^A, \boldsymbol{\gamma}^B\} \qquad \boldsymbol{\eta} = (\boldsymbol{\Gamma}[1,], \ldots, \boldsymbol{\Gamma}[d,])'$$

$$p(\boldsymbol{\eta}|\boldsymbol{\Theta}_I, \tau) \propto \exp\left\{ -\frac{1}{2\tau} \mathrm{R}(\boldsymbol{\eta}, \boldsymbol{\Theta}_I) \right\},$$

$$\mathrm{R}(\boldsymbol{\eta}, \boldsymbol{\Theta}_I) =$$

$$\sum_{i_1=1}^{d} \int_0^T \left(\frac{dx_{i_1}(t)}{dt} - \sum_{i_2=1}^{d} \delta(m_{i_1}, m_{i_2}) \cdot \gamma_{i_1 i_2}^A \cdot A_{i_1 i_2} \cdot x_{i_2}(t) \cdot (1 - u(t)) \right.$$
$$\left. - \sum_{i_2=1}^{d} \delta(m_{i_1}, m_{i_2}) \cdot \gamma_{i_1 i_2}^B \cdot B_{i_1 i_2} \cdot x_{i_2}(t) \cdot u(t) - C_{i_1} \cdot u(t) - D_{i_1} \right)^2 dt.$$

Prior specification for the basis coefficients

- Since $\mathrm{R}(\boldsymbol{\eta}, \Theta_I) = \boldsymbol{\eta}'\Omega_{\Theta_I}\boldsymbol{\eta} - 2\Lambda'_{\Theta_I}\boldsymbol{\eta} + \Xi_{\Theta_I}$,

$$\boldsymbol{\eta} | \Theta_I, \tau \sim \mathrm{MN}(\Omega_{\Theta_I}^{-1}\Lambda_{\Theta_I}, \tau \cdot \Omega_{\Theta_I}^{-1}).$$

Prior Specification for Indicators

$$p(\boldsymbol{\gamma}^A, \boldsymbol{\gamma}^B | \boldsymbol{\theta}, \boldsymbol{m}, \tau) \quad \propto$$

$$\det(\Omega_{\Theta_I})^{-1/2} \cdot \exp\{\frac{1}{2\tau}(\boldsymbol{\Lambda}'_{\Theta_I} \boldsymbol{\Omega}^{-1}_{\Theta_I} \boldsymbol{\Lambda}_{\Theta_I} - \Xi_{\Theta_I})\}$$
$$p_0^{\sum_{i,j} \gamma_{ij}^A + \sum_{i,j} \gamma_{ij}^B} \cdot (1 - p_0)^{2d^2 - \sum_{i,j} \gamma_{ij}^A - \sum_{i,j} \gamma_{ij}^B},$$

Priors for other MIDDM parameters

$$P(\boldsymbol{m}) \quad \propto \quad \exp\{-\mu \cdot \sum_{i_1,i_2=1}^{d} \delta(m_{i_1}, m_{i_2})\},$$

$$A_{ij}, B_{ij}, C_i, D_i \overset{\text{i.i.d}}{\sim} N(0, \xi_0^2) \quad \text{and}$$

$$p(\sigma_i^2) \propto 1/\sigma_i^2, \ \ \text{for } i, j = 1, 2, \ldots, d,$$

Joint Posterior Distribution

$$p(\boldsymbol{\eta}, \boldsymbol{\Theta}_I, \boldsymbol{\sigma}^2 | \mathbf{Y}, \tau, \mu) \propto$$

$$\prod_{i=1}^{d} \sigma_i^{-T} \exp\{-\frac{(Y_i - \boldsymbol{\Phi}\,\boldsymbol{\Gamma}[i,]')^2}{2\sigma_i^2}\} \cdot \exp\Big\{ - \frac{1}{2\tau} \mathrm{R}(\boldsymbol{\eta}, \boldsymbol{\Theta}_I) \Big\}$$

