Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Joshua Lukemire 4/24/2021

DISSERTATION COMMITTEE APPROVAL SHEET

Statistical Methods for Brain Network Estimation

By

Joshua Lukemire
Doctor of Philosophy

Biostatistics

Ying Guo
Advisor

Suprateek Kundu
Advisor

Jennifer Stevens
Committee Member

Hao Wu
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Statistical Methods for Brain Network Estimation

By

Joshua Lukemire
M.S., Emory University, 2019

Advisors: Ying Guo, Ph.D. and Suprateek Kundu Ph.D.

Committee

Ying Guo, Ph.D

Suprateek Kundu, Ph.D

Hao Wu, Ph.D

Jennifer Stevens, Ph.D

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2021

Abstract

Statistical Methods for Brain Network Estimation

Functional Magnetic Resonance Imaging (fMRI) allows researchers to study neural activity by measuring changes in blood oxygen levels throughout the brain at rest or in response to experimental tasks/stimuli. In recent years, there has been great interest in using fMRI data to study brain networks. In this dissertation, we develop new statistical methods for providing reliable and robust estimation and characterization of brain networks across different cognitive states, subpopulations and imaging study designs.

In the first topic, we propose the Bayesian Joint Network Learning (BJNL) approach to joint brain network estimation that pools information across groups to estimate group-specific brain networks under a graph theoretic approach. BJNL uses information from multiple groups to estimate the probability of edges between nodes without forcing similarity in the edge strengths. The BJNL approach is shown to outperform other individual and joint estimation techniques in simulations. The technique is then applied to a Stroop fMRI data set.

Next, we propose a Sparse Bayesian Independent Component Analysis (SparseBayes ICA) for reliable estimation of individual differences in brain networks. We model the population-level ICA source signals for brain networks using a Dirichlet process mixture of Gaussians. To reliably capture individual differences on brain networks, we propose sparse estimation of the covariate effects in a hierarchical ICA model via a horseshoe prior. Through extensive simulation studies, we show our approach has improved performance in detecting covariate effects in comparison with the current group ICA methods. We then use it to perform an ICA decomposition of a motivating Zen meditation resting-state study.

In our third topic, we introduce a general framework of repeated measures Sparse Bayesian ICA (RM-SparseBayes ICA). This method provides a rigorous and much needed tool for investigating brain networks in imaging studies with complex study designs including longitudinal and/or multi-center studies. Through simulations, we show that the proposed method has considerably improved performance as compared to other potential approaches. We apply the RM-SparseBayes ICA to investigate brain network changes using the longitudinal multi-center Alzheimer's disease data from the ADNI2 study.

Statistical Methods for Brain Network Estimation

By

Joshua Lukemire
M.S., Emory University, 2019

Advisors: Ying Guo, Ph.D. and Suprateek Kundu Ph.D.

Committee

Ying Guo, Ph.D

Suprateek Kundu, Ph.D

Hao Wu, Ph.D

Jennifer Stevens, Ph.D

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2021

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) images the brain (or other tissues) by emitting a strong magnetic field that lines up the protons in the tissues. After lining up the protons, the MR scanner emits a radio frequency (RF) pulse that knocks the protons out of alignment. As the protons return to their natural positions (a process known as relaxation), they emit their own RF signals, which are then measured by the scanner to image the brain (Weishaupt et al., 2008).

There are two forms of relaxation measured by the MR scanner: T1 weighted and T2 weighted (Weishaupt et al., 2008). T1 weighted relaxation, or longitudinal relaxation, shows up as very bright white pixels for white matter, varying degrees of grey pixels in grey matter, and dark pixels for cerebrospinal fluid. T2 weighted relaxation or transverse relaxation has roughly the opposite pattern (Pooley, 2005).

Functional Magnetic Resonance Imaging (fMRI) is one of the most popular tools for studying the human brain due to its combination of high spatial resolution and relatively high temporal resolution (Hennig et al., 2003). The technique was first introduced in 1991 (Belliveau et al., 1991), following the common adaptation of structural MRI in the early 80s (Bandettini, 2012).

fMRI measures a quantity called the blood-oxygenated-level-dependent (BOLD) signal (Buxton, 2013). When a region of the brain is being used, the oxygen in the blood flowing through that region is used up and the cardiovascular system responds by increasing blood flow to that region. Crucially, the cardiovascular system sends *more* oxygenated blood than is required (Fox and Raichle, 1986). Thus, fMRI is able to detect regions of the brain that exhibit increased oxygen levels (higher BOLD signal) over the course of the scan, which is associated with increased neural activation (Buxton, 2013). Thus, fMRI uses the BOLD signal as a proxy for neural activity.

1.2 Brain Networks

The early years of fMRI primarily saw interest in identifying brain regions that were correlated with task performance, known as *localization* studies. However, in recent years the interest of the neuroimaging community has largely shifted to understanding how the brain behaves as a coherent entity via the study of functional brain networks. These functional brain networks are most commonly studied during resting-state scans in which subjects in the scanner are imaged while not performing any specific task.

The first study of resting-state functional connectivity revealed that motor regions of the brain were correlated at rest (Biswal et al., 1995). Many studies followed, and the study of resting-state functional connectivity has resulted in the discovery of a number of highly reproducible brain networks. The most widely known of the networks is the default mode network (DMN), which is hypothesized to be involved in self-referential thought, and has been identified as a potential biomarker for Alzheimer's disease (Greicius et al., 2004; Jones et al., 2011).

There are numerous other brain functional networks which are highly reproducible across subjects and sessions. In a large-scale study involving almost 30,000 subjects and over a thousand experimental conditions, Smith et al. (2009) identified ten functional networks that could be easily matched across multiple datasets. These networks have clear purposes based on previous experimental activation studies and include three visual networks, the DMN, and several networks related to cognitive function. We use several of these networks in the work that follows. Some of the key networks involve the executive control network, which is associated with cognition and action inhibition, and the left and right frontoparietal networks, which are associated with language functions.

There are two main approaches to studying brain networks. The first, *functional connectivity*, involves the study of how the observed BOLD time courses in anatomically distinct brain regions change together over time (Van Den Heuvel and Pol, 2010). This approach is generally concerned with the connections between regions. The second, related, approach

involves investigating the spatial network organization of the brain. In what follows, we go into further detail on both approaches and discuss some common statistical approaches to each.

1.2.1 Functional Connectivity

In functional connectivity studies researchers examine the time courses of different regions of the brain during an fMRI scan. Regions exhibiting similar time courses are said to be "functionally connected," and collections of functionally connected regions form brain networks. In functional connectivity studies these aspect of interest is usually the strength of the connection between regions. Research questions of interest often involve how these connection strengths differ in different clinical populations. In the next sections we discuss some of the commonly used statistical approaches to studying functional connectivity in fMRI studies.

