A Bayesian Model for Brain Network

Functional Connectivity using PyMC3

By

Rui Wang

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Biostatistics

August 10, 2018

Nashville, Tennessee

Approved:

Hakmook Kang, Ph.D.

Qingxia Chen, Ph.D.

To my beloved family.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Our brain network, as a complex integrative system, consists of many different regions. Each region has its own task and function and simultaneously shares structural and functional information. With the developed imaging techniques such as functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI), researchers can investigate the underlying brain functions related to human behaviors and some diseases or disorders in the nervous system such as major depressive disorder (MDD).

In this thesis, we developed a Bayesian hierarchical spatiotemporal model that combined fMRI and DTI data jointly to enhance the estimation of resting-state functional connectivity. Structural connectivity from DTI data was utilized to construct an informative prior for functional connectivity based on resting-state fMRI data through the Cholesky decomposition in a mixture model. The analysis took the advantages of probabilistic programming package as PyMC3 and next-generation Markov Chain Monte Carlo (MCMC) sampling algorithm as No-U-Turn Sampler (NUTS). The simulation study with this advanced algorithm, illustrated reduced mean squared errors (MSEs) of estimation. Furthermore, through a case study of MDD, we applied our model to examine how the estimated functional connectivity was associated with tasks of episodic memory, executive function, processing speed and working memory.

CHAPTER 1

INTRODUCTION TO BRAIN CONNECTIVITY

1.1 Brain Imaging

Our brain consists of many different regions that have each own task and function [1]. As a complex integrative system, the brain network is formed when information is simultaneously processed and transported between structurally and functionally linked brain regions. The studies of brain network apply various techniques to either directly or indirectly image the structure and function of the nervous system. It is highly associated with the fields of medicine, neuroscience and psychology. Currently there are two main categories: one is structural imaging, involving with the structure of the nervous system and the diagnosis of intracranial diseases and disorders on a larger scale; the other is functional imaging, related with the diagnosis of metabolic diseases on a finer scale such as Alzheimer's disease. It is fundamental for humans to understand how a healthy nervous system can produce activities such as thought, emotion and physical behavior, and regulate body functions to investigate some diseases and disorders.

In the recent decades, magnetic resonance imaging (MRI), a highly versatile imaging technique, was created by Peter Mansfield and Paul Lauterbur to capture the pictures of anatomy and the physiological process of the body in both health and disease. Due to their advanced breakthrough, they were awarded the 2003 Nobel Prize in Physiology or Medicine for their "discoveries concerning magnetic resonance imaging". In the 1990s, functional magnetic resonance imaging (fMRI) was developed and it has been leading the brain mapping field because of its low invasiveness, lack of radiation exposure, relatively wide availability [2].

With all the developed imaging techniques, the studies of brain are focused on the structural and functional connectivity between two or more brain regions. Structural connectivity (also called anatomical connectivity) is referred as a structural link or the existence of neural connections between two regions while functional connectivity is defined as the temporal correlation between spatially remote neurophysiological events [3]. In other words, two systems can be assumed to be functionally connected if they illustrate synchronized or correlated patterns of activity [4]. The relationship between structural and functional connectivity is illustrated in Figure 1.1 in the form of a Venn diagram [4]. It highlights that the structural information could offer plausible arguments for the functional information.
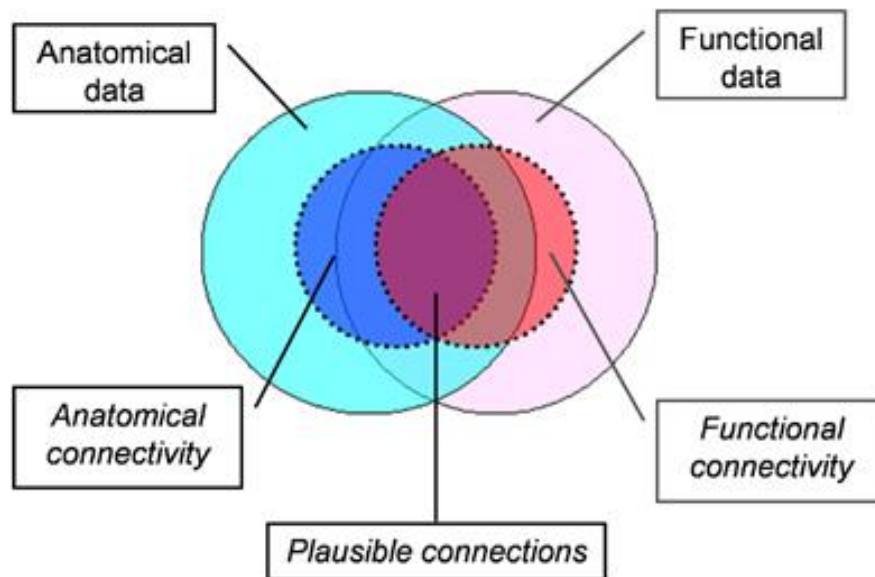
Figure 1.1 Venn diagram showing the studies combining functional and anatomical data, with focus on anatomical and functional connectivity data. Image reprinted with permission from Ref [4]. Copyright © 2007, John Wiley and Sons.

1.2 Functional Connectivity
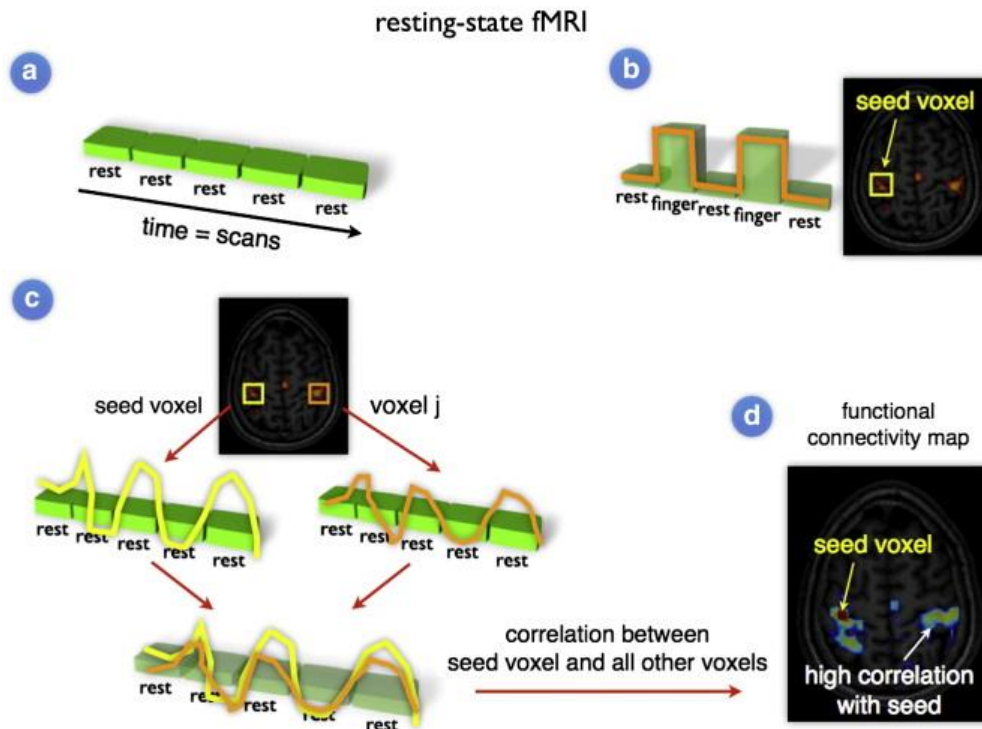
Figure 1.2 Resting-state fMRI studies capture the correlation between spontaneous activation patterns of brain regions. (a) The blood oxygen level-dependent (BOLD) signal is measured throughout the experiment. (b) Conventional task-dependent fMRI can be used to select a seed region of interest. (c) To examine the level of functional connectivity, the resting-state time-series of the seed voxel i is correlated with the resting-state time-series of region j. (d) Furthermore, to map out all functional connections of the selected seed region, the time-series of the seed voxel i can be correlated with the time-series of all other voxels in the brain, resulting in a functional connectivity map that reflects the regions that show a high level of functional connectivity with the selected seed region. Image reprinted with permission from Ref [1]. Copyright © 2010, Elsevier Inc.

Functional connectivity is typically measured by analyzing the patterns of concurrent activity between various brain areas that share functional properties. It is also defined as the

temporal dependency between spatially remote neurophysiological events [5, 6]. With deep understanding of fMRI in the recent two decades, more and more studies imply functional connectivity between brain areas as the level of co-activation of spontaneous fMRI time-series recorded at rest [7-9]. In the resting-state experiments, their level of spontaneous brain activity was recorded when subjects were required to relax without thinking anything. Typically, they are placed into the scanner, asked to close their eyes and to think nothing without falling asleep. Experimental evidences demonstrate that the left and right hemispheric regions of the primary motor network at rest show a high correlation between their fMRI time-series [9, 10], indicating the ongoing information transporting and ongoing functional connectivity between these areas at rest [7, 8, 10, 11]. Figure 1.2 illustrates that the resting-state timeseries of a voxel in the motor network was correlated with the resting-state time-series of all other brain voxels, suggesting a high correlation between the spontaneous neuronal activation patterns of these areas. Therefore, resting-state fMRI research focuses on mapping functional communication channels between different brain regions by leveraging the correlated dynamics of fMRI time-series.

1.3 Structural Connectivity

Structural connectivity is typically pointed as the illustration of fiber tracts directly connecting different brain regions. Because of the neuronal axons in these fiber tracts, they can transmit the neural signals across all the brain areas, allowing for communications between brain regions. Previously, it has been applied to study animal with histological methods while it could not reveal much information in humans. Thanks to the advanced development of *in vivo* imaging techniques, researchers are now able to visualize the white matter. The volumetric estimates of while matter can be used as measures of structural

connectivity [12]. For example, it was found that the degree of while matter preservation in the groups of older adults is highly associated with their performance on tasks that require functional integration involving interhemispheric interactions [4].
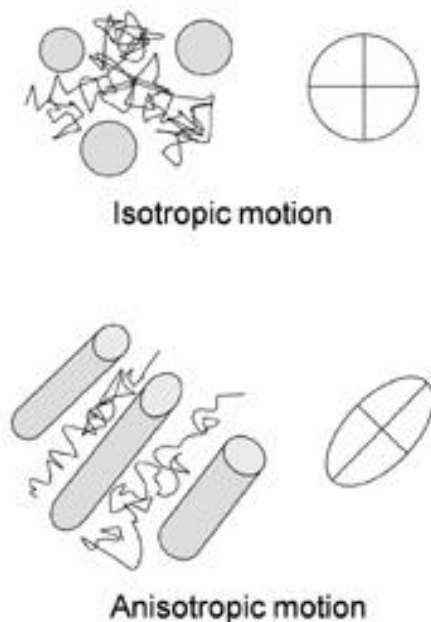
Figure 1.3 Schematic of the basic principles underlying DTI: isotropy and anisotropy of water motion in tissue. The ellipsoids represent the directionality and the degree of anisotropy. The axes in these ellipsoids are oriented along the diffusion tensor eigenvectors, and the lengths of these axes are proportional to the amount of diffusivity (corresponding eigenvalues) in the respective dimensions. Image reprinted with permission from Ref [4]. Copyright © 2007, John Wiley and Sons.

In the recent few years, a new imaging technology, diffusion tensor imaging (DTI) has been developed as a promising method for describing the structural connectivity *in vivo*. DTI is significant when the neural axons of while matter in the brain has an internal fibrous structure. Water will then diffuse more rapidly in the direction aligned with the internal structure while it moves more slowly in the perpendicular direction. So, DTI yields images

of the anisotropy of water diffusion in the living tissue. Figure 1.3 demonstrates the isotropy and anisotropy of water motion in tissue. Due to the increasingly accurate estimation of fiber orientation/strength and the widespread potential implications, DTI technology will shed light on the development of brain research in the future.

CHAPTER 2

SPATIOTEMPORAL HIERARCHICAL MODEL

2.1 Existing Methods

As introduced in CHAPTER 1, functional connectivity is used as a biomarker for neurological and psychiatric disorders such as Alzheimer's diseases [13, 14] and bipolar disorders [15, 16]. White matter tracts are the structural pathways of our brain, allowing the information to transmit between various brain areas. With the assistance of DIT data to reconstruct the white matter pathways, it demonstrates the links between structural connectivity and resting-state functional networks [17]. Many existing statistical methods have been developed to analyze fMRI data [18-21] and DTI data [22-25] within a Bayesian framework, separately. For example, a spatial and temporal independent component analysis (ICA) together with Bayesian approximation was applied to process large scale resting-state fMRI data from 200 subjects [21]. A multi-tensor Bayesian model with a new parameterization method was developed for DTI data from a healthy subject, allowing to be suitable for model selection in the post analysis via thresholding the Bayes factor [25].

Some efforts have been made to develop statistical models to combine fMRI data and DTI data jointly. A combined analysis of DTI and fMRI data was conducted to explore whether there were networks of regions where maturation of white matter and changes in brain activity showed similar developmental trends during childhood [26]. In this fMRI data analysis, functional anisotropy, as an indicator of myelination and axon thickness, was used as a covariate in a multiple regression model to find brain regions where functional anisotropy values and blood oxygen level-dependent (BOLD) response were correlated [26].

Another statistical method implementing a hierarchical clustering algorithm, combined various sources of data including anatomically weighted functional connectivity (awFC), fMRI and DTI data, to determine the functional connectivity [27]. What's more, DTI data was utilized as a supplement of fMRI information, to estimate functional connectivity in a multimodal approach. This Bayesian model can determine the hierarchy among functional connected pairs of brain regions based on the associated probabilities of elevated activity for each node [28]. Other studies also demonstrated the superior advantages of using both fMRI and DTI data to investigate the functional connectivity in brain networks [29-33].

2.2 Spatiotemporal Hierarchical Model

We developed a novel Bayesian hierarchical modeling framework using resting-state fMRI and DTI data to improve the precision and accuracy of estimation of functional connectivity [34]. To mention, we defined a term called "naïve FC" as the correlation between two average time-series across voxels within each region of interest (ROI) without considering the temporal correlation. In general, our approach not only applied the intrinsic spatial and temporal correlation in resting-state fMRI data, but also considered the weighted average of structural connectivity from DTI data and the naïve FC as a source of prior information for functional connectivity. Furthermore, two sources of structural connectivity, direct information and indirect information, were utilized in a mixture model to estimate the functional connectivity. Due to the natural incorporation of functional and structural information from resting-state fMRI and DTI data, we can improve the estimation accuracy and lead to more reliable inference.