$$\cdot \quad \exp\{-\mu \sum_{i_1, i_2 = 1}^{d} \delta(m_{i_1}, m_{i_2})\}$$

$$\cdot\, p_0^{\sum_{i,j} \gamma_{ij}^A + \sum_{i,j} \gamma_{ij}^B} \cdot (1 - p_0)^{2d^2 - \sum_{i,j} \gamma_{ij}^A - \sum_{i,j} \gamma_{ij}^B}$$

$$\cdot \quad \prod_{i,j=1}^{d} \phi(\frac{A_{ij}}{\xi_0}) \cdot \prod_{i,j=1}^{d} \phi(\frac{B_{ij}}{\xi_0}) \cdot \prod_{i=1}^{d} \phi(\frac{C_i}{\xi_0}) \cdot \prod_{i=1}^{d} \phi(\frac{D_i}{\xi_0}) \cdot \prod_{i=1}^{d} \frac{1}{\sigma_i^2},$$

Partially Collapsed Gibbs Sampler (PCGS; van Dyk and Park, 2008)

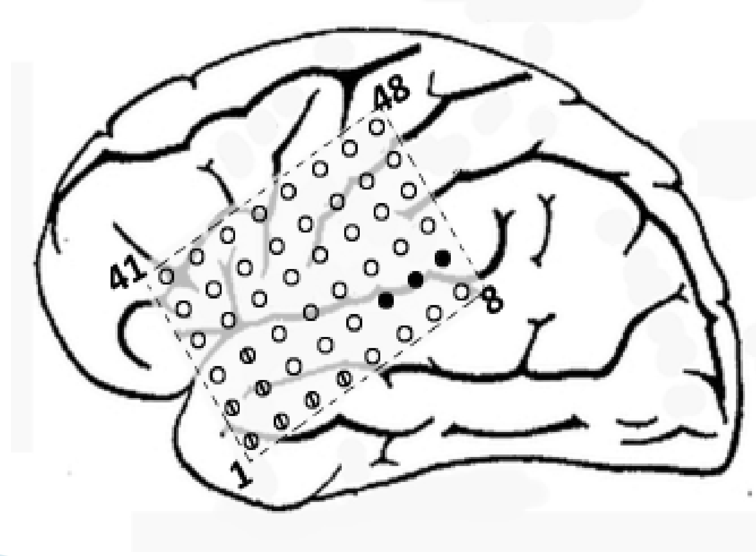
$\Theta_I = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \boldsymbol{m}, \gamma^A, \gamma^B\}$ $\boldsymbol{\eta} = (\boldsymbol{\Gamma}[1,], \ldots, \boldsymbol{\Gamma}[d,])'$