1.2.1.1 Seed Voxel Analyses

Traditionally functional connectivity has been studied via simple correlation analyses. In many of the earliest studies investigators would chose a region of the brain, known as a seed region, that was thought to be associated with some brain function of interest. They would then correlate the rest of the regional time courses in the brain with this seed region in order to identify brain regions that exhibited temporally coherent activity patterns (Biswal et al., 1995). Figure 1.1 displays an example of this kind of procedure.

The seed voxel approach has several drawbacks. First, it requires the researcher to know which region of the brain they wish to use as the seed. Second, the results of a seed analysis can vary significantly depending on the selected region, even within similar functional networks. Even a pair of nearby voxels can yield very different connectivity estimates in some cases (Cole, Smith and Beckmann, 2010). Finally, seed analyses only provide information about the relationship between the rest of the brain and the selected

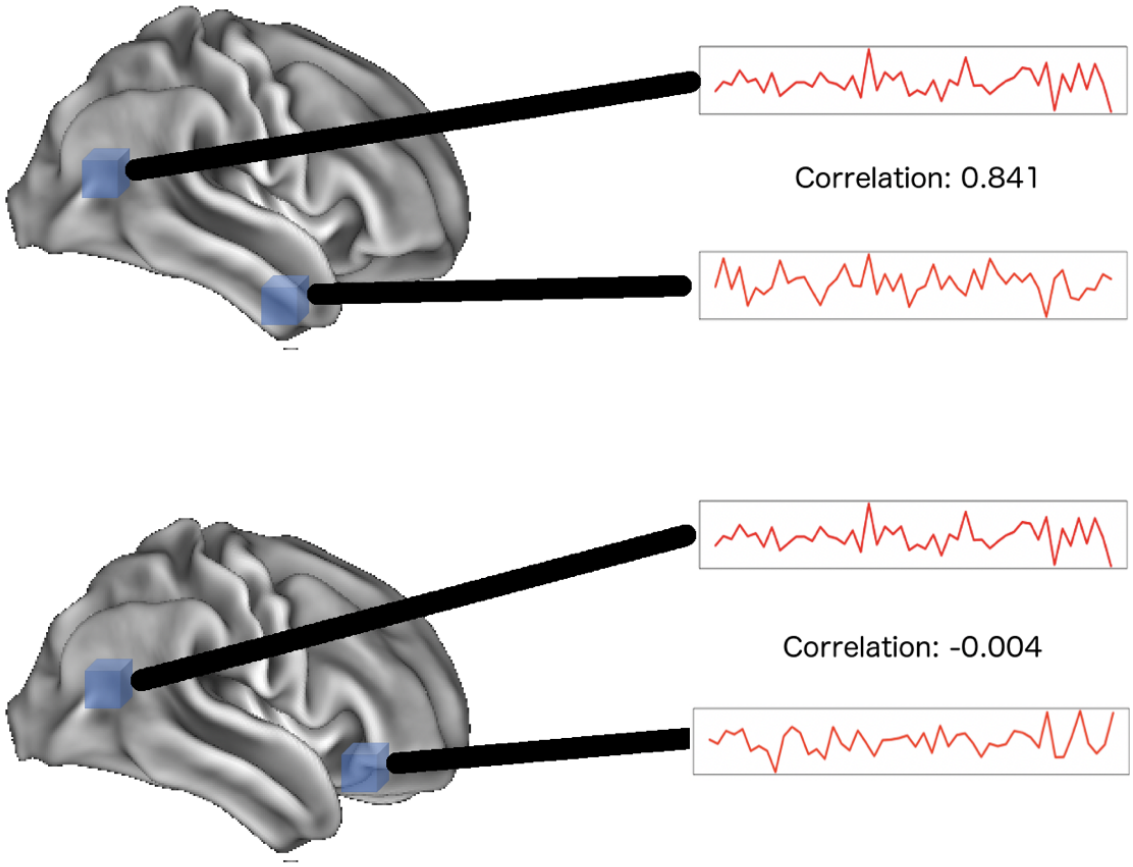
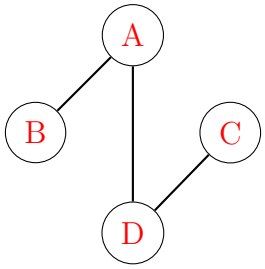


Figure 1.1: An illustration of the seed voxel approach. The top row shows the time course of a seed voxel and another voxel, and they appear to have a strong functional connection. The bottom row shows the time course of the same seed voxel, but with another voxel that does not appear to be functionally connected to the seed region.

Figure 1.2: An example graph with $V = \{A, B, C, D\}$ and $E = \{(A, B), (A, D), (C, D)\}$.



seed voxel/region. In many cases researchers are interested in multiple networks or how the brain behaves as a whole.

1.2.1.2 Gaussian Graphical Models

A popular alternative to the seed voxel approach is to do network modeling by viewing the brain from a graph theoretical perspective. Under these approach, functional connectivity is assessed by examining the connections (edge) between distinct brain regions (nodes). This approach has the advantage of enabling graph metrics use to describe different nodes or even entire graphs, as will be discussed in Topic 1.

A graph $\mathcal{G} = \{V, E\}$ is a set of vertices, or nodes (V), and the edges between those nodes (E). An edge between two nodes indicates that there is direct information transmission between them. In the context of a brain network, this is a direct functional connection. On the other hand, if there is not an edge between two nodes, then the only way they can share information is via a third (or more) intermediate node(s). Figure 1.2 provides an illustration of this. The graph $\mathcal{G}$ in Figure 1.2 contains four nodes, $A - D$, and three edges. Notice that while node A cannot communicate directly with node C, it can still reach it via node D.