2.2.1 Spatiotemporal Structure

In a resting-state fMRI study, we define the time-series data at voxel v in ROI c as

$Y_{cv}(t)$, where $t = 1, \ldots, T$. In the same ROI c, a spatiotemporal model for the resting-state fMRI time-series can be expressed as the following:

$$Y_{cv}(t) = \beta_c + b_c(v) + d_c + \epsilon_{cv}(t).$$

In the formula above, β_c is the grand mean in the ROI c. $b_c(v)$ represents a zero-mean voxel-specific random effect in the ROI c and captures the local spatial dependency between voxels. A kernel function $K_c(\cdot)$ is defined as the covariance structure for local spatial covariance. It is a function of Euclidean distance:

$$Cov\big(b_c(v), b_c(v')\big) = K_c(\|v - v'\|).$$

Note that the voxel-specific random effect b values are uncorrelated when two voxels are in different ROIs $(c \neq c')$, which means the expression below:

$$Cov\big(b_c(v), b_{c'}(v')\big) = 0 \text{ if } c \neq c'.$$

This kernel function can be any valid spatial covariance function. Table 2.1 lists the common kernel functions for the covariance structure [35]. In our model, we apply the exponential function to represent the covariance structure between voxels in ROI c:

$$Cov\big(b_c(v), b_c(v')\big) = \sigma_{b_c}{}^2 \exp(-\|v - v'\|\varphi_c),$$

where $\sigma_{b_c}{}^2$ is defined as the spatial variance at each voxel in the ROI c and $\|v - v'\|$ denotes the Euclidean distance between two voxels, v and v'. φ_c represents the inversed ROI-specific decaying parameter in the exponential structure.

d_c is a zero-mean ROI-specific random effect. Its covariance structure is used to model functional connectivity and expressed as $Cov\big(d_c(v), d_{c'}(v')\big)$. We will explain how this effect results from naïve FC and DTI data with a series of prior information.

Table 2.1 Various common covariance functions from Ref [35].

Constant	$K(x, x') = c$
Linear	$K(x, x') = x^T x'$
Gaussian noise	$K(x, x') = \sigma^2 \delta_{x,x'}$
Squared exponential	$K(x, x') = \exp\left(-\dfrac{\|x - x'\|^2}{2l^2}\right)$
Exponential	$K(x, x') = \exp\left(-\dfrac{|x - x'|}{l}\right)$
Matérn	$K(x, x') = \dfrac{2^{1-v}}{\Gamma(v)} \left(\dfrac{\sqrt{2v}|x - x'|}{l}\right)^v B_v\left(\dfrac{\sqrt{2v}|x - x'|}{l}\right)$
Periodic	$K(x, x') = \exp\left(-\dfrac{2\sin^2\left(\dfrac{x - x'}{2}\right)}{l^2}\right)$
Rational quadratic	$K(x, x') = (1 + |x - x'|^2)^{-\alpha}, \alpha \geq 0$

Finally, $\epsilon_{cv}(t)$ is the noise part. We assume this voxel-specific noise follows an autoregressive (AR) temporal process with order one, that is AR (1). So, the expression of the noise follows:

$$\epsilon_{cv}(t) = \delta_c + \phi_{cv} \, \epsilon_{cv}(t - 1) + w(t),$$

where δ_c is the constant shift, ϕ_{cv} is the AR (1) coefficient with a requirement of $|\phi_{cv}| < 1$. And $w(t)$ is Gaussian random noise with a distribution as $N(0, \sigma_{cv}^2)$ and is independent of $\epsilon_{cv}(t)$. It is straightforward to calculate the mean and variance of $\epsilon_{cv}(t)$ as the following:

$$E[\epsilon_{cv}(t)] = \frac{\delta_c}{1 - \phi_{cv}}$$

$$Var[\epsilon_{cv}(t)] = \frac{\sigma_{cv}^2}{1 - \phi_{cv}^2}$$

2.2.2 Hierarchical Structure

Our goal is to estimate each functional connectivity through its corresponding posterior distribution. To obtain the posterior distribution, each component in the spatiotemporal structure from last section can be rewritten as a hierarchical structure in one ROI level:

$$Y_c(t) = \beta_c + b_c + d_c + \epsilon_c(t).$$

$Y_c(t)$ denotes a vector ($1 \times V$) of signals at each voxel as $[Y_{c1}(t), Y_{c2}(t), \ldots, Y_{cV}(t)]^T$. We use J and I to indicate the all-one vector and identity matrix, respectively. Therefore, each component can be vectorized as:

$$\beta_c = \beta_c J_{(1 \times V)}$$

$$b_c = [b_{c1}, b_{c2}, \ldots, b_{cV}]^T$$

$$d_c = d_c J_{(1 \times V)}$$

$$\epsilon_c(t) = [\epsilon_{c1}(t), \epsilon_{c2}(t), \ldots, \epsilon_{cV}(t)]^T.$$

And the hierarchical structure follows:

$$\beta_c \sim N\left(0, \sigma_{\beta_c}^2\right)$$

$$b_c \sim N\left(0, \Sigma_{b_c}\right)$$

$$d_c \sim N(0, \Sigma_d)$$

$$\epsilon_{cv}(t) \sim N\left(\frac{\delta_c}{1 - \phi_{cv}}, \frac{\sigma_{cv}^2}{1 - \phi_{cv}^2}\right)$$

In details, each term β_c follows a Gaussian distribution with mean zero and variance $\sigma_{\beta_c}^2$. In addition, for different ROIs ($c \neq c'$), β_c is independent of $\beta_{c'}$. For the term b_c, it follows a Gaussian distribution with the covariance Σ_{b_c}, which applies the distant-dependent exponential function. For the term d_c, we assume it to follow a Gaussian distribution as

$N(0, \Sigma_d)$. Here the correlation matrix of d_c represents the functional connectivity among all ROIs, which is considered as the most important parameter estimated in the whole Bayesian framework. For the noise part with AR (1) time-series structure, it follows a voxel-specific Gaussian distribution as $N(\frac{\delta_c}{1-\phi_{cv}}, \frac{\sigma_{cv}^2}{1-\phi_{cv}^2})$.

2.2.3 Double Fusion

Σ_d, as the covariance matrix of d_c, is obtained through a novel method considering both structural and functional information. In other words, the prior distribution of the correlation matrix can be established from the structural and naïve functional connectivity of each pair of ROIs in two steps and we name this method as "a double fusion model".
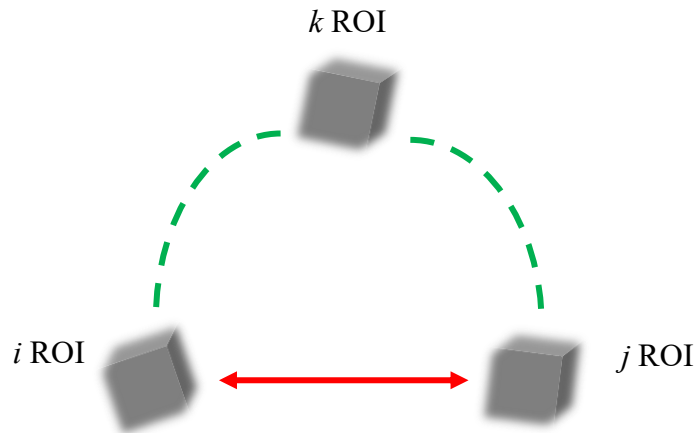
Figure 2.1 Schematic of a direct structural connection between i and j ROIs and a possible indirect structural connection via k ROI.

We combine the structural connectivity and naïve functional connectivity together because the effect of direct structural connectivity is different from that of indirect structural connectivity. For example, relatively lower values of structural connectivity imply no direct

correlated pathways between two ROIs. However, it is likely that there exist indirect structural connections between two ROIs, resulting in high functional coupling (Figure 2.1). And the low structural connectivity will indicate very low functional connectivity if no structural connection exists. Therefore, we should treat the indirect structural connectivity differently from direct structural connectivity in the fusion step.

Then the prior distribution of the covariance matrix Σ_d is further considered to be a function of structural and naïve functional connectivity matrix. With the Cholesky decomposition, we name L_{sc}, L_{nfc} and L_d as the lower triangular matrix from the structural covariance matrix, naïve functional covariance matrix and functional covariance matrix. To identify the different effects from direct structural connectivity and indirect structural connectivity, we denote $L_d(direct)$ and $L_d(indirect)$ as the lower triangular matrix from direct and indirect structural information, separately. To regulate each source of information, we assume a weighted combination to represent $L_d(direct)$, $L_d(indirect)$ and L_d:

$$L_d(direct) = \lambda L_{sc} + (1 - \lambda)L_{nfc}$$

$$L_d(indirect) = M_{sc}\lambda L_{sc} + (1 - M_{sc}\lambda)L_{nfc}$$

$$L_d = \omega L_d(direct) + (1 - \omega)L_d(indirect)$$

where λ and ω are weight parameters. M_{sc} is the measurement of structural connectivity.

Finally, Σ_d is reconstructed as $L_d \times L_d{}^T$ and the corresponding correlation matrix ρ_d can be obtained to denote the resting-state functional connectivity through a normalization step. It is also important to mention that the estimated Σ_d and ρ_d are demonstrated to be positive semidefinite due to the Cholesky decomposition and reconstruction. For a correlation matrix within n ROIs, it follows:

$$\rho_d = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ & 1 & \ddots & \vdots \\ & & 1 & \rho_{(n-1)n} \\ & & & 1 \end{pmatrix}_{n \times n}.$$

The elements in the upper triangular part can be vectorized as:

$$\left[\rho_{12}, \ldots, \rho_{1n}, \rho_{23}, \ldots, \rho_{2n}, \ldots, \rho_{(n-1)n}\right]_{n\_vec}.$$

The total number of estimation is $n\_vec = (n-1) + (n-2) + \cdots + 1 = (n-1)\,n/2$.

2.2.4 Prior Distribution

Because we have no prior information about the values of each parameter, we decide to apply uninformative priors. In the exponential function $Cov\big(b_c(v), b_c(v')\big) = \sigma_{b_c}{}^2 \exp(-\|v - v'\|\varphi_c)$, we assume the corresponding parameters follow:

$$\varphi_c \sim \text{Unif}(0, 20)$$

$$\sigma_{b_c} \sim \text{Unif}\,(0, 100).$$

In the temporal correlation $\epsilon_{cv}(t) = \delta_c + \phi_{cv}\,\epsilon_{cv}(t-1) + w(t)$, we assume the prior distribution of each parameter as:

$$\phi_{cv} \sim \text{Unif}(0, 1)$$

$$\sigma_{cv} \sim \text{Unif}(0, 100).$$

And the grand mean β_c:

$$\beta_c \sim \text{N}(0, 100^2).$$

Two weight parameters λ and ω in the double fusion model:

$$\lambda \sim \text{Beta}(1, 1)$$

$$\omega \sim \text{Beta}(1, 1).$$

Also, the covariance matrix for functional connectivity and structural connectivity is constructed via a prior diagonal matrix. The diagonal element is generated from a function

of a σ_{d_c} parameter in a logarithmic scale:

$$\sigma_{d_c} \sim \text{Unif}(-8, 8).$$

Finally, with adding all the separate components, we assume the observed values Y_{obs} follow a Gaussian distribution $Y_{obs} \sim \text{N}(Y_{cv}, \sigma^2)$ under the Bayesian framework and σ has a prior distribution:

$$\sigma \sim \text{Unif}(0, 100).$$

2.3 Introduction to PyMC3 and NUTS

Probabilistic programming is designed for flexible specification and fitting of Bayesian statistical model. PyMC3 is new, open-source framework with a readable but powerful syntax close to the natural syntax statisticians will use to describe models [36]. It includes the new-generation Markov Chain Monte Carlo (MCMC) sampling algorithms as the No-U-Turn Sampler (NUTS) [37], which avoids the random walk behavior and sensitivity to correlated parameters by taking a series of steps informed by first-order gradient information. In other words, the NUTS method is a self-tuning variant of Hamiltonian Monte Carlo (HMC) [38].

Many simulations have demonstrated that the new-generation sampler is good for high dimensional and complex posterior distributions such as the spatiotemporal hierarchical model we built. Because HMC and NUTS apply the gradient information from the likelihood, they can achieve much faster convergence than traditional sampling methods. Taking our spatiotemporal hierarchical model as an example, to ensure the convergence of model parameters, we need to set the sampling size as 300000 in the Metropolis-Hastings algorithm, which is another MCMC method. The Metropolis-Hastings method also requires a long burn-in period, where an initial number of samplers are thrown away. In our model,

the burn-in number of samplers is usually set to one third of the sampling size as 100000. However, with the adoption of NUTS in the model, we only need to set the sampling size as 1000 and the burn-in (or tuning) number as 1000. The NUTs also have a few self-tuning strategies for adaptively setting the parameters of HMC and it allows many complex models to be fit without specialized knowledge about fitting algorithms [36]. Therefore, the work of our spatiotemporal model is mainly maintained by PyMC3.

It is also important to mention that PyMC3 take the advantage of Theano [39, 40] as backend to transparently transcode models to C and compile them to machine code. So, it can boost the performance of sampling procedure by taking the advantage of graphical processing units (GPU) architectures. Theano is a numerical computation library for Python, which allows expressions to be like NumPy syntax. Here I illustrate a simple example of linear regression in PyMC3.

The model includes a predicting outcome Y with normal-distributed observations with mean μ and variance σ^2. The expected value μ is a linear combination of two predictor variables, X_1 and X_2:

$$Y \sim \mathrm{N}(\mu, \sigma^2)$$

$$\mu = \alpha + \beta_1 X_1 + \beta_2 X_2.$$

Each parameter ($\alpha, \beta_1, \beta_2, \sigma$) corresponds to the following prior distribution:

$$\alpha \sim \mathrm{N}(0, 100)$$

$$\beta_1 \text{ or } \beta_2 \sim \mathrm{N}(0, 20)$$

$$\sigma \sim \mathrm{HalfNormal}(0, 1).$$

16

This simple model of linear regression is specified in PyMC3 as the following:

```python
import pymc3 as pm
with pm.Model() as basic_model:

    # Priors for unknown model parameters
    alpha = pm.Normal('alpha', mu=0, sd=100)
    beta = pm.Normal('beta', mu=0, sd=20, shape=2)
    sigma = pm.HalfNormal('sigma', sd=1)

    # Expected value of outcome
    mu = alpha + beta[0]*X1 + beta[1]*X2

    # Likelihood (sampling distribution) of observations
    Y_obs = pm.Normal('Y_obs', mu=mu, sd=sigma, observed=Y)
```

And the tuning step goes through with NUTS algorithm in 1000 draws from the posterior:

```python
with basic_model:

    # instatiate sampler
    step = pm.NUTS()

    # draw 1000 posterior samples and tune 500 as default
    trace = pm.sample(1000, step = step)
```

When the sampling is complete, the posterior analysis can be inspected through trace plot of each parameters and various diagnostics such as Geweke statistics [41] and Gelman Rubin statistics [42, 43]. The Geweke score on some series x is computed by:

$$\frac{E[x_s] - E[x_e]}{\sqrt{V[x_s] + V[x_e]}}$$

where E stands for the mean, V the variance, x_s a section at the start of the series and x_e a section at the end of the series. And the Gelman Rubin diagnostic is computed by:

$$\hat{R} = \frac{\hat{V}}{W}$$

where W is the within-chain variance and $\hat{V}$ is the posterior variance estimate for the pooled traces. This is the potential scale reduction factor, which converges to unity when each of the traces is a sample from the target posterior. Values greater than one indicate that one or more chains have not yet converged. In practice, we look for values of $\hat{R}$ close to one (say, less than 1.1) to be confident that a particular estimate has converged.

What's more, the effective sample size is computed to check the model diagnostics:

$$\hat{n}_{eff} = \frac{mn}{1 + 2\sum_{t=1}^{T}\hat{\rho}_t}$$

Where m is the number of chains, n is the sample draws, $\hat{\rho}_t$ is the estimated autocorrelation at lag t, and T is the first odd positive integer for which the sum $\hat{\rho}_{T+1} + \hat{\rho}_{T+2}$ is negative [44].