$\boldsymbol{\theta} = \{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$

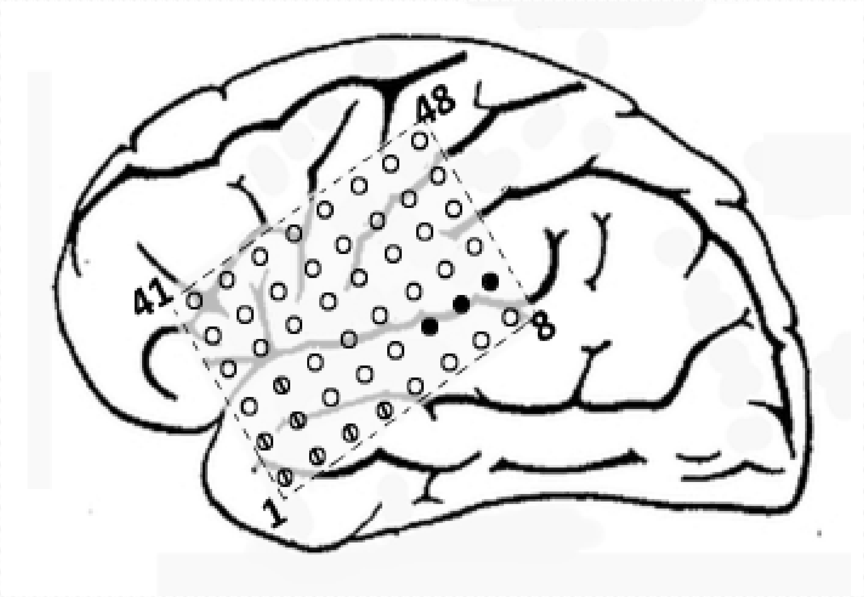
1. Draw from $p(m_i | \boldsymbol{m}_{-i}, \boldsymbol{\eta}, \boldsymbol{\sigma}^2, \gamma^A, \gamma^B, \mathbf{Y})$
2. Draw from $p(\gamma^A_{ij} | \boldsymbol{m}, \boldsymbol{\eta}, \boldsymbol{\sigma}^2, \gamma^A_{-ij}, \gamma^B, \mathbf{Y})$
3. Draw from $p(\gamma^B_{ij} | \boldsymbol{m}, \boldsymbol{\eta}, \boldsymbol{\sigma}^2, \gamma^A, \gamma^B_{-ij}, \mathbf{Y})$
4. Draw $\boldsymbol{\theta}$ from $p(\boldsymbol{\theta} | \boldsymbol{m}, \boldsymbol{\eta}, \boldsymbol{\sigma}^2, \gamma^A, \gamma^B, \mathbf{Y})$
5. Draw $\sigma_1^2, \ldots, \sigma_d^2$ from $p(\boldsymbol{\sigma}^2 | \Theta_I, \boldsymbol{\eta}, \mathbf{Y})$,
6. Draw $\boldsymbol{\eta}$ from $p(\boldsymbol{\eta} | \Theta_I, \boldsymbol{\sigma}^2, \mathbf{Y})$

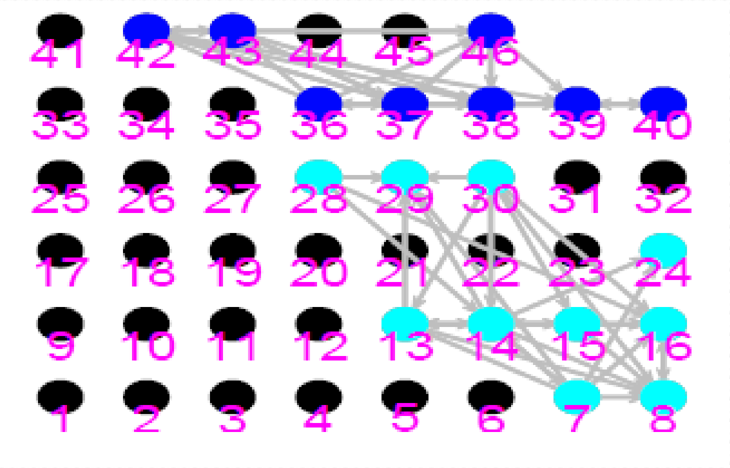
Real ECoG Data

- 45 Recording Channels/Nodes
- Cover Auditory Cortex
- Cover Epileptic Areas
- Use auditory stimulus of 1000 Hz
- Consist of 254 Trials
- Each trial lasted 250 ms including 50 ms of auditory stimulus



Identified Clusters





Each edge represents a top 5% selection probability.

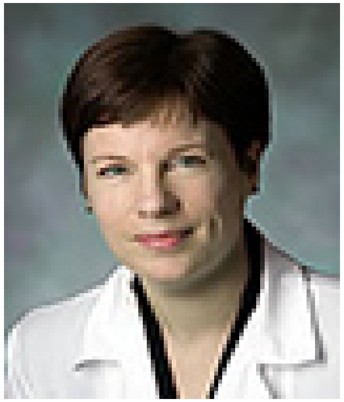
Summary

- Propose a new ODE model for the brain’s functional integration.

- Develop a new Bayesian framework for inferring a high-dimensional ODE model for a complex system.

Acknowledgement

- Brian Caffo and Dana Boatman from Johns Hopkins University



- My Student Qiannan Yin

Acknowledgement

- Support from SAMSI.

- ***Thank you!***

- ***Questions?***