Gaussian Graphical models provide a convenient statistical approach to studying graphs such as the one above. Recall that a $P-$dimensional random variable $\mathbf{Y}$ that follows a multivariate Gaussian distribution has density function

$$f(\mathbf{Y}; \boldsymbol{\mu}, \boldsymbol{\Omega}^{-1}) = (2\pi)^{-\frac{P}{2}} |\boldsymbol{\Omega}| \exp\{-\frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu})^T \boldsymbol{\Omega} (\mathbf{Y} - \boldsymbol{\mu})\}. \tag{1.1}$$

Here, $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ is the inverse covariance matrix, also known as the precision matrix. A key property of the multivariate Gaussian distribution is that the conditional distribution of a subset of its dimensions can be expressed as another multivariate Gaussian distribution. Define $\mathbf{Y}_1$ to be the P^* dimensional subvector of $\mathbf{Y}$, where $P^* \leq P$ and let $\mathbf{Y}_2$ be the P^- dimensional subvector containing the remaining elements of $\mathbf{Y}$, where $P^* + P^- = P$. The conditional distribution of $\mathbf{Y}_1$ given $\mathbf{Y}_2$ is then:

$$f(\mathbf{Y}_1|\mathbf{Y}_2; \boldsymbol{\mu}, \mathbf{\Sigma}) = N_{P^*}(\boldsymbol{\mu}_{1|2}, \mathbf{\Sigma}_{1|2}), \tag{1.2}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{1|2} &= \boldsymbol{\mu}_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{Y}_2 - \boldsymbol{\mu}_2) \\ \mathbf{\Sigma}_{1|2} &= \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}, \end{aligned} \tag{1.3}$$

where $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$ are defined as the $\mathbf{Y}_1$, $\mathbf{Y}_2$ above and $\mathbf{\Sigma}_{11}$ is the $P^* \times P^*$ sub-matrix of elements of the covariance matrix of $\mathbf{Y}$ corresponding to the sub-vector $\mathbf{Y}_1$. $\mathbf{\Sigma}_{12}$ and $\mathbf{\Sigma}_{22}$ are defined similarly. From this, it can be shown that one dimension of $\mathbf{Y}$, say i, is independent of another dimension, say j, conditioned on the remaining dimensions if and only if $\mathbf{\Omega}_{ij} = 0$. This relationship can be used to describe a graph. Specifically, each dimension of $\mathbf{Y}$ corresponds to a node. If an off-diagonal element of $\mathbf{\Omega}$ is zero, then there is no edge between the corresponding nodes.

1.2.2 Spatial Network Organization

The functional connectivity approaches described above are primarily concerned with examining the connections between different brain regions. The other general approach to studying brain networks focuses the spatial extents of the regions involved in these connected regions. Specifically, while the functional connectivity approaches examine the connections

between the voxels, the spatial network organization approaches examine where in the brain these voxels are located. Additionally, spatial extent studies are concerned with the intensities *within* the regions during the fMRI scan.

The collections of these spatial regions form the brain networks of interest, and brain network studies often concern how different clinical or demographic characteristics affect the extent or intensity of the activation in these regions. For example, Greicius et al. (2007) identified abnormalities in the default mode network associated with major depressive disorder.

1.2.2.1 Independent Component Analysis

Independent Component Analysis (ICA) is the most popular approach for studying the spatial network organization of the brain. ICA models decompose the observed fMRI time courses into a set of spatial components (brain networks) that are combined together via a mixing matrix to generated the observed data. The single subject noisy ICA model is given by:

$$\mathbf{Y} = \mathbf{AS} + \mathbf{E}, \tag{1.4}$$

where $\mathbf{Y}$ is a matrix with columns containing the observed (potentially preprocessed) time courses for each voxel, $\mathbf{A}$ is a mixing matrix where each column contains the time course of the corresponding brain network, $\mathbf{S}$ is a matrix with each row containing the network-specific activation for each voxel, and $\mathbf{E}$ is an error term. This was the first type of ICA approach applied to fMRI studies of a single subject (McKeown, Jung, Makeig, Brown, Kindermann, Lee and Sejnowski, 1998).

In most settings investigators will be interested in using ICA to study the brain network across an entire group or groups of subjects. In this case, a group ICA approach will be required in order to enable comparison of brain networks across subjects. The most popular group ICA approach is temporal concatenation group ICA, or TC-GICA (Calhoun et al.,

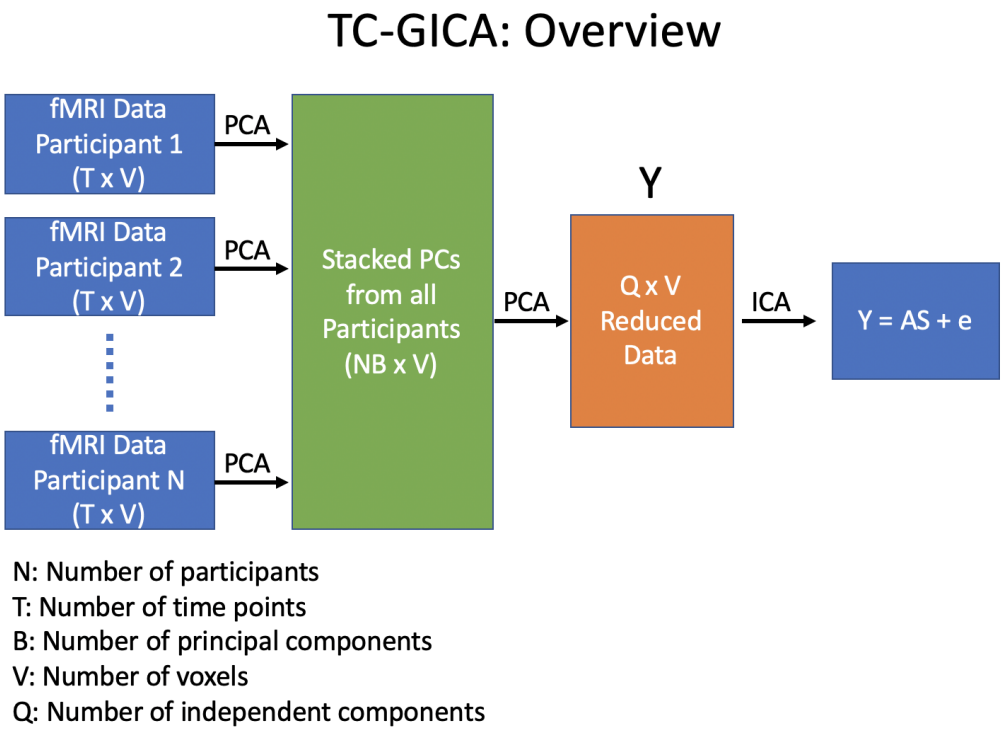


Figure 1.3: Illustration of the temporal concatenation group ICA approach.

2001). TC-GICA works by stacking each subjects' data along the time doing to create a single, large data set. Standard single subject ICA methods can then be applied to the concatenated data, and back-reconstruction can be performed to obtain subject-specific brain networks (see Figure 1.3 for an illustration).