2.4 Optimization and Decomposition

To reduce the computation cost in this spatiotemporal hierarchical model using PyMC3, we avoid defining the prior distribution of parameters by each ROI. Instead, we optimize all the parameters through vectorization. For example, the normal distribution of the noise in the voxel-specific AR (1) temporal correlation is defined through a mean vector $(n \times 1)$ and a diagonal matrix $(n \times n)$ as covariance for n ROIs. Another advantage is to define the module without using the hard coding style. This example can be referred in the APPENDIX.

In addition, Cholesky decomposition is utilized in our MCMC method to generate

correlated random variables. For a positive definite covariance matrix Σ, Cholesky decomposition expresses Σ as $U^T U$, where U is a unique upper-triangular matrix with positive diagonal entries. For example, to generate correlated random variables that follow a n dimension multivariate normal distribution $X \sim N(\mu, \Sigma)$ with a mean vector μ ($n \times 1$) and a covariance matrix Σ (a $n \times n$ positive definite matrix), we can decompose the covariance matrix Σ into $U^T U$ and generate a vector as Z with n independent $N(0,1)$ random variables. Therefore, X can be generated as:

$$X = \mu + U^T Z$$

We illustrate this technique by generating the exponential covariance structure between two voxels. What's more, we can generate a matrix ($m \times n$) from the univariate random variable $Z \sim N(0,1)$ to speed up the sampling procedure. Here m is the number of voxels in each ROI. This optimization greatly helps reduce the computation cost from ~30 hours to ~20 hours for modeling each subject.

2.5 Simulation Study

2.5.1 Data Generation

We generated time-series data with a length of $T = 128$ scans using AR (1) at 5 ROIs and each ROI contains 100 voxels. The AR (1) temporal correlation coefficient is assume as 0.6. Then we imposed correlation between each ROI using a multivariate normal distribution with zero mean and the correlation matrix as the following:

$$\rho_d = \begin{pmatrix} 1 & 0.6 & 0 & 0.5 & 0 \\ & 1 & 0.2 & 0.1 & 0 \\ & & 1 & 0 & 0.1 \\ & & & 1 & 0.2 \\ & & & & 1 \end{pmatrix}$$

or the vector $[0.6, 0, 0.5, 0, 0.2, 0.1, 0, 0, 0.1, 0.2]$ as in the upper triangular part. Moreover,

we assume that the structural connectivity is the same as the above correlation matrix, but the functional connectivity is sampled from a Wishart distribution with mean covariance matrix ρ_d as the above matrix and six degrees of freedom:

$$SC \sim W_p(6, \rho_d).$$

2.5.2 Estimation

Table 2.2 Median (standard deviation), credible intervals (2.5% and 97.5% quantiles) of estimated functional connectivity, the Gelman Rubin diagnostic and the effect sample size under the condition of true structural connectivity and structural independence.

FC	Correct	Bayesian correct SC				Bayesian independence			
		Median (SD)	[2.5% 97.5%]	$\hat{R}$	$\hat{n}_{eff}$	Median (SD)	[2.5% 97.5%]	$\hat{R}$	$\hat{n}_{eff}$
ρ_1	0.6	0.645 (0.014)	[0.617 0.669]	0.999	1675.438	0.531 (0.125)	[0.170 0.652]	1.002	979.951
ρ_2	0.0	0.310 (0.080)	[0.074 0.378]	0.999	1110.936	0.305 (0.078)	[0.091 0.379]	0.999	1303.385
ρ_3	0.5	0.546 (0.017)	[0.513 0.576]	0.999	1877.821	0.463 (0.108)	[0.151 0.572]	0.999	1033.009
ρ_4	0.0	0.145 (0.042)	[0.039 0.204]	0.999	1341.864	0.146 (0.042)	[0.040 0.204]	0.999	1034.929
ρ_5	0.2	0.478 (0.053)	[0.340 0.547]	0.999	1152.252	0.432 (0.090)	[0.197 0.537]	1.001	1285.429
ρ_6	0.1	0.123 (0.020)	[0.088 0.165]	1.000	1721.880	0.017 (0.106)	[-0.194 0.216]	1.000	1218.262
ρ_7	0.0	0.399 (0.083)	[0.180 0.522]	0.999	1102.607	0.419 (0.098)	[0.174 0.558]	0.999	1381.804
ρ_8	0.0	0.173 (0.049)	[0.047 0.235]	0.999	1293.770	0.158 (0.068)	[0.016 0.283]	1.000	1251.840
ρ_9	0.1	0.105 (0.070)	[-0.026 0.254]	0.999	1417.283	0.057 (0.063)	[-0.066 0.177]	1.000	1458.304
ρ_{10}	0.2	0.316 (0.077)	[0.165 0.484]	1.000	1524.805	0.348 (0.118)	[0.054 0.532]	0.999	1095.445

The data were analyzed with two different priors for functional connectivity: one was the informative prior based on true correlation matrix ρ_d and the other prior was based on the identity matrix $I_{5\times5}$, which assumed no structural dependence. The posterior of functional connectivity was obtained via our Bayesian model using NUTS sampler implemented in PyMC3. We draw 500 posterior samples with 500 tuning steps. Table 2.2 lists the estimated functional connectivity and some

important statistics of the model diagnostics. We define three different strength of functional connectivity as zero, low (0.1, 0.2) and strong (0.5 or 0.6). The mean squared errors (MSEs) are very close to two conditions. For zero functional connectivity, two MSEs are 0.277 and 0.281; for low functional connectivity, two MSEs are 0.151 and 0.145; for high functional connectivity, two MSEs are 0.045 and 0.055. This result based on the NUTS algorithm in PyMC3 is quite different from that from previous report using Metropolis-Hastings algorithm in PyMC2 [34], implying the superior advantage of NUTS algorithm even though the Bayesian independence assumption might be incorrect.

To explore more about the estimated functional connectivity, we plot the histograms of 10 parameters from their posterior distributions in Figure 2.2. We can clearly observe smaller variances in the correct Bayesian structural connectivity than those in the Bayesian independence assumption. Furthermore, to ensure the convergence with the right number of sample draws using NUTS algorithms, we conduct an experiment of 3000 posterior samples under the correct correlation matrix in PyMC3. The histogram of each functional connectivity is shown in Figure 2.3. The posterior distributions of each parameter under 1000, 2000 and 3000 sample draws are almost identical. In practice, the typical sample draw using NUTS is 500 or 1000, which is more advantageous than the Metropolis-Hastings algorithm (usually more than 100000).
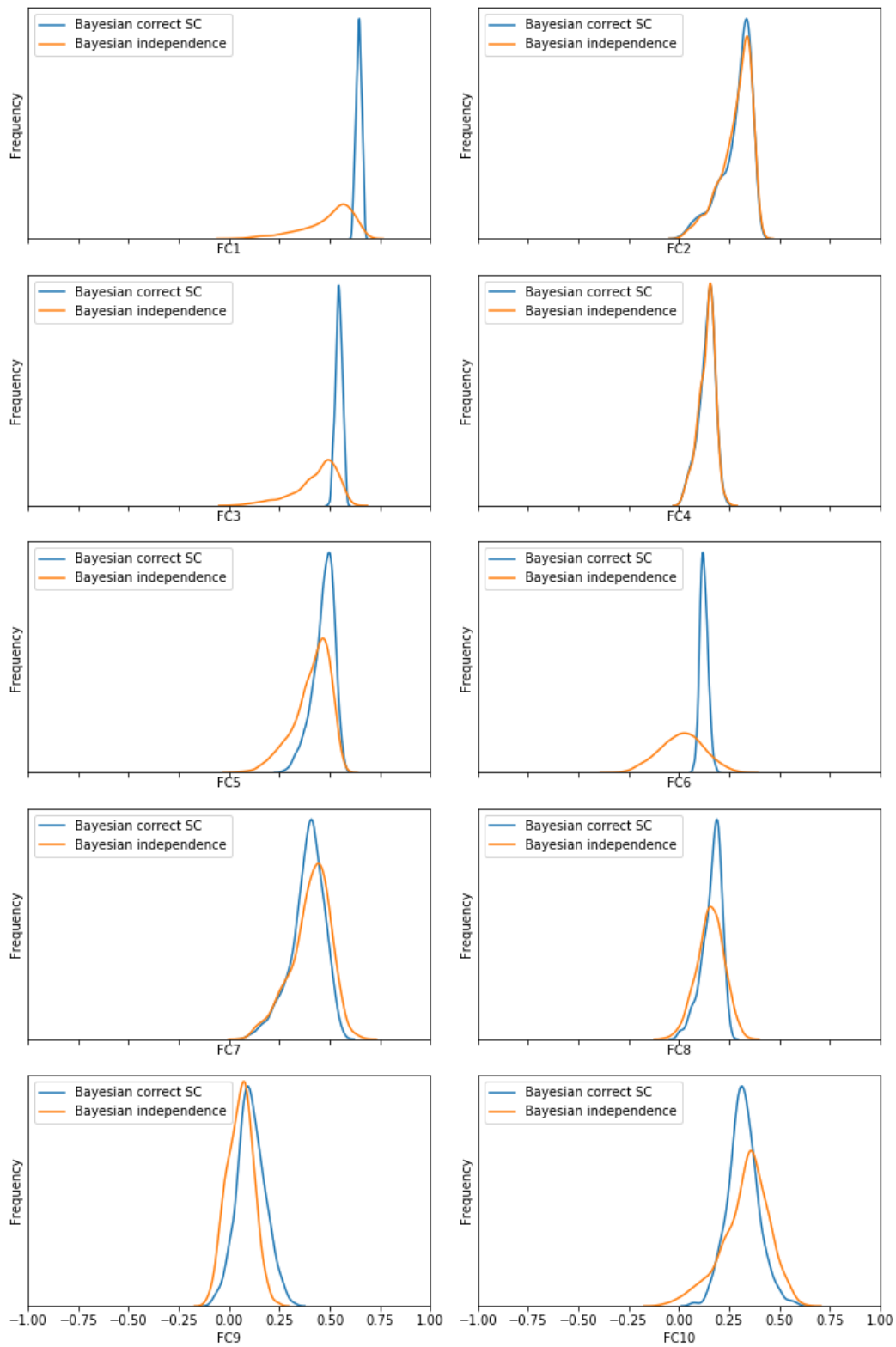
Figure 2.2 Histogram plots of 10 parameters from their posterior distribution.
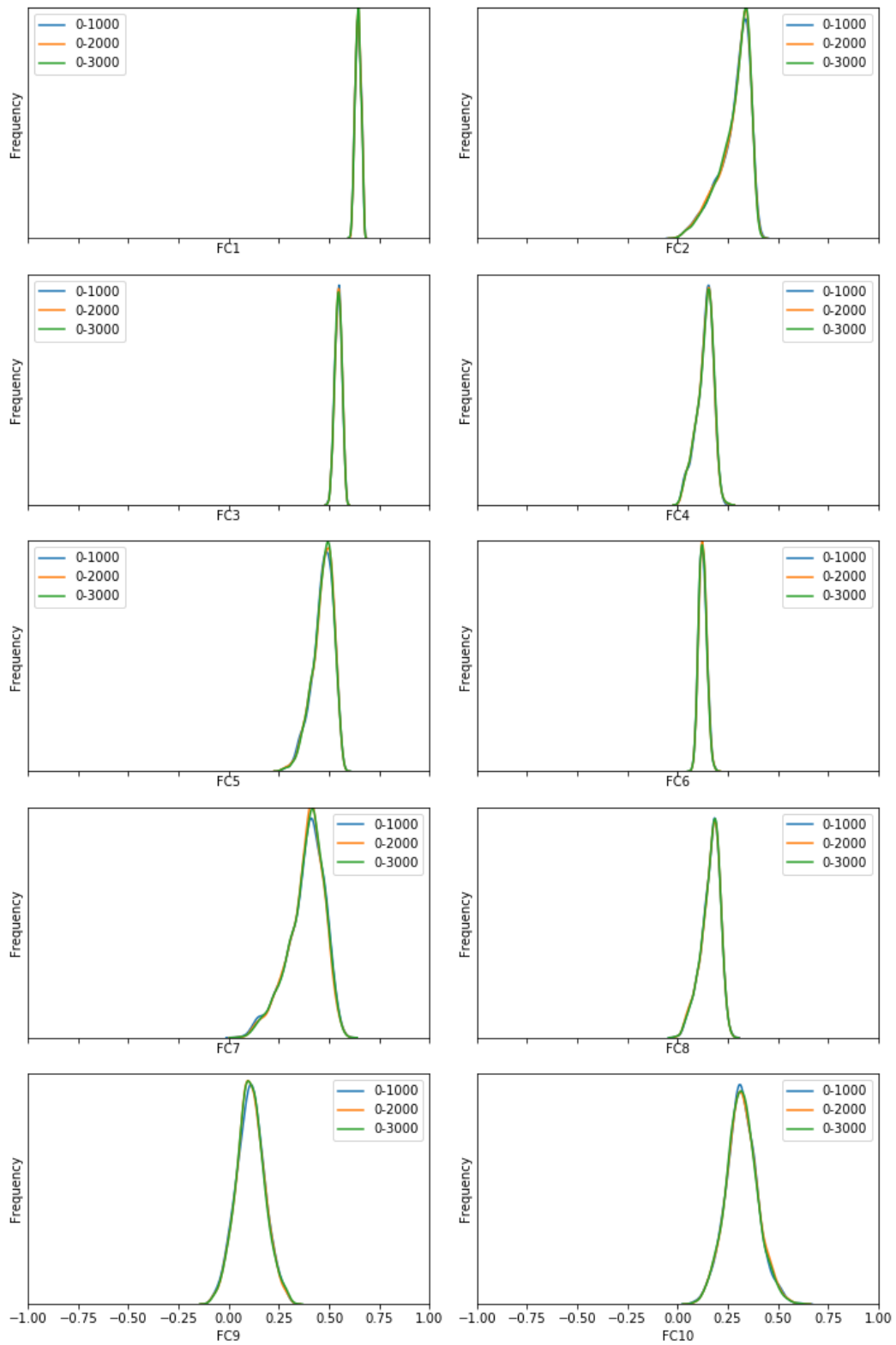
Figure 2.3 Histogram plots of 10 parameters with 1000, 2000 and 3000 sample draws.

2.6 Case Study

2.6.1 Background

Major depressive disorder (MDD), also known as depression, is a mental disorder characterized by a period of low mood, low self-esteem, loss of interest, low energy or pain [45]. Although MDD is diagnosed based on behaviors, deficient signs in cognitive performance are also common. Changes in the connectivity and function of the brain networks are likely to affect the emotional processes directly related to depressive symptoms but also negatively affect cognitive function. Recent studies have demonstrated that intrinsic networks exhibit altered connectivity between brain regions associated with emotional processes and cognitive function [46]. In this case study, we are interested to establish the relationship between altered resting-state fMRI connectivity and cognitive function in depressed individuals, which can facilitate understanding the role of network connectivity in MDD.

Forty-one subjects were enrolled and completed resting-state fMRI. In the current analysis, we have 23 subjects that are non-depressed and 18 subjects that are in MDD group. From Table 2.3, there were no significant difference between the control and MDD groups in the covariates as age, sex or education. The MDD group had significantly higher Montgomery–Asberg Depression Rating Scale (MADRS) and Beck Depression Inventory (BDI) scores as expected.

Table 2.3 Descriptive statistics of 41 subjects including control (n=23) and MDD (n=18) groups.

	Control (n=23)	MDD (n=18)	Wilcoxon Rank Sum Tests
	Mean (SD)	Mean (SD)	
Age(years)	31.78 (10.16)	32.06 (8.55)	t = -0.512, p = 0.608
Sex (% female)	65%	50%	t = 0.828, p = 0.408
Education (years)	15.78 (1.73)	16.28 (1.90)	t = -0.512, p = 0.608
Beck Depression Inventory (BDI)	1.90 (2.62)	22.11 (9.38)	t = -4.085, p <0.001
Montgomery–Asberg Depression Rating Scale (MADRS)	0.70 (1.06)	25.29 (3.20)	t = -5.438, p < 0.001
Processing Speed Domain	0.36 (0.67)	0.20 (0.61)	t = 0.841, p = 0.401
Working Memory Domain	0.10 (0.88)	0.02 (0.81)	t = 0.158, p = 0.875
Episodic Memory Domain	0.23 (0.55)	0.07 (0.75)	t = 0.578, p = 0.563
Executive Function Domain	0.20 (0.55)	0.23 (0.59)	t = -0.053, p = 0.958

2.6.2 Exploratory Analysis

Based on fMRI and DTI data, we estimated the functional connectivity between 14 ROIs, which are main regions in the default mode network, for each subject through the established Bayesian spatiotemporal model using PyMC3 and NUTS. Each ROI contains 300 voxels and the length of time-series per voxel is 150. Because the large dimension leads to more parameters in the MCMC process, the computation costs for each subject range from 16 hours to 30 hours. The sampling draw is 1000 with 1000 tuning step. We also conducted sampling for 3000 draws on three subjects and check the trace plot between 0-1000, 0-2000 and 0-3000, which is consistent with the convergence illustrated in Figure 2.3. The plots of 3000 draws on one subject are shown in the APPENDIX.

The number of the estimated functional connectivity on each subject is ninety-one and each is named with prefix "FC" such as "FC24". The correlation among age, sex, education, depress (control or MDD) and functional connectivity is visualized in Figure 2.4

based on the hierarchical clustering order.

Figure 2.4 Correlation plot of covariates and estimated functional connectivity.

Furthermore, we conducted the correlation tests between the estimated functional connectivity and multiple cognitive domains including episodic memory, executive function, processing speed and working memory. Here p values from the tests were adjusted to control the false discovery rate (FDR), which is the expected proportion of false discoveries amongst

the rejected hypotheses. Across all the ninety-one functional connectivity, we did not find any adjusted p value below 0.1, implying that there is no significant evidence that each of functional connectivity is correlated with any of the cognitive domains at FDR = 0.1. Later, similar tests were conducted on the control and MDD groups, separately. In the MDD group, there is no adjusted p value below 0.1, indicating that we don't see correlation between the functional connectivity and cognitive domains at FDR = 0.1. However, in the control group, we find that "FC80" is correlated with executive function domain after we adjusted the correlation tests using the FDR method at 0.1.

2.6.3 Regression Analysis

In order to explore the relationship between MDD and multiple cognitive domains, we applied the following regression models to check how the interaction between the estimated functional connectivity and depression:

$$\text{Cognitive domain} \sim \text{Age} + \text{Sex} + \text{Education} + \text{FC}_i + \text{Depress} + \text{FC}_i * \text{Depress}$$

where FC_i is the i th estimated functional connectivity of the upper triangular elements in the correlation matrix. The regression coefficients can be referred in the APPENDIX. For one cognitive domain regressed with each of ninety-one functional connectivity measures, p values obtained from the interaction term were adjusted using the FDR method. For processing speed domain, working memory domain and episodic memory domain, we observed that no adjusted p value was below 0.1, indicating that there was no significant interaction between functional connectivity and depression status. However, for executive function domain, we found that the adjusted p value for the interaction term between "FC80" and depress was below 0.1. In other words, there exists significant evidence that "FC80" were associated with depression. The coefficient of that interaction is 1.9, meaning higher association between executive function domain and "FC80" for subjects who are suffering from MDD

27

than that for subjects in control while all the other conditions hold the same.

What's more, we examined which of the estimated functional connectivity could affect each cognitive function through several methods of variable selection. First, we employed four different subset selection approaches, which are exhaustive, forward, backward and sequential methods. Due to the large number of explanatory variables including four covariates (age, sex, education and depress), we decide to force the four covariates always in the model setting and then choose six best variables from "FC1" to "FC91". The plots of residual sum of squares and adjusted R^2 for each search algorithm are illustrated in APPENDIX.

Another comparable method of variable selection is to apply LASSO (least absolute shrinkage and selection operator) method. It utilizes the l_1 penalty for both fitting and penalization of the coefficients. The model is also validated through the bootstrapping technique with 1000 iterations. We calculated the percentage of each functional connectivity that was selected in each bootstrapping sample. The four covariates were also not allowed to be dropped out. The results of percentages can be referred in the APPENDIX.

After we obtained the variable selection results from all the methods, we listed six of the selected variables in Table 2.4. For processing speed domain, "FC10" and "FC29" were selected frequently in all the five methods. For working memory domain, "FC26" was selected five times, which warrants more research on the network connectivity between the corresponding brain regions. "FC69" is the second frequent of all the selected variables. For episodic memory domain, "F9" is the most frequent variable with four times. For executive function domain, "FC26" are included in five methods, the same frequency for working memory domain. All the selected functional connectivity could be of future research interests

to explore the relationship between brain connectivity – cognition association and depression.

Table 2.4 Variable selection results from each method.

	Processing Speed Domain	*Working Memory Domain*	*Episodic Memory Domain*	*Executive Function Domain*
Exhaustive	FC4, FC27, FC28, FC48, FC57, FC69	FC26, FC29, FC62, FC64, FC69, FC71	FC6, FC9, FC42, FC57, FC78, FC79	FC6, FC7, FC26, FC27, FC50. FC79
Forward	FC10, FC20, FC29, FC44, FC51, FC58	FC20, FC26, FC33, FC62, FC69, FC85	FC6, FC9, FC11, FC26, FC35, FC65	FC18, FC20, FC26, FC43, FC50, FC77
Backward	FC2, FC4, FC6, FC12, FC20, FC24	FC7, FC18, FC19, FC25, FC26, FC33	FC7, FC9, FC11, FC22, FC29, FC30	FC3, FC6, FC8, FC11, FC13, FC26
Sequential	FC10, FC11, FC17, FC28, FC29, FC78	FC26, FC29, FC62, FC64, FC69, FC71	FC1, FC5, FC6, FC9, FC11, FC35	FC18, FC20, FC26, FC43, FC50, FC77
Lasso	FC10, FC11, FC26, FC29, FC70, FC85	FC20, FC26, FC62, FC64, FC69, FC84	FC6, FC25, FC34, FC57, FC77, FC86	FC11, FC26, FC58, FC70, FC77, FC85

CHAPTER 3

SUMMARY AND FUTURE WORK

In this thesis, we developed a Bayesian hierarchical spatiotemporal model that combined fMRI and DTI data jointly to enhance the estimation of resting-state functional connectivity. Structural connectivity from DTI data was utilized to construct an informative prior for functional connectivity based on resting-state fMRI data through the Cholesky decomposition in a mixture model. The analysis took the advantages of probabilistic programming package as PyMC3 and next-generation Markov Chain Monte Carlo (MCMC) sampling algorithm as No-U-Turn Sampler (NUTS). The simulation study with this advanced algorithm, illustrated reduced mean squared errors (MSEs) of estimation. Furthermore, through a case study of MDD, we applied our model to examine how the estimated functional connectivity was associated with tasks of episodic memory, executive function, processing speed and working memory. Through five various methods of variable selection, certain estimated functional connectivity can be extended to investigate the correlated links between ROIs and depression in the future.

We have made significant progress in the development of estimating functional connectivity using PyMC3. However, abundant exciting opportunities are still available to further advance the capabilities and applications of our model. For example, we have listed a set of covariance function to capture local spatial correlation, which can be extended to compare the estimated results. Another explorable work is to combine all the variables including the covariates, estimated functional connectivity and multiple cognitive domains, and to build a classifier to identify whether this subject is under MDD or not. The case study in this thesis only have forty-one subjects, which is likely to be overfitted using the modern

machine learning techniques. For 200 ~ 300 subjects, multiple methods such as random forest, gradient boosting tree and support vector machine, can be applied to predict the classification of subjects or the value of cognitive domains.

In conclusion, we have proposed a Bayesian model for resting-state brain networks using the newly-developed probabilistic programming even though many challenges still exist. Hopefully, the work presented in this thesis will bring attention to future development of brain imaging.

APPENDIX

A.1 Setting of PyMC3 and Theano

The latest version of *pymc3* package can be installed from PyPI using *pip*:

```
pip install pymc3
```

Or via *conda-forge* channel if you have installed Anaconda to manage installations of various packages:

```
conda install -c conda-forge pymc3
```

To mention, other related packages like *numpy*, *pandas*, *theano* will be installed together with *pymc3*. In addition, to ensure running *pymc3* and *theano* in various OS systems such as Unix, Mac and Windows, other packages such as *mkl-service* and *m2w64-toolchain* need to be installed with the underlined notices during the setup. For Windows users, environmental variables need to be added in the system setting to use *conda* command in the terminal and *theano* package in Python environment. C++ compiler such as Cygwin needs to be installed before running *pymc3* and *theano* in Windows if it does not exist.

To use GPUs for intensive parallel computation purposes in *theano* on either Vanderbilt ACCRE or your own devices, a *theanorc* file is required and can be added based on *theano* GPU documentation or ACCRE python documentation in the GitHub pages. All the setting and examples are also illustrated in my GitHub repository named DoubleFusion (https://github.com/wangruinju/Double-Fusion).

A.2 Spatiotemporal Hierarchical Model in PyMC3

```python
import numpy as np
import pymc3 as pm
import theano.tensor as tt
import theano
import csv
import os
from datetime import date

def get_data(name):
    yreader = csv.reader(open(name + ".csv"))
    Y = np.array([row for row in yreader]).astype(float)
    return Y

def get_func(name, n):
    sreader = csv.reader(open(name + ".csv"))
    mFunc = np.array([row for row in sreader]).astype(float)
    func_new = np.array(mFunc[0, 0:n*(n-1)//2])
    func_temp = np.triu(np.ones([n, n]),1)
    func_temp[func_temp==1] = func_new
    Func_mat = func_temp.T + np.eye(n) + func_temp
    return Func_mat

def get_struct(name, n):
    sreader = csv.reader(open(name + ".csv"))
    S_read = np.array([row for row in sreader]).astype(float)
    struct_new = np.array(S_read[0:n*(n-1)//2, 0])
    Struct_temp = np.triu(np.ones([n, n]), 1)
    Struct_temp[Struct_temp ==1] = struct_new
    Struct_mat = Struct_temp.T + np.eye(n) + Struct_temp
    return Struct_mat

def get_dist(name, n):
    Dist = []
    for i in range(1, n+1):
        distreader = csv.reader(open(name + "_" + str(i) + ".csv"))
        Dist.append(np.array([row for row in distreader]).astype(float))
    return Dist

def run_model(index, in_dir, out_dir, data_filename, func_filename, struct_file-
name, dist_filename, n, sample_size, tune_size):
    os.chdir(in_dir + str(index))
    Y = get_data(data_filename)
    mFunc = get_func(func_filename, n)
    Struct = get_struct(struct_filename, n)
```

33

```python
    Dist = get_dist(dist_filename, n)
    m = Dist[0].shape[0]
    k = Y.shape[1]
    n_vec= n*(n+1)//2
    Y_mean = []
    for i in range(n):
        Y_mean.append(np.mean(Y[i*m:(i+1)*m, 0]))
    Y_mean = np.array(Y_mean)

    with pm.Model() as model_generator:
        # convariance matrix
        log_Sig = pm.Uniform("log_Sig", -8, 8, shape=(n, ))
        SQ = tt.diag(tt.sqrt(tt.exp(log_Sig)))
        Func_Covm = tt.dot(tt.dot(SQ, mFunc), SQ)
        Struct_Convm = tt.dot(tt.dot(SQ, Struct), SQ)

        # double fusion of structural and FC
        L_fc_vec = tt.reshape(tt.slinalg.chole-
sky(tt.squeeze(Func_Covm)).T[np.triu_indices(n)], (n_vec, ))
        L_st_vec = tt.reshape(tt.slinalg.chole-
sky(tt.squeeze(Struct_Convm)).T[np.triu_indices(n)], (n_vec, ))
        Struct_vec = tt.reshape(Struct[np.triu_indices(n)], (n_vec, ))
        lambdaw = pm.Beta("lambdaw", alpha=1, beta=1, shape=(n_vec, ))
        Kf = pm.Beta("Kf", alpha=1, beta=1, shape=(n_vec, ))
        rhonn = Kf*( (1-lambdaw)*L_fc_vec + lambdaw*L_st_vec ) + \
            (1-Kf)*( (1-Struct_vec*lambdaw)*L_fc_vec + Struct_vec*lamb-
daw*L_st_vec )

        # correlation
        Cov_temp = tt.triu(tt.ones((n,n)))
        Cov_temp = tt.set_subtensor(Cov_temp[np.triu_indices(n)], rhonn)
        Cov_mat_v = tt.dot(Cov_temp.T, Cov_temp)
        d = tt.sqrt(tt.diagonal(Cov_mat_v))
        rho = (Cov_mat_v.T/d).T/d
        rhoNew = pm.Deterministic("rhoNew", rho[np.triu_indices(n,1)])

        # temporal correlation AR(1)
        phi_T = pm.Uniform("phi_T", 0, 1, shape=(n, ))
        sigW_T = pm.Uniform("sigW_T", 0, 100, shape=(n, ))
        B = pm.Normal("B", 0, 100, shape=(n, ))
        muW1 = Y_mean - B # get the shifted mean
        mean_overall = muW1/(1.0-phi_T) # AR(1) mean
        tau_overall = (1.0-tt.sqr(phi_T))/tt.sqr(sigW_T) # AR (1) variance
        W_T = pm.MvNormal("W_T", mu = mean_overall, tau = tt.diag(tau_overall),
shape = (k, n))
```

```python
        # add all parts together
        one_m_vec = tt.ones((m, 1))
        one_k_vec = tt.ones((1, k))

        D = pm.MvNormal("D", mu=tt.zeros(n), cov=Cov_mat_v, shape = (n, ))
        phi_s = pm.Uniform("phi_s", 0, 20, shape = (n, ))
        spat_prec = pm.Uniform("spat_prec", 0, 100, shape = (n, ))
        H_base = pm.Normal("H_base", 0, 1, shape = (m, n))

        Mu_all_temp = []
        for i in range(n):
            # exponential covariance function
            H_temp = tt.sqr(spat_prec[i])*tt.exp(-phi_s[i]*Dist[i])
            L_H_temp = tt.slinalg.cholesky(H_temp)
            Mu_all_temp.append(B[i] + D[i] + one_m_vec*W_T[:,i] +
tt.dot(L_H_temp, tt.reshape(H_base[:,i], (m, 1)))*one_k_vec)
        MU_all = tt.concatenate(Mu_all_temp, axis = 0)

        sigma_error_prec = pm.Uniform("sigma_error_prec", 0, 100)
        Y1 = pm.Normal("Y1", mu = MU_all, sd = sigma_error_prec, observed = Y)

    with model_generator:
        step = pm.NUTS()
        trace = pm.sample(sample_size, step = step, tune = tune_size, chains = 1)

    # save as pandas format and output the csv file
    save_trace = pm.trace_to_dataframe(trace)
    save_trace.to_csv(out_dir + date.today().strftime("%m_%d_%y") + "_sam-
ple_size_" + str(sample_size) + "_index_" + str(index) + ".csv")
```

```python
# initializing parameters
index_list = [8007, 8012, 8049, 8050, 8068, 8072, 8077, 8080, \
              8098, 8107, 8110, 8146, 8216, 8244, 8245, 8246, \
              8248, 8250, 8253, 8256, 8257, 8261, 8262, 8263, \
              8264, 8265, 8266, 8273, 8276, 8279, 8280, 8282, \
              8283, 8284, 8285, 8288, 8290, 8292, 8293, 8295, \
              8299]
in_dir = "/Users/ruiwang/source/doublefusion/simulation/data/"
out_dir = "/Users/ruiwang/source/doublefusion/simulation/results/"
data_filename = "ROI_timeseries_data"
func_filename = "DMN_MeanFunctional_Connectivity"
struct_filename = "DMN_StructuralConnectivity"
dist_filename = "distMatrix_ROI"
n = 14
sample_size = 1000
tune_size = 1000

# run the model
for index in index_list:
    run_model(index, in_dir, out_dir, data_filename, func_filename, struct_file-
name, dist_filename, n, sample_size, tune_size)
```