An alternative approach has been developed in recent years called hierarchical covariate-adjusted ICA, or hc-ICA (Shi and Guo, 2016). hc-ICA utilizes a hierarchical statistical model to estimate a common set of brain networks that is modulated by individual covariate effects. These hc-ICA approaches have the advantage of providing a formal statistical framework for testing covariate effects. Further details on both the TC-GICA and hc-ICA models will be provided in Topics 2 and 3.

1.3 Bayesian Nonparmetric Approaches

The Dirichlet Process (DP) (Ferguson, 1973) is by far the most popular tool for Bayesian nonparametrics. The DP has two parameters: a scalar term $M > 0$, known as the precision, and a measure G_0, which is known as the center measure. Suppose that the center measure has support over the region $\mathcal{B}$, and that $B_1, \ldots, B_K$ form a disjoint partition of $\mathcal{B}$. The defining property of a draw G from the DP is that it is a probability distribution on $\mathcal{B}$ such that the joint distribution of $G(B_1), \ldots, G(B_K)$ is distributed Dirichlet$(MG_0(B_1), \ldots, MG_0(B_K))$. Note that G is a discrete measure, and can be generated through a stick breaking process (Sethuraman, 1994). To see this generation, first note that due to the discreteness of distributions generated from a Dirichlet Process we can write

$$G = \sum_{h=1}^{\infty} w_h \delta_{s_h}, \tag{1.5}$$

where the w_h are weights and the δ_{s_h} are point masses at values in the support of G_0. The stick-breaking procedure for generating samples from the DP uses the following steps, one h at a time.

1. Draw $\nu_h \sim Beta(1, M)$
2. Set $w_h = \nu_h \prod_{h'=1}^{h-1}(1 - w_{h'})$
3. Draw a new location $\delta_{s_h} \sim G_0$

Note that the sum in Equation (1.5) is infinite. In practice, some truncation must be performed in order to sample from this distribution (i.e. we must stop at some h).

One drawback (in some circumstances) of the discreteness of distributions drawn from the DP is that it makes the DP inappropriate as a prior for continuous distributions. To solve this, Ferguson (1983) proposed the Dirichlet Process Mixture (DPM). In a DPM, a continuous kernel $(f_\theta(y_i))$ is mixed using a DP random probability measure in order to produce a continuous distribution i.e.

$$f_G(y_i) = \int f_\theta(y_i) dG(\theta). \tag{1.6}$$

This extension allows the DP to be used for continuous distributions. One interesting and useful aspect of this technique is clustering; because samples from the DP are discrete, there is a non-zero probability of parameters drawn from the constructed prior being the same. Thus, there will be a limited number of unique values for the parameters $\boldsymbol{\theta}$. We will see later that this can lead to large computational savings when the clustering behavior is expected for the study question. Specifically, instead of having to update a parameter vector for each observation in the study, we only have to update the parameter vector for a finite number of clusters.

1.4 Research Topics

Chapter 2 introduces Bayesian Joint Network Learning, a Bayesian technique for the joint estimation of multiple brain networks under a graph theoretic approach. This method models the probability of individual edges as a function of shared and differential effects. We use information across all groups to estimate the shared effects, and the differential effects give us information about where different groups have a different graph structure. We use simulations to show that this technique is superior to other estimation approaches for this problem. We apply the method to a Stroop task fMRI data set.

Chapter 3 introduces Sparse Bayes Independent Component Analysis (SparseBayes ICA). SparseBayes ICA is a Bayesian extension of hierarchical covariate-adjusted ICA methods. Our proposed method introduces a novel approach to modeling the distribution of the independent component source signals. Additionally, it incorporates assumptions about covariate effect sparsity directly into the ICA decomposition. Through simulations, we show that this proposed method has a better ability to find significant covariate effects, as compared to current state-of-the-art approaches. The method is then applied to a Zen meditation data

set.

Chapter 4 extends the Sparse Bayes Independent Component Analysis from Chapter 2 to a repeated measures setting. This novel approach is applicable to data collected from multiple centers, repeated longitudinal scans on the same subjects, or both. Though simulations, we show that our proposed approach is superior to a naive dual regression and data harmonization approaches. Finally, we apply the method to an Alzheimer's disease data set with three measurements for each subject over a period of two years.

Chapter 2

Bayesian Joint Network Learning

Figure 2.1: Illustration of the brain's nodes under a graph theoretic approach.

2.1 Introduction

2.1.1 Graphical Models in fMRI

Graphical models provide an highly interpretable approach to the study of brain functional connectivity. Under the graph theoretic approach, the brain is viewed as a set of nodes and edges. Here, nodes are spatial regions of the brain (see Figure 2.1 for an example), and edges are the set of functional connections between these regions. The nodes are typically defined using a brain anatomical atlas. Some popular atlases include the automated anatomical labeling atlas (AAL) (Tzourio-Mazoyer et al., 2002) and the Power 264 node system (Power et al., 2011). Once a node atlas has been selected, it remains to estimate the edges. There are many techniques for this, as will be discussed in Section 2.1.2.

Once the brain graph has been obtained, there are a wealth of metrics that can be used to describe aspects of information transmission in the brain at different scales (He and Evans, 2010). At the macro scale, measures such as global efficiency and characteristic path length give a picture of how many steps are required for brain regions to communicate. At the more local node level, metrics such as the local efficiency provide information about the ability of *a specific* node to communicate with different regions of the brain. Finally, at the very fine edge level we have information about the strengths of individual connections in the brain.

Graph theoretical approaches to the study of functional connectivity have been very popular in recent years (He and Evans, 2010). These approaches have been especially valuable to the study of neurodegenerative disorders (Stam, 2014). There is significant interest in using brain network metrics to classify and predict subjects based on the topological features of their connections (Bullmore and Sporns, 2009). For example, global efficiency has been used to successfully classify patients as healthy controls or Alzheimer's disease patients (Supekar et al., 2008). Similarly, Bassett et al. (2012) used connected graph components to classify schizophrenia patients with a high degree of sensitivity and specificity.

2.1.2 Graph Estimation

The principal difficulty in applying a graph theoretical approach to the study of brain functional connectivity is identifying the edge set and quantifying the strength of the edges. The most basic approach to identifying the edge set is to calculate whole brain Pearson correlation, however this approach has a number of drawbacks. The primary issue with using a standard correlation analysis is that it does not account for the influence of intermediate nodes; a Pearson correlation analysis will not be able to distinguish a direct connection between two nodes from an indirect connection via an intermediary. Naturally this leads to the estimation of graphs that are too dense.