A.3 Histogram Plots of Parameters with 1000, 2000 and 3000 Sample Draws
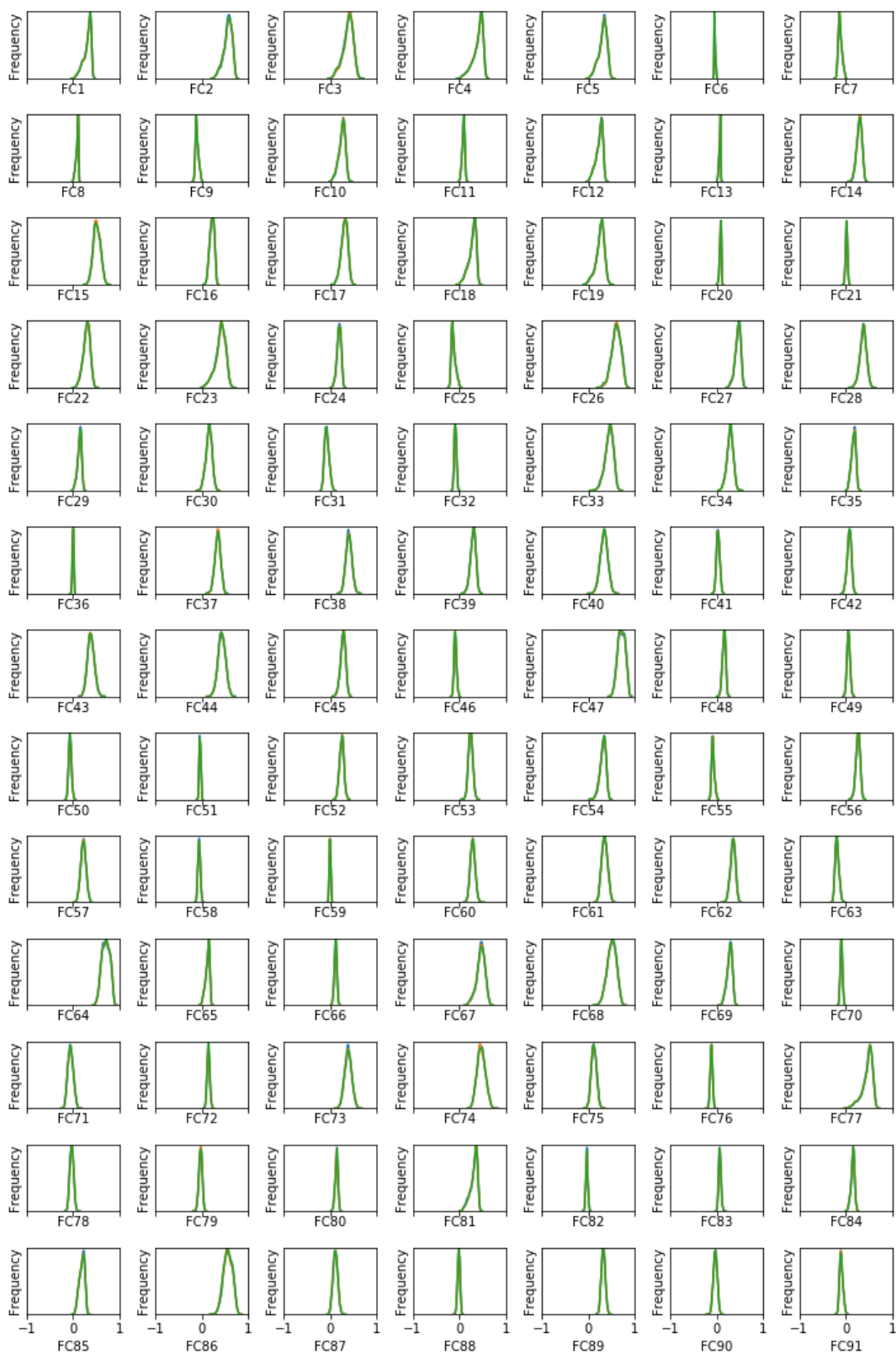
A.4 Regression Coefficients

	Intercept	Age	sex	Education	FC_i	Depress	FC_i*Depress
1	1.17	-0.01	-0.00	-0.03	0.26	0.09	-0.53
2	-0.41	-0.01	-0.13	-0.03	2.80	3.19	-5.91
3	0.68	-0.02	-0.07	-0.02	1.00	1.05	-2.70
4	0.86	-0.02	-0.15	-0.01	1.07	-0.37	0.63
5	1.28	-0.01	-0.10	-0.04	0.75	-0.09	-0.10
6	1.18	-0.02	0.07	-0.02	0.83	-0.27	0.62
7	1.38	-0.02	0.05	-0.03	-0.15	-0.28	1.48
8	1.29	-0.02	0.02	-0.03	0.21	-0.14	0.44
9	1.12	-0.02	0.06	-0.02	-0.03	-0.08	1.00
10	0.60	-0.01	-0.08	-0.03	2.03	0.08	-0.47
11	1.09	-0.02	0.03	-0.01	-1.07	-0.10	0.06
12	1.05	-0.01	0.04	-0.02	-0.58	-0.08	1.37
13	1.66	-0.02	-0.08	-0.05	-1.08	-0.13	1.74
14	0.52	-0.02	0.02	-0.01	1.87	0.71	-3.25
15	-0.15	-0.02	-0.06	-0.01	2.72	2.23	-5.16
16	1.27	-0.02	0.03	-0.04	1.28	0.42	-3.23
17	1.59	-0.02	0.03	-0.04	-0.11	0.27	-1.52
18	1.25	-0.01	-0.01	-0.03	-0.04	-0.10	1.19
19	1.18	-0.01	0.04	-0.03	-0.93	-0.17	1.78
20	1.51	-0.02	-0.02	-0.05	-0.91	-0.09	0.50
21	1.35	-0.02	-0.01	-0.04	-0.54	-0.06	0.72
22	1.38	-0.01	-0.04	-0.04	0.14	-0.11	0.45
23	1.42	-0.02	0.01	-0.03	-0.27	-0.15	-0.02
24	1.40	-0.02	0.02	-0.03	-1.06	-0.24	1.38
25	1.50	-0.02	-0.05	-0.05	-1.44	-0.02	0.98
26	-0.66	-0.01	-0.10	-0.07	4.01	1.89	-3.44
27	1.05	-0.01	0.00	-0.04	0.74	0.56	-1.60
28	0.16	-0.00	-0.06	-0.05	3.24	1.10	-3.57
29	1.16	-0.02	-0.01	-0.02	0.36	-0.65	1.75
30	1.46	-0.02	0.01	-0.03	-0.66	-0.63	2.47
31	1.64	-0.02	-0.06	-0.05	0.98	-0.19	0.35
32	1.09	-0.02	-0.02	-0.02	-0.63	-0.10	2.11
33	1.05	-0.02	-0.04	-0.02	0.58	-0.62	1.37
34	1.37	-0.02	-0.00	-0.02	-1.32	-0.47	2.16
35	0.52	-0.01	-0.06	0.02	-1.39	-0.25	2.86
36	1.50	-0.02	-0.12	-0.03	-1.97	-0.17	3.12
37	1.29	-0.01	0.02	-0.04	0.47	0.63	-2.56
38	1.04	-0.01	-0.02	-0.04	0.84	0.49	-1.61
39	1.02	-0.02	0.06	0.00	-0.56	-0.63	2.31
40	1.28	-0.02	0.05	-0.02	-1.07	-0.49	1.89
41	1.43	-0.02	-0.03	-0.03	0.44	-0.16	0.70
42	1.49	-0.02	-0.03	-0.04	0.10	-0.12	1.47
43	1.33	-0.02	-0.00	-0.02	-0.67	-0.63	2.12
44	1.61	-0.02	-0.02	-0.02	-1.24	-0.23	0.64
45	0.91	-0.01	-0.01	-0.00	-0.65	-0.38	1.74
46	1.06	-0.01	-0.05	-0.02	-1.48	-0.02	2.21
47	-1.14	-0.01	-0.05	-0.06	4.07	3.51	-5.65
48	0.92	-0.02	-0.03	0.01	0.14	-0.42	2.16
49	1.12	-0.02	-0.01	-0.01	-0.25	-0.29	1.58
50	1.57	-0.01	0.03	-0.05	1.08	-0.11	-0.95
51	1.39	-0.01	-0.03	-0.04	-0.52	-0.12	-0.33
52	1.29	-0.02	-0.16	-0.03	1.14	-0.18	0.53
53	1.32	-0.02	-0.02	-0.02	-0.98	-0.22	1.44
54	1.20	-0.01	-0.00	-0.03	-0.25	-0.15	0.56
55	1.26	-0.02	-0.03	-0.02	-0.13	-0.06	1.44
56	0.90	-0.02	0.04	0.00	-0.24	-0.39	2.59
57	1.20	-0.01	0.02	-0.02	-0.41	-0.22	1.07
58	1.43	-0.01	0.04	-0.04	1.19	-0.17	-0.74
59	1.41	-0.01	-0.01	-0.04	0.19	-0.13	-0.05
60	1.38	-0.02	-0.10	-0.04	0.75	-0.11	0.15
61	1.68	-0.01	0.02	-0.04	-1.41	-0.26	0.64
62	1.50	-0.01	-0.06	-0.04	0.45	-0.10	-0.15
63	1.34	-0.02	-0.00	-0.03	0.67	-0.08	0.53
64	1.56	-0.01	-0.05	-0.04	-0.07	-2.47	3.56
65	1.39	-0.02	-0.05	-0.03	0.00	-0.33	1.66
66	1.41	-0.02	-0.11	-0.03	0.11	-0.29	1.85
67	1.53	-0.01	0.01	-0.03	-0.80	-0.60	1.22
68	1.51	-0.01	-0.02	-0.04	-0.45	0.08	-0.93
69	1.39	-0.01	-0.03	-0.04	0.09	-0.14	0.02
70	1.22	-0.01	-0.11	-0.02	1.96	-0.22	-0.11
71	1.44	-0.02	-0.05	-0.04	1.25	-0.07	-0.01
72	1.30	-0.02	-0.06	-0.02	0.22	-0.30	1.87
73	1.43	-0.02	0.06	-0.02	-1.46	-0.61	2.49
74	1.53	-0.02	0.01	-0.03	-0.98	-0.16	0.22
75	1.35	-0.01	-0.01	-0.03	-0.38	-0.14	0.62
76	2.08	-0.02	-0.19	-0.06	2.26	-0.21	-1.09
77	0.76	-0.01	-0.12	-0.04	1.08	-0.43	0.69
78	1.21	-0.02	-0.02	-0.02	-0.74	-0.19	2.62
79	1.07	-0.02	0.01	-0.02	-1.12	-0.04	2.58
80	2.31	-0.01	-0.01	-0.08	-2.82	-0.46	3.30
81	1.73	-0.02	-0.08	-0.04	-0.84	-0.63	2.30
82	1.56	-0.02	0.01	-0.04	0.98	-0.21	-0.14
83	1.53	-0.01	0.00	-0.05	0.54	-0.14	0.01
84	0.85	-0.01	0.07	-0.02	1.34	0.10	-1.38
85	1.00	-0.01	-0.00	-0.02	1.06	-0.03	-0.03
86	2.43	-0.01	0.12	0.00	-3.98	-2.86	6.38
87	1.58	-0.02	-0.05	-0.04	0.05	0.10	-1.80
88	0.40	-0.02	-0.01	0.03	0.78	-0.23	1.18
89	1.68	-0.02	0.06	-0.03	-1.32	-0.12	0.12
90	0.95	-0.01	-0.02	-0.01	0.20	-0.12	1.66
91	1.15	-0.02	0.00	-0.02	0.47	-0.06	-1.13

	Intercept	Age	sex	Education	FC_i	Depress	FC_i*Depress
1	0.57	-0.03	-0.22	0.04	0.01	-0.68	1.31
2	0.80	-0.03	-0.06	0.06	-0.88	-2.60	4.33
3	0.97	-0.03	-0.10	0.04	-0.85	-1.81	3.77
4	0.20	-0.03	-0.12	0.08	-0.77	-0.68	1.78
5	0.29	-0.03	-0.09	0.07	-0.88	-0.47	1.37
6	-0.02	-0.04	-0.06	0.08	0.61	-0.37	1.06
7	0.26	-0.04	-0.08	0.07	0.80	-0.26	0.68
8	0.25	-0.03	-0.16	0.06	0.01	-0.16	-0.24
9	0.27	-0.03	-0.18	0.06	-0.15	-0.17	-0.29
10	-0.15	-0.03	-0.16	0.09	0.15	-0.46	0.99
11	0.00	-0.03	-0.10	0.08	-0.91	-0.30	1.32
12	0.00	-0.03	-0.12	0.08	-0.31	-0.13	0.78
13	0.02	-0.03	-0.10	0.08	0.59	-0.18	-0.25
14	0.28	-0.03	-0.16	0.06	-0.11	-0.24	0.34
15	0.38	-0.03	-0.14	0.07	-0.46	-0.03	-0.37
16	0.08	-0.03	-0.15	0.07	0.73	-0.00	-1.01
17	0.22	-0.03	-0.15	0.07	-0.07	-0.16	0.02
18	-0.01	-0.03	-0.16	0.08	0.30	-0.08	1.98
19	0.25	-0.03	-0.16	0.06	0.61	-0.17	-0.09
20	0.35	-0.03	-0.13	0.06	0.81	-0.33	-1.55
21	0.32	-0.03	-0.15	0.06	0.15	-0.23	-0.59
22	0.09	-0.03	-0.21	0.07	-0.81	-0.17	3.43
23	0.29	-0.03	-0.11	0.07	-0.62	-0.20	0.07
24	-0.06	-0.03	-0.16	0.08	0.02	-0.17	0.76
25	0.08	-0.03	-0.12	0.08	1.29	-0.28	-0.99
26	0.04	-0.03	-0.24	0.06	0.69	-1.36	2.08
27	0.13	-0.03	-0.12	0.06	0.28	0.46	-1.43
28	0.33	-0.03	-0.11	0.07	-0.31	0.04	-0.65
29	0.05	-0.03	-0.15	0.07	0.55	-0.31	0.47
30	0.18	-0.03	-0.15	0.07	0.09	-0.20	0.14
31	0.12	-0.03	-0.13	0.07	0.85	-0.15	-1.25
32	0.27	-0.03	-0.13	0.06	-0.05	-0.16	-1.00
33	0.24	-0.04	-0.13	0.10	-1.31	-1.10	2.66
34	0.35	-0.03	-0.18	0.08	-1.40	-0.32	1.26
35	-0.20	-0.03	-0.20	0.09	0.38	-0.17	0.60
36	-0.09	-0.03	-0.10	0.09	0.75	-0.22	0.13
37	0.26	-0.03	-0.13	0.07	-0.19	-0.08	-0.26
38	0.60	-0.03	-0.12	0.06	-0.87	-0.10	-0.18
39	0.18	-0.03	-0.15	0.07	0.06	-0.17	0.05
40	0.34	-0.03	-0.17	0.06	-0.13	-0.02	-0.62
41	0.34	-0.03	-0.15	0.06	0.73	-0.17	-1.14
42	0.37	-0.03	-0.17	0.06	0.41	-0.16	0.03
43	0.15	-0.04	-0.13	0.09	-0.89	-0.80	2.75
44	0.37	-0.03	-0.15	0.07	-0.86	-0.23	0.47
45	0.15	-0.03	-0.16	0.06	0.66	-0.11	-0.30
46	0.25	-0.03	-0.13	0.07	0.81	-0.21	-0.52
47	1.34	-0.03	-0.20	0.05	-1.18	-2.53	3.79
48	0.13	-0.03	-0.18	0.08	0.61	-0.21	0.18
49	0.16	-0.03	-0.15	0.07	-0.17	-0.19	0.40
50	0.50	-0.03	-0.11	0.04	1.47	-0.08	-2.74
51	0.36	-0.03	-0.17	0.06	0.57	-0.14	-0.97
52	0.08	-0.03	-0.19	0.09	-0.54	-0.48	2.44
53	0.33	-0.03	-0.12	0.06	0.07	-0.13	-0.57
54	0.56	-0.03	-0.17	0.05	0.52	-0.13	-1.06
55	0.20	-0.03	-0.10	0.07	0.65	-0.21	-0.81
56	0.48	-0.03	-0.18	0.04	0.12	-0.01	-1.50
57	0.48	-0.03	-0.14	0.04	-0.28	-0.03	-1.38
58	0.31	-0.03	-0.14	0.06	0.89	-0.17	-1.16
59	0.23	-0.03	-0.13	0.07	0.44	-0.14	0.35
60	0.36	-0.03	-0.19	0.06	-0.44	-0.39	2.16
61	0.22	-0.03	-0.13	0.06	0.14	-0.05	-0.87
62	0.18	-0.03	-0.21	0.07	0.27	-0.17	0.45
63	0.21	-0.03	-0.15	0.07	0.05	-0.17	-0.14
64	0.37	-0.03	-0.18	0.06	-0.10	-1.95	2.72
65	0.15	-0.03	-0.13	0.08	-1.04	-0.30	0.81
66	0.12	-0.03	-0.14	0.07	-0.61	-0.23	0.72
67	0.07	-0.03	-0.09	0.08	0.11	-1.17	2.48
68	0.26	-0.03	-0.15	0.07	-0.45	-0.16	-0.02
69	0.24	-0.03	-0.23	0.07	0.37	-0.17	0.26
70	-0.00	-0.03	-0.18	0.08	0.03	-0.18	1.05
71	0.19	-0.03	-0.14	0.07	-0.40	-0.20	0.25
72	0.15	-0.03	-0.13	0.08	-1.02	-0.29	1.11
73	0.18	-0.03	-0.11	0.08	-0.69	-0.51	1.73
74	0.32	-0.03	-0.12	0.08	-1.21	-0.31	0.81
75	0.19	-0.03	-0.15	0.07	-0.30	-0.16	0.47
76	0.40	-0.03	-0.21	0.06	0.57	-0.19	0.29
77	0.57	-0.03	-0.34	0.07	-0.61	-1.77	3.38
78	0.42	-0.03	-0.15	0.06	1.55	-0.15	-2.04
79	0.28	-0.03	-0.17	0.06	0.35	-0.21	-1.46
80	0.54	-0.03	-0.19	0.05	-0.48	-0.26	1.65
81	0.55	-0.03	-0.22	0.07	-0.77	-0.65	2.31
82	0.46	-0.03	-0.16	0.05	1.00	-0.19	-2.41
83	0.24	-0.03	-0.18	0.06	-0.04	-0.16	-2.05
84	0.09	-0.03	-0.13	0.07	0.30	-0.10	-0.32
85	0.47	-0.03	-0.20	0.07	-0.91	-0.45	2.14
86	1.20	-0.03	-0.09	0.07	-2.56	-1.08	1.96
87	0.21	-0.03	-0.17	0.07	0.31	-0.08	-0.70
88	-0.22	-0.03	-0.14	0.09	0.60	-0.17	0.04
89	0.54	-0.03	-0.10	0.07	-1.27	-0.33	0.91
90	-0.19	-0.03	-0.16	0.09	-0.36	-0.12	2.07
91	0.41	-0.03	-0.16	0.06	-0.37	-0.16	-0.30

	Intercept	Age	sex	Education	FC_i	Depress	FC_i*Depress
1	0.39	-0.00	0.24	-0.01	-0.20	0.05	-0.13
2	0.50	-0.00	0.19	-0.01	-0.22	0.57	-1.01
3	1.23	0.00	0.30	-0.03	-1.61	-0.90	1.99
4	0.15	-0.00	0.19	0.01	-0.41	-0.68	2.28
5	0.44	-0.00	0.25	-0.01	-0.76	-0.34	1.57
6	0.31	-0.00	0.25	-0.01	0.48	-0.01	0.04
7	0.38	0.00	0.22	-0.02	0.01	0.00	0.04
8	0.49	0.00	0.20	-0.02	0.21	0.00	-0.62
9	0.68	-0.00	0.18	-0.03	0.65	-0.07	-1.05
10	-0.04	-0.00	0.20	-0.00	0.61	-0.13	0.50
11	0.13	-0.00	0.25	0.00	-0.93	0.03	0.10
12	-0.02	0.00	0.29	0.00	-0.80	0.08	1.77
13	0.54	-0.00	0.18	-0.03	-0.55	0.02	0.47
14	0.00	-0.00	0.25	0.00	0.55	0.40	-1.65
15	0.40	-0.00	0.23	-0.01	-0.23	0.55	-1.29
16	0.29	0.00	0.21	-0.02	0.36	0.03	-0.20
17	0.35	0.00	0.21	-0.01	-0.13	-0.14	0.56
18	0.43	-0.00	0.19	-0.02	0.82	0.03	-0.48
19	0.49	-0.00	0.19	-0.02	0.60	0.02	-1.01
20	0.66	-0.00	0.21	-0.04	-0.71	-0.04	-0.35
21	0.51	0.00	0.21	-0.02	0.41	-0.10	-0.97
22	0.41	0.00	0.18	-0.02	0.45	0.05	-0.10
23	0.44	-0.00	0.24	-0.02	-0.18	0.05	-0.34
24	0.28	0.00	0.24	-0.01	-1.10	-0.11	1.79
25	0.43	-0.00	0.19	-0.02	-1.07	0.03	0.26
26	-1.67	0.01	0.10	-0.05	4.16	1.56	-2.64
27	1.33	-0.00	0.26	-0.02	-1.82	-0.74	1.72
28	0.10	0.00	0.22	-0.02	0.73	0.38	-1.11
29	0.21	-0.00	0.21	-0.02	0.86	0.05	-0.27
30	0.38	-0.00	0.22	-0.02	-0.04	-0.06	0.30
31	0.43	-0.00	0.21	-0.02	0.79	-0.01	-0.55
32	0.54	-0.00	0.18	-0.02	0.70	-0.04	-0.30
33	0.04	-0.00	0.18	-0.02	1.44	0.20	-0.55
34	0.28	-0.00	0.24	-0.01	0.07	-0.11	0.60
35	-0.56	0.00	0.14	0.04	-0.38	-0.07	2.29
36	0.24	-0.00	0.19	0.00	-0.66	-0.07	2.03
37	0.24	-0.00	0.18	-0.02	0.53	0.04	-0.13
38	0.46	-0.00	0.22	-0.02	-0.19	-0.10	0.26
39	0.27	-0.00	0.17	-0.02	1.05	0.12	-0.64
40	0.32	-0.00	0.20	-0.02	0.72	0.09	-0.55
41	0.56	-0.00	0.21	-0.02	1.06	-0.02	-1.11
42	1.06	-0.00	0.15	-0.05	1.98	-0.02	-1.55
43	0.34	-0.00	0.23	0.00	-0.68	-0.49	2.11
44	0.51	-0.00	0.22	-0.01	-0.66	-0.11	0.63
45	0.16	0.00	0.21	-0.00	-0.04	-0.08	0.64
46	0.11	0.00	0.20	-0.00	-0.78	0.08	1.67
47	-1.69	0.01	0.15	-0.05	3.66	2.62	-4.00
48	0.14	-0.00	0.19	0.01	0.35	-0.14	1.01
49	0.18	-0.00	0.21	-0.00	0.08	-0.11	0.94
50	0.59	0.00	0.27	-0.03	1.22	0.01	-0.53
51	0.27	-0.00	0.24	-0.01	0.05	-0.02	0.92
52	0.27	-0.00	0.22	0.00	-0.97	-0.33	2.47
53	0.08	-0.00	0.15	0.01	0.35	-0.03	0.76
54	0.32	0.00	0.12	-0.01	0.36	0.06	0.76
55	0.25	-0.00	0.28	-0.01	0.98	-0.01	-0.06
56	0.44	-0.00	0.18	-0.01	0.66	-0.00	-0.06
57	0.38	-0.00	0.20	-0.01	0.39	0.00	-0.12
58	0.37	-0.00	0.28	-0.01	1.03	-0.04	-0.13
59	0.39	-0.00	0.25	-0.02	0.49	0.04	0.56
60	0.43	0.00	0.23	-0.02	-0.32	-0.08	0.65
61	0.23	-0.00	0.19	-0.01	0.59	-0.00	0.21
62	0.61	-0.00	0.09	-0.03	1.03	0.07	-0.15
63	0.38	-0.00	0.21	-0.02	-0.15	0.04	0.34
64	-1.04	0.00	0.26	-0.02	1.93	1.13	-1.71
65	0.40	-0.00	0.21	-0.02	0.31	0.01	0.03
66	0.40	-0.00	0.17	-0.01	0.08	-0.07	0.84
67	0.26	-0.00	0.21	-0.02	0.57	0.27	-0.72
68	0.43	-0.00	0.18	-0.04	1.59	0.74	-2.89
69	0.33	0.00	0.25	-0.01	-0.30	-0.01	0.21
70	0.19	0.00	0.16	-0.00	0.81	-0.04	0.56
71	0.40	-0.00	0.20	-0.02	1.03	0.13	-1.02
72	0.34	-0.00	0.19	-0.01	0.21	-0.07	0.90
73	0.43	-0.00	0.23	-0.01	-0.45	-0.04	0.34
74	0.44	0.00	0.19	-0.04	1.01	0.38	-1.75
75	0.34	-0.00	0.19	-0.01	-0.57	-0.00	1.25
76	0.66	-0.00	0.15	-0.03	0.91	-0.02	-0.56
77	-0.40	0.01	0.03	-0.02	1.35	-0.75	1.68
78	0.27	0.00	0.22	-0.01	-0.89	0.01	0.89
79	0.06	0.00	0.21	-0.00	-0.99	0.05	0.88
80	1.62	0.00	0.18	-0.07	-3.22	-0.41	4.85
81	0.87	-0.00	0.14	-0.01	-1.50	-0.59	2.59
82	0.59	0.00	0.21	-0.03	0.93	-0.04	-1.65
83	0.52	0.00	0.20	-0.03	0.36	-0.00	-1.79
84	0.27	0.00	0.24	-0.02	0.42	0.04	-0.19
85	0.01	-0.00	0.20	-0.00	0.96	0.02	0.87
86	1.09	0.00	0.28	-0.01	-2.12	-1.04	2.37
87	0.64	-0.00	0.19	-0.03	-0.28	0.22	-1.53
88	-0.02	-0.00	0.19	0.02	-0.54	-0.12	2.07
89	0.67	-0.00	0.26	-0.02	-1.09	-0.18	0.94
90	-0.12	0.00	0.22	0.01	-1.05	0.08	3.25
91	0.44	0.00	0.21	-0.02	-0.16	-0.05	0.84