In light of this difficulty, partial correlation analysis has become a more popular approach to estimating the edge set. Recall that if the klth element of $\mathbf{\Omega}$ is 0, then the nodes k and l are conditionally independent (see Section 1.2.1.2 for a review). This is directly related to the partial correlation between nodes k and l, which can be calculated from the precision matrix as:

$$\rho_{kl} = -\frac{\omega_{kl}}{\sqrt{\omega_{kk}\omega_{ll}}}. \tag{2.1}$$

Clearly when $\omega_{kl} = 0$ the partial correlation is 0. This motivates the estimation of a sparse

precision matrix $\mathbf{\Omega}$ as a technique for estimating the edge set. This approach has been shown to be very effective at identifying true connections between two nodes, as opposed to a connection that appears to exist due to another, intermediate, node (Smith et al., 2011).

Estimation of a sparse precision matrix is key to the above partial correlation approach. Because the graph interpretation requires exact zeros, these problems naturally lend themselves to frequentist penalized approaches. The most popular graph estimation technique is the graphical lasso (Friedman et al., 2008). The graphical lasso estimates the precision matrix for a given group or individual as:

$$\mathbf{\Omega} = \arg\min_{\mathbf{\Omega}} \left\{ tr(\mathbf{S\Omega}) - \log\det(\mathbf{\Omega}) + \lambda \sum_{k \neq l} |\omega_{kl}| \right\},$$

where λ is a penalty that can be chosen using cross-validation or a measure such as AIC. The penalty on the magnitude of the ω_{kl} terms encourages exact zeros, and the larger λ is, the more exact zeros that will be obtained.

Penalized techniques such as the graphical lasso have become very popular in the fMRI graph theoretical literature (Rosa et al., 2015; Schmittmann et al., 2015). Despite their popularity, they have a significant drawback: they estimate the graph for one group or experimental condition at a time. Many research questions involve more than one group of subjects, such as our Stroop task example. In such cases, the focus is often identifying similarities and differences between groups. Estimating the graphs separately for each group/condition ignores the likely presence of a large degree of similarity between the two graphs. For problems like this, joint estimation - estimating the graphs for all groups of interest together - is preferred, as it allows us to leverage information from each group of subjects when estimating the graphs for the other subjects.

There are numerous frequentist extensions to joint estimation of multiple graphical models (Guo et al., 2011; Danaher et al., 2014). These approaches pool information across groups by smoothing over the edge strengths. This enforces shared edges by penalizing differences in the edge strengths between two groups. For example, the penalty term:

$$P(\boldsymbol{\Omega}) = \lambda_1 \sum_{j \neq k} \sum_{g=1}^{G} |\omega_{g,jk}| + \lambda_2 \sum_{j \neq k} \sum_{g \neq g'} |\omega_{g,jk} - \omega_{g',jk}|$$

pools information across groups by penalizing the distance $|\omega_{g,jk} - \omega_{g',jk}|$, with new tuning parameter λ_2. Such procedures are convenient from a computational perspective. However this approach adds the assumption that the shared edges between graphs have similar edge strengths. That is, by smoothing over the edge strengths, these techniques force shared structure in the graph by encouraging the edge strengths to be similar. This assumption is rather restrictive, and might be quite unrealistic in practice. Another downside of these penalized approaches is that they only provide point estimates for the precision matrices; they do not provide any measures of uncertainty. This makes quantifying our uncertainty about the values of the matrix difficult, requiring additional computation stages such as bootstrap or permutation tests.

As an alternative to the frequentist penalized approaches, there are numerous Bayesian approaches to sparse precision matrix estimation (Wang, 2012; Li et al., 2019). These approaches place a prior on the off-diagonal elements of the precision matrix in order to encourage sparsity. The priors chosen generally are based on either a spike-and-slab approach (Yu and Dauwels, 2016) or a continuous shrinkage approach (Carvalho et al., 2010; Polson and Scott, 2010; Piironen et al., 2017). The spike and slab approach will be discussed in more detail shortly. In short, spike and slab approaches assume that the observed data comes from two distinct groups, the groups are modeled under a diffuse distribution (the slab) and a distribution that is tight around zero (the spike). The contribution of each distribution is determined by a weight term. On the other hand, the continuous shrinkage approaches are "one-group" models, that generally use a normal prior for the model parameters, where the precision is encouraged to be large/small for each parameter depending on the information provided in the data. While the two-group model under the spike and slab is likely more realistic, the one-group models can often achieve similar/better performance because they have an easier parameter space to explore.

There are several additional difficulties introduced by using a Bayesian approach. First, these approaches do not give exact zeros, and thus some thresholding must be performed in order to obtain exact zeros. Common approaches involve thresholding based on the magnitude of the precision matrix elements, or on the posterior probability of an edge for models that calculate edge probabilities (Wang, 2012). Second, the computation is generally more difficult under the Bayesian approaches. These approaches generally require the use of Markov chain Monte Carlo (MCMC) techniques to carry out posterior inference. Despite the additional difficulties encountered when using Bayesian approaches, they provide some significant advantages over the frequentist techniques. First, they provide measures of uncertainty in the form of credible intervals. This allows us to quantify our certainty about the presence or absence of an edge. Second, they are extremely flexible; we can select priors with varying degrees of strength in our prior belief.

Despite the advantages of using Bayesian approaches to sparse precision matrix estimation, the above Bayesian approaches have generally only been developed for estimation of a single graph. An exception is the recent work by Peterson et al. (2015) using a Markov random field approach to jointly estimate multiple networks. However, this approach was intended for multiple protein-protein interaction networks, and generally only appears to be applicable to networks with a small number of nodes. Clearly there is a need for Bayesian approaches to joint estimation of multiple networks.

In the first topic, we propose a novel Bayesian approach to joint estimation of multiple brain networks: Bayesian Joint Network Learning (BJNL). Our approach pools information across experimental conditions or groups in order to estimate the probability of an edge between connections, instead of the strengths of the connections themselves. We do this by modeling the probability of an edge using a function of shared and differential components. Specifically, we use a logistic regression model to describe the probability of an edge between any two nodes. The regression coefficients in this model are clustered using a Dirichlet process mixture (DPM) of Gaussians approach (Müller et al., 1996). The edge probabilities

allow us to threshold the find edge set using the posterior probability of an edge. They also enable interpretation of group effects on the odds of an edge. The off-diagonal precision elements are modeled under a spike and slab prior, where the weights are given by the edge probabilities. Posterior computation is carried out via MCMC using Gibbs sampling.