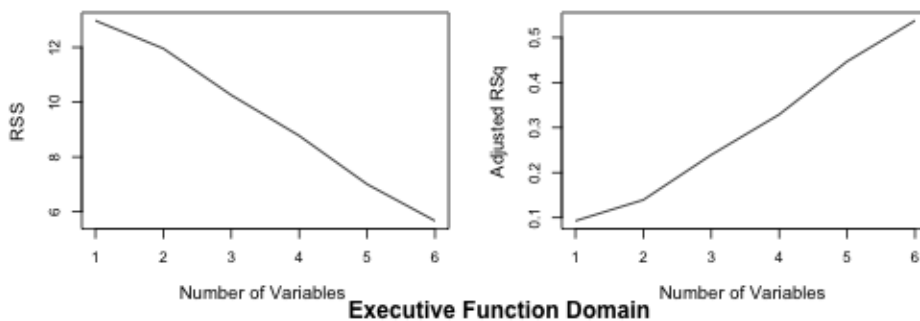|	Intercept	Age	sex	Education	FC_i	Depress	FC_i*Depress
1	1.78	-0.03	0.33	-0.04	-0.42	-0.15	0.09
2	0.53	-0.03	0.19	-0.04	1.72	2.90	-5.33
3	1.25	-0.03	0.26	-0.05	0.81	0.29	-0.85
4	1.69	-0.03	0.28	-0.07	0.74	0.46	-1.87
5	1.70	-0.03	0.35	-0.05	-0.19	0.02	-0.61
6	1.79	-0.03	0.26	-0.07	1.93	0.44	-2.96
7	1.75	-0.03	0.27	-0.05	1.55	0.10	-2.48
8	2.08	-0.03	0.23	-0.07	0.82	-0.10	-2.27
9	2.31	-0.03	0.19	-0.08	0.86	-0.24	-2.34
10	1.41	-0.03	0.22	-0.09	2.71	1.04	-3.41
11	1.54	-0.03	0.34	-0.04	-0.63	-0.29	1.71
12	2.18	-0.03	0.28	-0.08	0.10	-0.16	-1.27
13	1.71	-0.03	0.29	-0.05	-0.15	-0.09	-0.08
14	1.33	-0.03	0.33	-0.03	0.48	0.25	-1.44
15	1.81	-0.03	0.31	-0.05	-0.31	-0.19	0.18
16	1.49	-0.03	0.32	-0.05	1.17	0.25	-2.10
17	1.69	-0.03	0.32	-0.05	0.18	0.08	-0.67
18	1.69	-0.03	0.34	-0.05	-1.01	-0.15	-0.35
19	1.48	-0.03	0.34	-0.04	-1.21	-0.11	1.39
20	2.00	-0.03	0.28	-0.08	-1.78	-0.04	0.66
21	1.77	-0.03	0.30	-0.06	-1.04	-0.01	0.81
22	1.72	-0.03	0.34	-0.05	0.42	-0.09	-1.95
23	1.60	-0.03	0.30	-0.05	-0.11	-0.19	0.64
24	1.62	-0.03	0.30	-0.05	-0.01	-0.10	0.11
25	1.71	-0.03	0.27	-0.06	-1.06	-0.07	0.24
26	-0.44	-0.02	0.21	-0.08	4.12	2.02	-3.60
27	1.74	-0.03	0.37	-0.06	0.10	0.83	-2.11
28	1.37	-0.03	0.32	-0.05	0.76	0.47	-1.69
29	1.61	-0.03	0.29	-0.07	1.18	0.65	-2.73
30	1.63	-0.03	0.30	-0.05	0.29	0.15	-1.22
31	1.47	-0.03	0.36	-0.04	2.25	-0.07	-3.18
32	1.76	-0.03	0.31	-0.05	0.24	-0.11	-0.72
33	1.78	-0.02	0.30	-0.08	0.43	0.59	-1.95
34	1.60	-0.03	0.33	-0.02	-1.65	-0.59	3.10
35	2.09	-0.03	0.35	-0.07	-0.12	-0.07	-0.85
36	2.10	-0.03	0.17	-0.07	-2.18	-0.04	1.80
37	0.99	-0.02	0.18	-0.06	2.69	0.76	-2.94
38	0.94	-0.03	0.28	-0.05	1.76	1.04	-2.90
39	1.98	-0.02	0.23	-0.08	0.76	0.41	-2.37
40	1.66	-0.03	0.31	-0.05	-0.14	-0.12	0.17
41	1.72	-0.03	0.31	-0.06	0.22	-0.09	-1.33
42	1.52	-0.03	0.32	-0.05	-0.27	-0.10	-0.79
43	1.69	-0.02	0.29	-0.07	0.65	0.35	-1.91
44	1.89	-0.03	0.33	-0.05	-0.89	-0.69	2.63
45	2.17	-0.03	0.30	-0.09	1.10	0.21	-2.18
46	1.59	-0.02	0.26	-0.06	-1.99	0.03	1.11
47	-0.80	-0.02	0.19	-0.10	4.65	2.67	-4.17
48	2.07	-0.02	0.29	-0.08	0.49	0.16	-2.13
49	1.74	-0.03	0.30	-0.06	0.11	-0.04	-0.53
50	1.83	-0.03	0.33	-0.07	0.89	-0.04	-1.70
51	1.85	-0.03	0.28	-0.06	0.48	-0.07	-1.31
52	1.72	-0.03	0.18	-0.07	2.24	0.31	-2.80
53	1.63	-0.03	0.30	-0.05	0.36	-0.07	-0.32
54	2.17	-0.03	0.15	-0.08	1.34	0.03	-0.75
55	1.74	-0.02	0.49	-0.06	2.25	-0.32	-4.01
56	1.95	-0.03	0.27	-0.07	0.23	0.06	-1.55
57	1.61	-0.03	0.30	-0.04	0.03	-0.12	0.23
58	1.82	-0.02	0.32	-0.07	1.31	-0.11	-1.84
59	1.86	-0.03	0.28	-0.06	0.64	-0.13	-1.29
60	1.46	-0.03	0.15	-0.05	2.33	0.28	-2.42
61	1.60	-0.03	0.25	-0.04	-0.14	-0.32	1.73
62	2.36	-0.03	0.13	-0.09	2.01	0.12	-1.61
63	1.69	-0.02	0.36	-0.06	1.65	-0.47	-4.08
64	1.00	-0.03	0.40	-0.03	0.55	4.95	-7.69
65	1.71	-0.03	0.31	-0.06	1.17	0.20	-2.10
66	1.65	-0.03	0.33	-0.05	0.01	-0.05	-0.50
67	1.59	-0.03	0.24	-0.06	0.81	1.18	-3.23
68	1.75	-0.03	0.32	-0.04	-1.08	-0.21	0.37
69	1.61	-0.03	0.51	-0.05	-0.79	-0.04	-1.08
70	2.00	-0.03	0.30	-0.07	1.27	-0.11	-2.33
71	1.70	-0.03	0.27	-0.07	2.31	0.24	-2.79
72	1.64	-0.03	0.30	-0.05	-0.04	-0.12	0.31
73	1.80	-0.02	0.26	-0.07	0.17	0.33	-1.93
74	1.75	-0.03	0.31	-0.05	-0.44	-0.07	-0.07
75	1.65	-0.03	0.41	-0.05	-0.28	-0.09	-0.78
76	2.38	-0.03	0.15	-0.08	2.55	-0.13	-3.94
77	0.43	-0.02	0.45	-0.06	2.11	1.80	-3.89
78	1.73	-0.03	0.31	-0.06	0.17	-0.05	-1.59
79	1.09	-0.03	0.29	-0.02	-1.80	0.00	1.92
80	2.18	-0.03	0.30	-0.07	-1.51	-0.27	1.90
81	1.68	-0.03	0.33	-0.04	-0.54	0.02	-0.77
82	1.89	-0.03	0.29	-0.06	0.92	-0.12	-2.18
83	1.75	-0.03	0.31	-0.05	0.32	-0.10	-0.13
84	1.35	-0.03	0.38	-0.05	1.25	-0.01	-0.47
85	1.35	-0.03	0.36	-0.05	1.05	0.22	-2.24
86	2.73	-0.02	0.34	-0.06	-2.34	-0.49	0.66
87	1.92	-0.03	0.18	-0.06	1.77	0.65	-6.20
88	1.87	-0.03	0.31	-0.06	0.17	-0.04	-0.88
89	1.65	-0.03	0.32	-0.04	-0.10	0.02	-0.50
90	1.88	-0.03	0.32	-0.06	-1.32	-0.05	0.40
91	1.21	-0.03	0.33	-0.02	0.91	0.02	-1.60

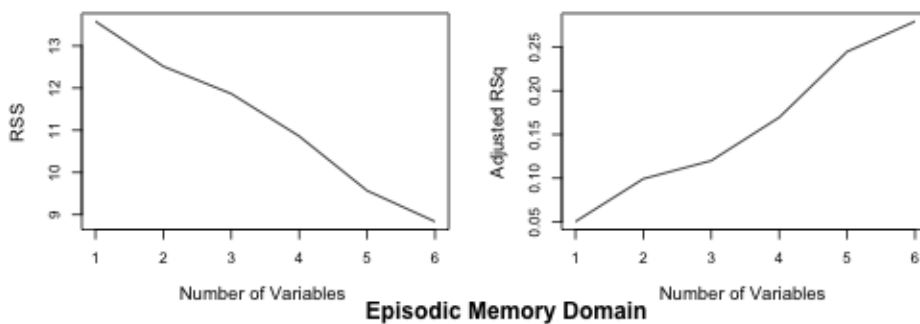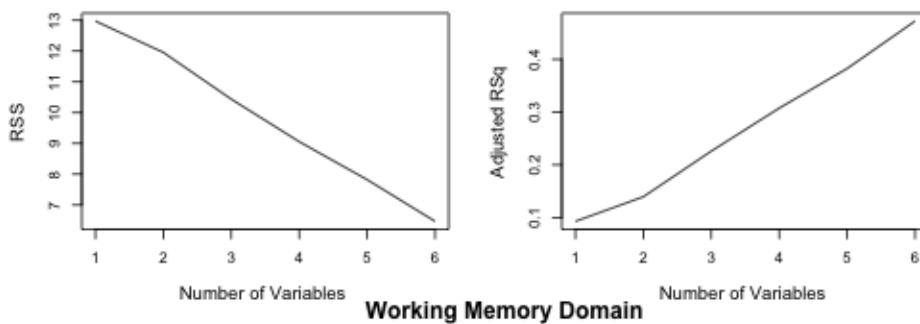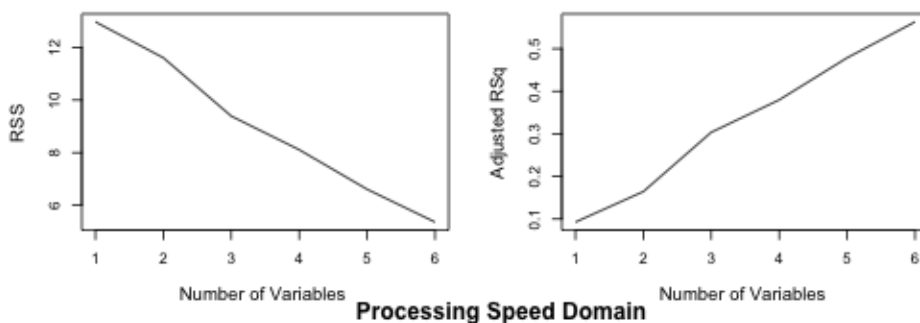## A.5 Variable Selection with Lasso Method[1]

```r
require(boot)
require(glmnet)

get_lasso_selection <- function(data, fixed_indices, selected_indices, re-
sponse_indice, n_boot) {
  df <- as.matrix(data)
  pnlty <- c(rep(0, length(fixed_indices)), rep(1, length(selected_indices)))
  sim <- function(x,y,z) {
    cv <- cv.glmnet(x, y, nfold = z, type.measure = "mse", penalty.factor =
pnlty)
    return(cv$lambda.1se)
    }
  boot_sim <- function(dat, indices){
    data <- dat[indices,]
    result <- sim(x = data[, c(fixed_indices, selected_indices)],
                  y = data[, response_indice], z = 10)
    return(result)
  }
  boot_result <- boot(df, boot_sim, R = n_boot)
  lambda <- boot_result$t
  # create the table for the result
  mat <- matrix(0, nrow = 1, ncol = length(fixed_indices) + length(selected_indi-
ces))
  for(i in 1:length(lambda)){
    beta_r <- glmnet(x = df[, c(fixed_indices, selected_indices)],
                     y = df[, response_indice], lambda = lambda[i],
                     penalty.factor=pnlty)$beta
    for(j in 1:ncol(mat)){
      if (beta_r[j] != 0){
        mat[1,j] <- mat[1,j] + 1
      }
    }
  }
  # change the count to percentage
  mat <- mat/(n_boot/100)
  colnames(mat) <- colnames(df)[c(fixed_indices, selected_indices)]
  rownames(mat) <- "Percent"
  return(mat)
}
```
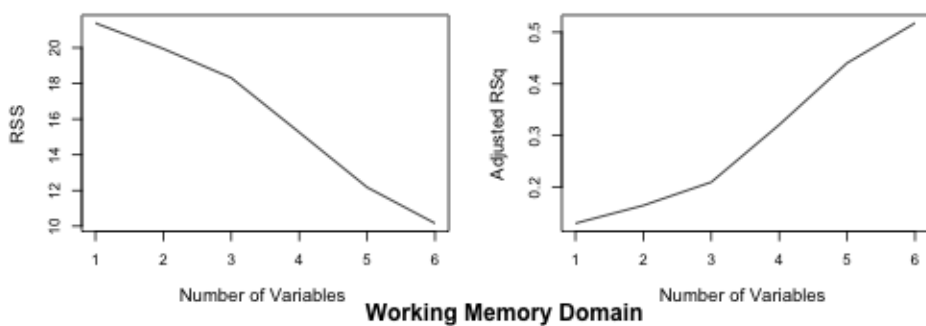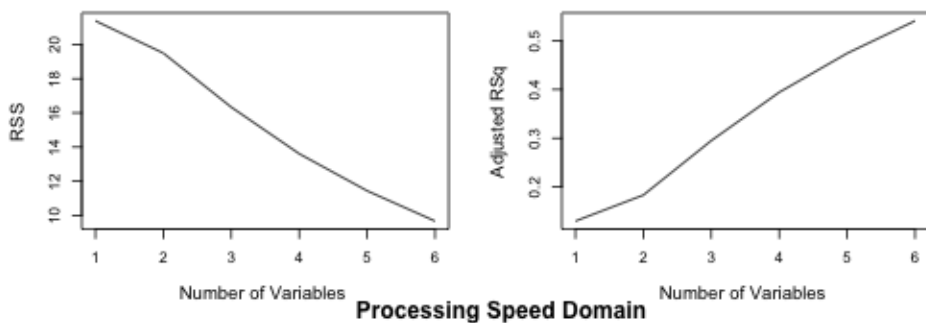
---

[1] Modified codes. Ya-Chen Lisa Lin credits.

A.6 Plots of Subset Selection

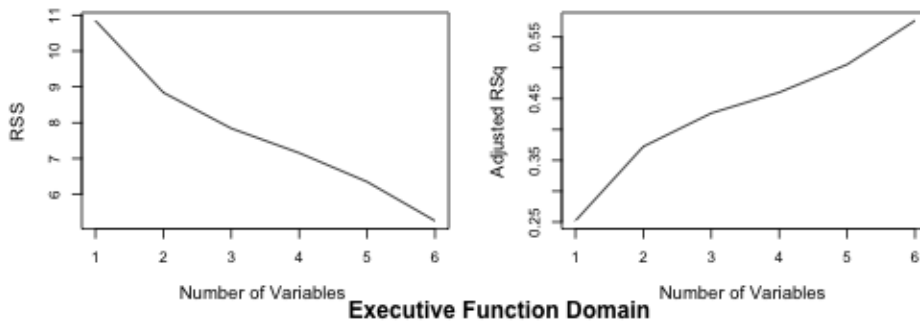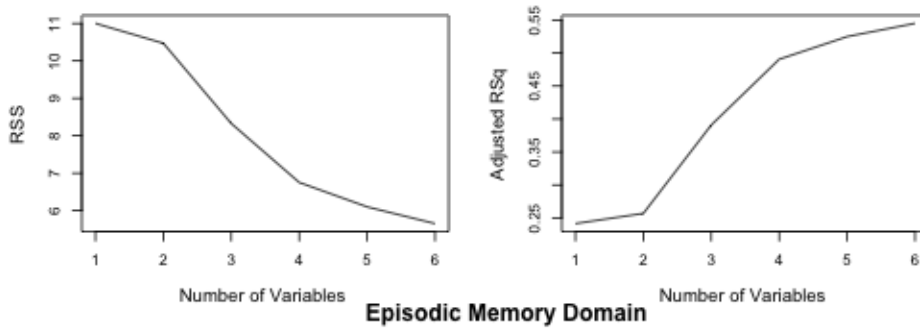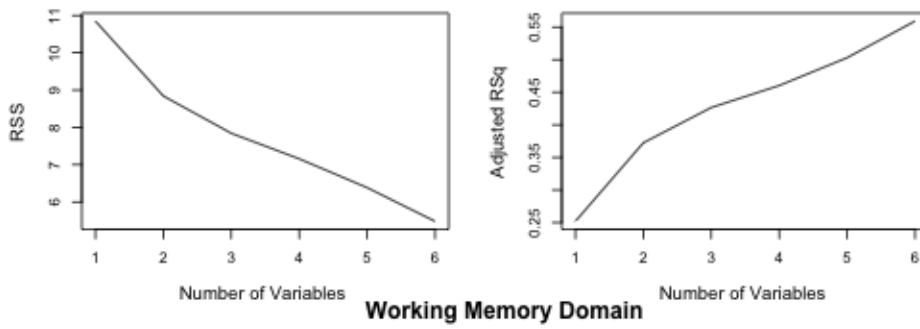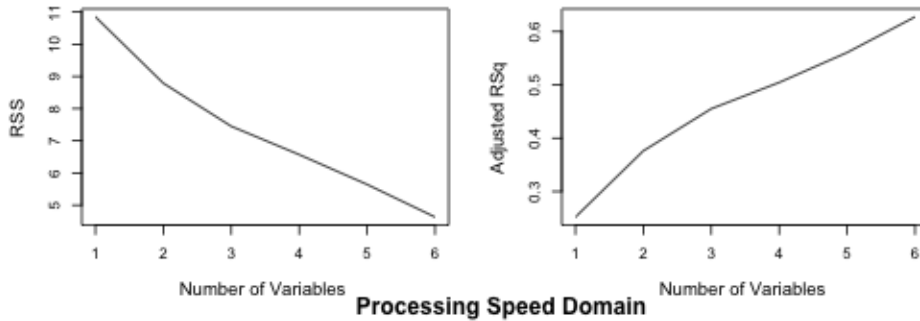Exhaustive method