2.2 Methods

2.2.1 Stroop Task Data

In this work we consider an fMRI Stroop Task study data (Stroop, 1935; Khachouf et al., 2017). All MRI data were collected from the N.O.C.S.A.E Hospital in Baggiovara (MO), Italy, using a 3T Philips Achieva scanner. Data were collected from forty-five right-handed college students (mean age = 21.9 years, sd = 2.2). The imaging session consisted of the collection of 6 echo-planar imaging (EPI) runs (112 volumes each, TR=2.5s, 25 axial slice, $3 \times 3 \times 3$ mm voxels) and a T1-weighted high-resolution volume (180 sagittal slices, 1mm isotropic voxels) for anatomical reference.

Figure 2.2 provides an illustration of the 4-color version of the Stroop task the subjects performed in the scanner (Gianaros et al., 2005). For each trial in the task, a word corresponding to a color appears on the screen in a colored font. The subjects are instructed to press a button corresponding to the color of the word that is being presented on the screen. There are two types of trials - congruent and incongruent. In the congruent trials, the color of the font matches the text of the word on the screen. In the incongruent trials, the color of the word does not match. This task has been found to elicit a slower response time in the incongruent trials as compared to the congruent trials (Stroop, 1935).

Trials were presented in task blocks. Each block was 30 seconds and contained a total of 12 trials, 6 congruent and 6 incongruent. The trials were administered in a (pseudo) random order, and there were 2.5 seconds between trials. Each task block was followed by a 25 second resting-state block in which subjects were instructed to focus on a crosshair at

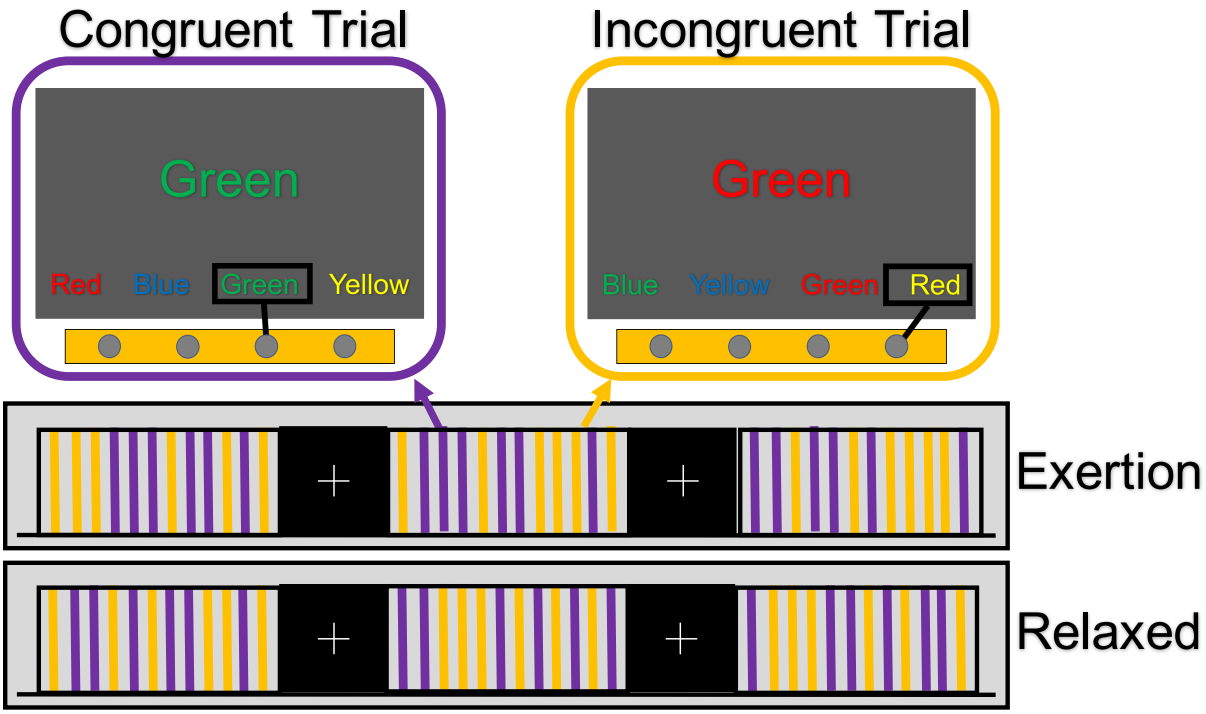


Figure 2.2: An illustration of the Stroop task involving task blocks of congruent and incongruent trials, indicated by purple bars and yellow bars respectively, and fixation blocks denoted by a centrally fixated cross. The purple and yellow bars are expanded into two boxes, and the correct button presses are indicated with a rectangle within each box.

the center of the viewing screen. A total of 6 runs were collected per subject. Each run had 4 task blocks and 5 rest blocks. Subjects were instructed to perform the odd numbered blocks "with maximum exertion" and the even numbered blocks "as relaxed as possible." Some subjects received the instructions in the opposite order, in order to check for possible order-effects. We are interested in the following two research questions:

1. How does the brain network differ between task performance and resting-state?

2. Within the task performance data, how does approaching the task "with maximum exertion" change the structure and function of the brain network as compared to approach the task from a relaxed state?

Note that both questions involves estimation of two graphs - one for each experimental condition. As discussed above, there is likely significant overlap in the graph structure

between the two conditions, and thus it is natural to want to pool information across groups to aid the estimation procedure. In what follows, we describe our proposed BJNL approach for this type of joint-estimation problem.

2.2.2 Bayesian Joint Network Learning

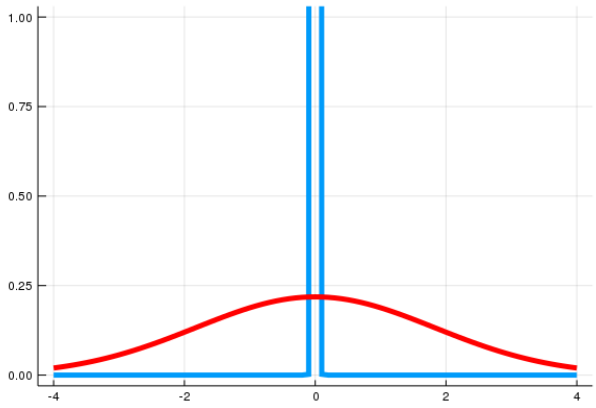
Consider an experiment with G groups, where the groups may correspond to different study populations or different experimental conditions, such as task performance versus rest. Our BJNL approach models the pre-whitened fMRI time courses for the gth cohort as $\mathbf{y}_{it}(g) \sim N_p(\mathbf{0}, \mathbf{\Omega}_g^{-1})$, $i = 1, \ldots, n, t = 1, \ldots, T_{ig}, g = 1, \ldots, G$. Here $N_p(\cdot; \mathbf{0}, \mathbf{\Sigma})$ denotes a p-dimensional normal distribution. As discussed above, in order to ensure sparsity in the precision matrices, we place the following prior on the precision matrix:

$$\pi(\mathbf{\Omega}_g) = C_g^{-1} \prod_{k=1}^{p} E(\omega_{g,kk}; \frac{\alpha}{2}) \Big\{ \prod_{k<l} w_{g,kl} N(\omega_{g,kl}; 0, \tau_{g,kl}^{-1}) + (1 - w_{g,kl}) DE(\omega_{g,kl}; \lambda_0) \Big\} I(\mathbf{\Omega}_g \in M^+), \tag{2.2}$$

where $\omega_{g,kl}$ is the strength of the edge (precision off diagonal) between nodes k and l in cohort g and $w_{g,kl}$ is the corresponding weight or edge probability. M^+ denotes the space of all positive definite matrices, which is required because we are modeling a precision matrix and thus it must be invertable. $E(\alpha)$ denotes an exponential distribution with scale parameter α^{-1}. $DE(\lambda)$ denotes a double exponential distribution with scale parameter λ^{-1}. Finally, $I(\cdot)$ denotes the indicator function and C_g is the intractable normalizing constant, which can easily be shown to be finite (Wang, 2012).

We note that the weighted mixture of the normal and double exponential distributions as a prior for the precision off-diagonals corresponds to the spike and slab prior (George and McCulloch, 1993) and thus the entire prior in Equation 2.2 is known as the *spike and slab Bayesian graphical lasso*. The $DE(\lambda)$ distribution is the spike, which has a large value for the hyperparameter λ_0 corresponding to a sharp peak at 0. The $N(\omega_{g,kl}; 0, \tau_{g,kl}^{-1})$ distribution is

Figure 2.3: Illustration of the spike and slab prior. The blue line corresponds to the double exponential distribution (spike) and the red line corresponds to the normal distribution (slab).



the slab component, with small values of the precision $\tau_{g,kl}^{-1}$ resulting in a diffuse distribution that accommodates non-zero values for locations in the brain with a connection. Figure 2.3 provides an illustration of the spike and slab prior. The weight terms $w_{g,kl}$ control how much each distribution contributes to the prior, with large $w_{g,kl}$ being indicative of strong prior belief that an edge is present between nodes k and l in cohort g.

As discussed in the introduction, our goal is to pool information across cohorts to estimate the brain networks without enforcing similarity in the edges strengths. We do this by borrowing information across cohorts to estimate the weights in the spike and slab prior. We model the weights using a generalized linear model with a parametric link function that includes a shared intercept term and differential cohort effects:

$$\begin{aligned} w_{g,kl} &= h(\eta_{0,kl}, \eta_{g,kl}) \\ \eta_{0,kl} &\sim f_0 \\ \eta_{g,kl} &\sim f_g \\ f_0 &\sim DP(MP_0) \\ f_g &\sim DP(MP_0). \end{aligned} \tag{2.3}$$

For our purposes we use the logistic link function for $h(\cdot)$, however other link functions such as the probit could also be considered. The prior structure on the regression coefficients is DPM with precision parameter M and center measure $P_0 \equiv N(0, \sigma_\eta^2)$. There are two main advantages to this approach to modeling the regression coefficents. First, the DPM is highly flexible as a prior, and thus we are not making strong assumptions about the prior distribution of the regression coefficients. Second, as discussed in Section 1.3, this prior induces a clustering on the regression coefficients. In doing so, we are able to avoid estimating an extremely large number of regression coefficients (one set for every possible edge) and instead can estimate a much smaller number.

Under our model, the $\eta_{0,kl}$ term is a baseline effect that is common across all cohorts corresponding to the overall probability of an edge between nodes k and l. It is estimated by pooling information across all of the cohorts of interest. On the other hand, the $\eta_{g,kl}$ terms correspond to cohort-specific modifications of the edge probability, or differential effects. Thus strong differences between $\eta_{g,kl}$ and $\eta_{g',kl}$, $g \neq g'$ suggest that there is a difference in the probability of an edge between nodes k and l in the two groups. Under the logistic link function, we can interpret $\eta_{g,kl} - \eta_{g',kl}$ as the log odds ratio of having edge kl in the gth network versus the g'th network. Note that the model is overcomplete, and thus the individual parameters are not identifiable since $h(\eta_{0,kl}, \eta_{g,kl}) = h(\eta_{0,kl} + c, \eta_{g,kl} - c)$ for any real constant c. However, the log odds ratios of interest are identifiable.

A schematic representation of the proposed model is illustrated in Figure 2.4.

2.2.3 Posterior Computation

Posterior computation for the BJNL model is carried out via MCMC. All updates are Gibbs. First, we augment the likelihood by introducing edge-inclusion indicators $\delta_{g,kl}$ such that,

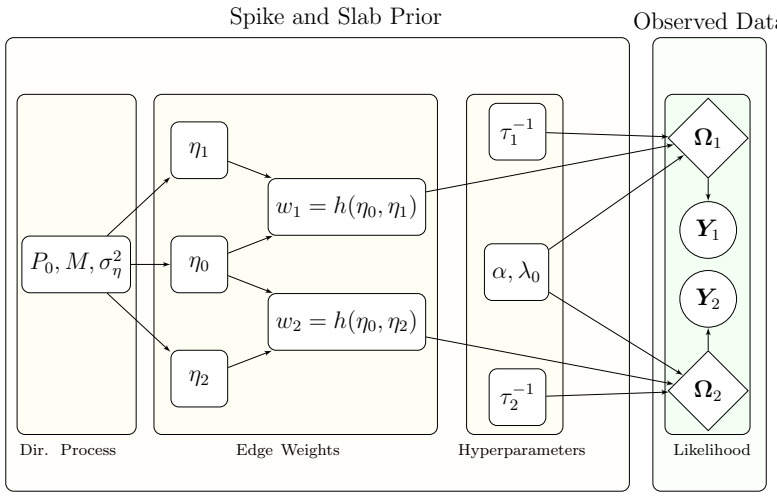


Figure 2.4: Directed graph illustrating the relationships between the model parameters for the case of two experimental conditions represented by fMRI data matrices $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$. Rectangular nodes correspond to parameters which are updated or tuned, diamond-shaped nodes correspond to parameters involved in the likelihood, and the circular nodes correspond to the observed data.

$$\delta_{g,kl} = \begin{cases} 0 & \text{if edge } kl \text{ is not present in group } g \\ 1 & \text{if edge } kl \text{ is present in group } g \end{cases}$$