Processing Speed Domain

Working Memory Domain

Episodic Memory Domain

Executive Function Domain

Forward method

Processing Speed Domain

Working Memory Domain

Episodic Memory Domain

Executive Function Domain

44

Backward method

Processing Speed Domain

Working Memory Domain

Episodic Memory Domain

Executive Function Domain

Sequential method

Processing Speed Domain

Working Memory Domain

Episodic Memory Domain

Executive Function Domain

46
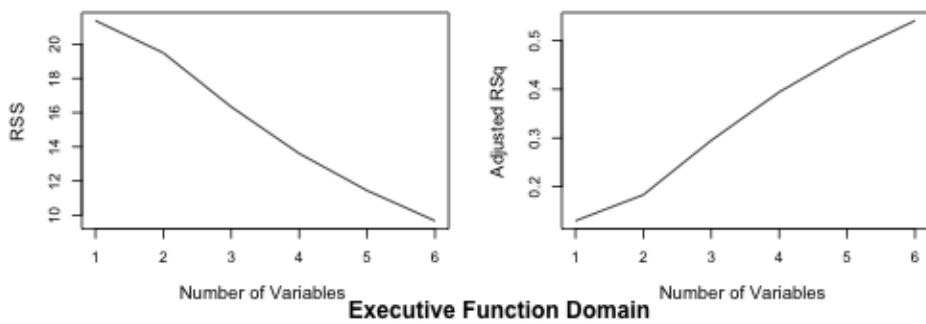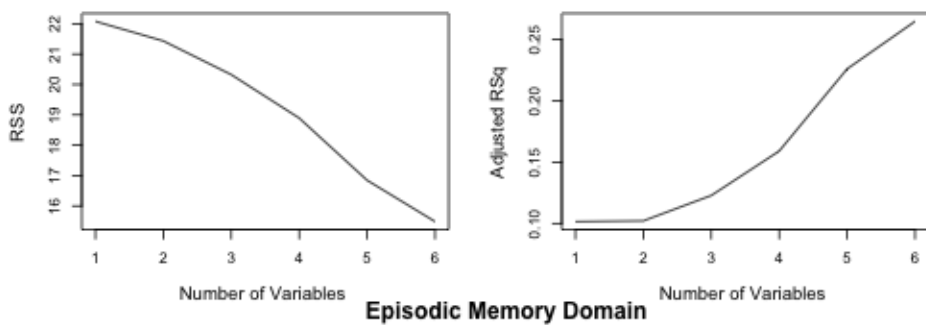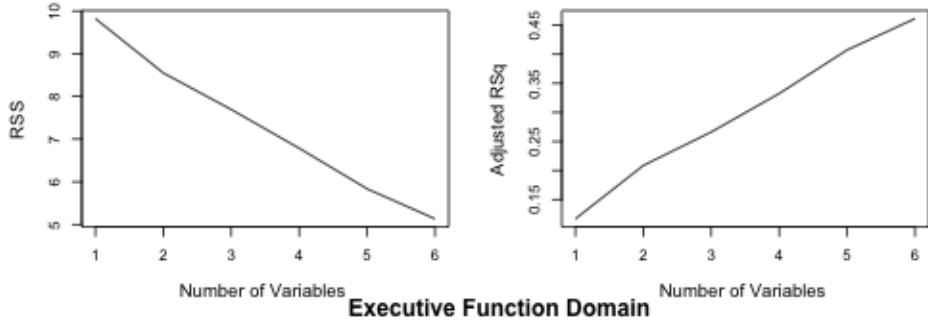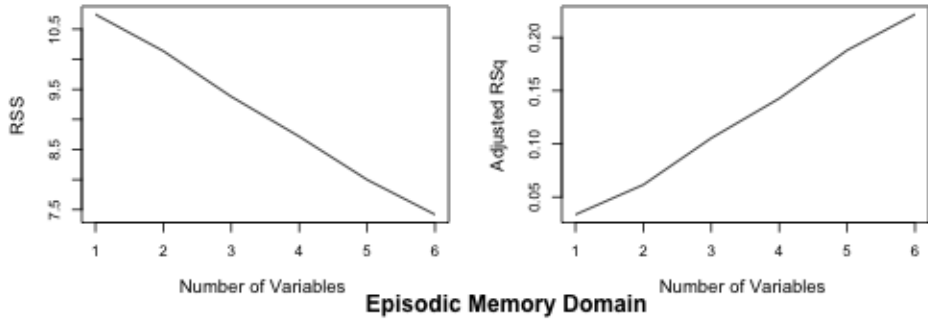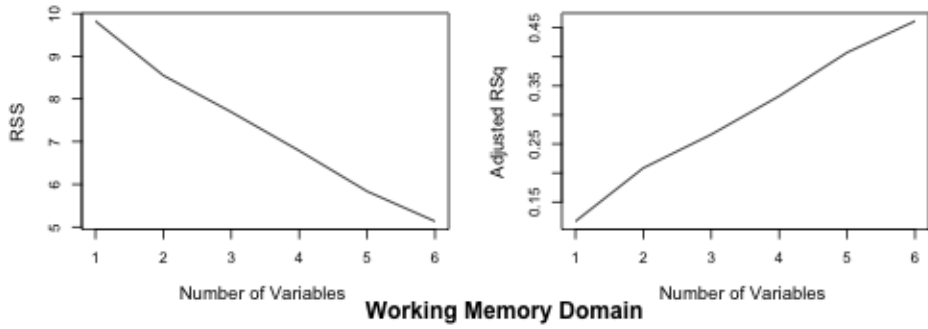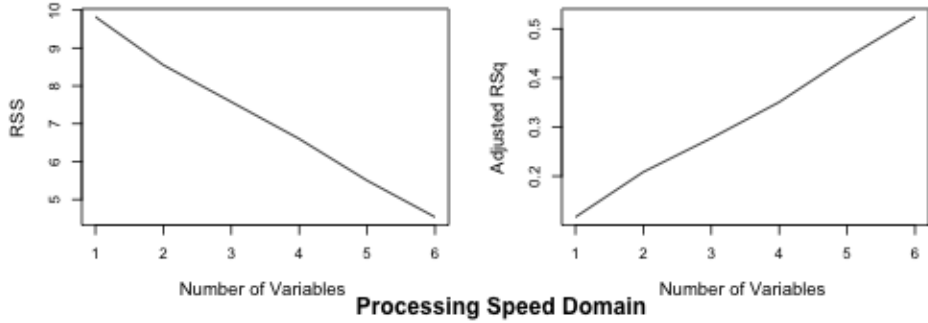
REFERENCES

1.   van den Heuvel, M.P. and H.E.H. Pol, *Exploring the brain network: A review on resting-state fMRI functional connectivity.* European Neuropsychopharmacology, 2010. **20**(8): p. 519-534.

2.   Mier, W. and D. Mier, *Advantages in functional imaging of the brain.* Frontiers in Human Neuroscience, 2015. **9**.

3.   Friston, K.J., et al., *Functional Connectivity - the Principal-Component Analysis of Large (Pet) Data Sets.* Journal of Cerebral Blood Flow and Metabolism, 1993. **13**(1): p. 5-14.

4.   Rykhlevskaia, E., G. Gratton, and M. Fabiani, *Combining structural and functional neuroimaging data for studying brain connectivity: A review.* Psychophysiology, 2008. **45**(2): p. 173-187.

5.   Andreasen, N.C., S. Paradiso, and D.S. O'Leary, *"Cognitive dysmetria" as an integrative theory of schizophrenia: A dysfunction in cortical subcortical-cerebellar circuitry?* Schizophrenia Bulletin, 1998. **24**(2): p. 203-218.

6.   Aertsen, A.M.H.J., et al., *Dynamics of Neuronal Firing Correlation - Modulation of Effective Connectivity.* Journal of Neurophysiology, 1989. **61**(5): p. 900-917.

7.   Lowe, M.J., et al., *Correlations in low-frequency BOLD fluctuations reflect cortico-cortical connections.* Neuroimage, 2000. **12**(5): p. 582-587.

8.   Greicius, M.D., et al., *Functional connectivity in the resting brain: A network analysis of the default mode hypothesis.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(1): p. 253-258.

9.   Biswal, B., et al., *Functional Connectivity in the Motor Cortex of Resting Human Brain Using Echo-Planar Mri.* Magnetic Resonance in Medicine, 1995. **34**(4): p. 537-541.

10.   Biswal, B.B., J. VanKylen, and J.S. Hyde, *Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps.* Nmr in Biomedicine, 1997. **10**(4-5): p. 165-170.

11.   Cordes, D., et al., *Mapping functionally related regions of brain with functional connectivity MR imaging.* American Journal of Neuroradiology, 2000. **21**(9): p. 1636-1644.

12.   Paus, T., et al., *Maturation of white matter in the human brain: A review of magnetic resonance studies.* Brain Research Bulletin, 2001. **54**(3): p. 255-266.

13.   Wang, K., et al., *Altered functional connectivity in early Alzheimer's disease: A resting-state fMRI study.* Human Brain Mapping, 2007. **28**(10): p. 967-978.

14.   Joo, S.H., H.K. Lim, and C.U. Lee, *Three Large-Scale Functional Brain Networks from Resting-State Functional MRI in Subjects with Different Levels of Cognitive Impairment.* Psychiatry Investigation, 2016. **13**(1): p. 1-7.

15.   Dickstein, D.P., et al., *Fronto-Temporal Spontaneous Resting State Functional*

Connectivity in Pediatric Bipolar Disorder. Biological Psychiatry, 2010. **68**(9): p. 839-846.

16.   Altinay, M.I., et al., *Differential resting-state functional connectivity of striatal subregions in bipolar depression and hypomania.* Brain connectivity, 2016. **6**(3): p. 255-265.

17.   van den Heuvel, M.P., et al., *Functionally Linked Resting-State Networks Reflect the Underlying Structural Connectivity Architecture of the Human Brain.* Human Brain Mapping, 2009. **30**(10): p. 3127-3141.

18.   Zhang, L.L., et al., *A spatio-temporal nonparametric Bayesian variable selection model of fMRI data for clustering correlated time courses.* Neuroimage, 2014. **95**: p. 162-175.

19.   Zhang, L.L., M. Guindani, and M. Vannucci, *Bayesian models for functional magnetic resonance imaging data analysis.* Wiley Interdisciplinary Reviews-Computational Statistics, 2015. **7**(1): p. 21-41.

20.   Yu, Z., et al., *Understanding the Impact of Stroke on Brain Motor Function: A Hierarchical Bayesian Approach.* Journal of the American Statistical Association, 2016. **111**(514): p. 549-563.

21.   Harrison, S.J., et al., *Large-scale Probabilistic Functional Modes from resting state fMRI.* Neuroimage, 2015. **109**: p. 217-231.

22.   Behrens, T.E.J., et al., *Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?* Neuroimage, 2007. **34**(1): p. 144-155.

23.   Behrens, T.E.J., et al., *Characterization and propagation of uncertainty in diffusion-weighted MR imaging.* Magnetic Resonance in Medicine, 2003. **50**(5): p. 1077-1088.

24.   Basser, P.J., J. Mattiello, and D. Lebihan, *Mr Diffusion Tensor Spectroscopy and Imaging.* Biophysical Journal, 1994. **66**(1): p. 259-267.

25.   Zhou, D.W., et al., *A Bayesian method with reparameterization for diffusion tensor imaging.* Medical Imaging 2008: Image Processing, Pts 1-3, 2008. **6914**.

26.   Olesen, P.J., et al., *Combined analysis of DTI and fMRI data reveals a joint maturation of white and grey matter in a fronto-parietal network.* Cognitive Brain Research, 2003. **18**(1): p. 48-57.

27.   Bowman, F.D., et al., *Determining functional connectivity using fMRI data with diffusion-based anatomical weighting.* Neuroimage, 2012. **62**(3): p. 1769-1779.

28.   Xue, W.Q., et al., *A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity.* Frontiers in Computational Neuroscience, 2015. **9**.

29.   Wieshmann, U.C., et al., *Combined functional magnetic resonance imaging and diffusion tensor imaging demonstrate widespread modified organisation in malformation of cortical development.* Journal of Neurology Neurosurgery and Psychiatry, 2001. **70**(4): p. 521-523.

30.	Werring, D.J., et al., *The structural and functional mechanisms of motor recovery: complementary use of diffusion tensor and functional magnetic resonance imaging in a traumatic injury of the internal capsule.* Journal of Neurology Neurosurgery and Psychiatry, 1998. **65**(6): p. 863-869.

31.	Morgan, V.L., et al., *Integrating Functional and Diffusion Magnetic Resonance Imaging for Analysis of Structure-Function Relationship in the Human Language Network.* Plos One, 2009. **4**(8).

32.	Honey, C.J., et al., *Predicting human resting-state functional connectivity from structural connectivity.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(6): p. 2035-2040.

33.	Greicius, M.D., et al., *Resting-State Functional Connectivity Reflects Structural Connectivity in the Default Mode Network.* Cerebral Cortex, 2009. **19**(1): p. 72-78.

34.	Kang, H., et al., *A bayesian double fusion model for resting-state brain connectivity using joint functional and structural data.* Brain connectivity, 2017. **7**(4): p. 219-227.

35.	Rasmussen, C.E. and C.K.I. Williams, *Gaussian Processes for Machine Learning.* Gaussian Processes for Machine Learning, 2005: p. 1-247.

36.	Salvatier, J., T.V. Wiecki, and C. Fonnesbeck, *Probabilistic programming in Python using PyMC3.* PeerJ Computer Science, 2016. **2**: p. e55.

37.	Hoffman, M.D. and A. Gelman, *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.* Journal of Machine Learning Research, 2014. **15**: p. 1593-1623.

38.	Duane, S., et al., *Hybrid Monte-Carlo.* Physics Letters B, 1987. **195**(2): p. 216-222.

39.	Bastien, F., et al., *Theano: new features and speed improvements.* arXiv preprint arXiv:1211.5590, 2012.

40.	Bergstra, J., et al. *Theano: A CPU and GPU math compiler in Python.* in *Proc. 9th Python in Science Conf.* 2010.

41.	Geweke, J., *Bayesian Statistics 4 ed JM Bernardo et al.* 1992, Oxford: Oxford Univ. Press.

42.	Gelman, A., et al., *Bayesian data analysis.* 1995: Chapman and Hall/CRC.

43.	Gelman, A. and D.B. Rubin, *Inference from iterative simulation using multiple sequences.* Statistical science, 1992: p. 457-472.

44.	Gelman, A., *Bayesian data analysis: what it is and what it is not.* 2014.

45.	Wikipedia, *Major Depression Disorder.*

46.	Mulders, P.C., et al., *Resting-state functional connectivity in major depressive disorder: A review.* Neuroscience and Biobehavioral Reviews, 2015. **56**: p. 330-344.

47.	Albert, K.M., et al., *Cognitive performance in antidepressant-free recurrent major depressive disorder.* Depression and anxiety, 2018.