The augmented likelihood can now be written:

$$\begin{aligned}
\pi(\boldsymbol{\Omega}_g \mid \lambda_0, \boldsymbol{\tau}_g, \boldsymbol{\tau}_g^*) &= C_{\tau,g}^{-1} 1(\boldsymbol{\Omega}_g \in M^+) \prod_{l=1}^{p} Exp(\omega_{g,ll}; \alpha/2) \times \prod_{l=1}^{p} \prod_{k<l} w_{g,kl}^{\delta_{g,kl}} (1 - w_{g,kl})^{1-\delta_{g,kl}} \\
&\quad \times \prod_{l=1}^{p} \prod_{k<l} [N(\omega_{g,kl}; 0, \tau_{g,kl}^{-1})]^{\delta_{g,kl}} [\int N(\omega_{g,kl}; 0, \tau_{g,kl}^{-1}) Exp(\tau_{g,kl}; \frac{\lambda_0^2}{2}) d\tau_{g,kl}]^{1-\delta_{g,kl}}, \\
\pi(\boldsymbol{\tau}_g, \boldsymbol{\tau}_g^*) &\propto C_{\tau,g} \Big(\prod_{l=1}^{p} \prod_{k<l} Ga(\tau_{g,kl}; a_\tau, b_\tau) \times Exp(\tau_{g,kl}^*; \lambda_0^2/2) \Big)
\end{aligned}$$

where $\boldsymbol{\tau}_g = \{\tau_{g,kl}, k \neq l, k, l = 1, \ldots, p\}$, $\boldsymbol{\tau}_g^* = \{\tau_{g,kl}^*, k \neq l, k, l = 1, \ldots, p\}$, $Ga(\cdot; a_\tau, b_\tau)$ corresponds to a Gamma distribution with mean a_τ/b_τ, and $C_{\tau,g}$ is the intractable normalizing constant. Note that $C_{\tau,g}$ cancels out in the expression for $\pi(\boldsymbol{\Omega}_g, \lambda_0, \boldsymbol{\tau}_g, \boldsymbol{\tau}_g^*)$. After integrating out $\boldsymbol{\tau}_g^*$ this results in a marginal prior $\pi(\boldsymbol{\Omega}_g, \lambda_0, \boldsymbol{\tau}_g)$ as shown in (2.2).

For the double exponential distribution to yield a sharp spike at zero, we require that λ_0 be large. For our purposes we select $\lambda_0 = 100$ as a hyperparameter, although a prior distribution could instead be used to learn λ_0 from the data. The sharp spike at zero in turn provides strong shrinkage for the off-diagonal elements of the precision matrix that correspond to absent edges. In a similar way, for the slab component to effectively be able to model very dispersed values, we need the latent scale parameters for the slab component to be large. We select a_τ and b_τ such that a_τ/b_τ is small, which allows for long tails on the Gamma prior for the scale parameters.

As discussed in Section 2.2, we choose the logistic link function to relate the linear predictor to the probability of an edge being present between nodes k and l in group g, $\pi_{g,kl}$. To carry out posterior computation under this approach, we use the an approximation to the logistic link function. Following O'Brien and Dunson (2004) we use,

$$\frac{e^{\mu*}}{(1+e^{\mu*})} \approx \int_0^\infty \mathbf{t}\big(u; \mu*, \frac{\pi^2(\phi-2)}{3\phi}\big)du = \int_0^\infty N(u; \mu*, \frac{\pi^2(\phi-2)}{3\phi}\sigma^2_\phi)\pi(\sigma^2_\phi; \frac{\phi}{2}, \frac{\phi}{2})du,$$

where $\mathbf{t}(\cdot)$ denotes a t-distribution, $\pi(\sigma^2_\phi)$ corresponds to a inverse Gamma distribution, $\phi = 7.3$, and u is a Gaussian latent variable used for data augmentation.

Note that the use of this approximation to the logistic link function implies that we are sampling from a posterior that is an approximation to the original posterior specification, which is based on the logistic link function. However, our posterior computation is still fully Gibbs, as all MCMC samples are being drawn from exact posterior distributions. We could consider an alternative to this approximation scheme. For example, Polson et al. (2013) proposed a Polya-gamma data augmentation scheme for Bayesian logistic regression. However, in our experience the approximation in O'Brien and Dunson (2004) works quite well.

Posterior computation under the DPM is more complicated. As discussed in the intro-

duction, the DPM can be written using a stick breaking representation (Sethuraman, 1994) as:

$$\begin{aligned}
\eta_{g,kl} &\sim f_g \\
f_g &= \sum_{h=1}^{\infty} \nu_{g,h} \delta_{\eta^*_{g,h}} \\
\eta^*_{g,h} &\sim N(0, \sigma^2_\eta) \\
\nu_{g,h} &= v_{g,h} \prod_{l<h} [1 - v_{g,l}] \\
v_{g,h} &\sim Beta(1, M)
\end{aligned} \tag{2.4}$$

for $g = 0, \ldots, G$. We can then use slice sampling techniques (Walker, 2007; Kalli et al., 2011) to sample from the posterior as follows.

First, compute the prior inclusion probability for edge (k, l) as

$$\begin{aligned}
\theta_{g,kl} = \frac{\exp\{-(u_{g,kl} - \theta^n_{g,kl})\}}{1 + \exp\{-(u_{g,kl} - \theta^n_{g,kl})\}} &= \int_0^\infty \frac{\exp\{-(u_{g,kl} - \theta^n_{g,kl})\}}{(1 + \exp\{-(u_{g,kl} - \theta^n_{g,kl})\})^2} du_{g,kl} \\
&\approx \int_0^\infty \int_0^\infty \mathbf{t}(u_{g,kl}; \theta^n_{g,kl}, \frac{\pi^2(\phi - 2)}{3} \sigma^2_{\phi,g,kl}) \pi(\sigma^2_{\phi,g,kl}, \frac{\phi}{2}, \frac{\phi}{2}) d\sigma^2_\phi du_{g,kl},
\end{aligned}$$

under the approximation to the logistic function where $\theta^n_{g,kl} = \eta_{0,kl} + \eta_{g,kl}$, and $\pi(\sigma^2_\phi)$ follows a inverse-Gamma distribution with parameters $\phi/2$ and $\phi/2$, with $\phi = 7.3$. In order to sample the common and differential effects, we perform the data augmentation described earlier by introducing latent variables $u_{g,kl} \sim N(\eta_{0,kl} + \eta_{g,kl}, \sigma^2_{\phi,g,kl} \frac{\pi^2(\phi-2)}{3})$. In particular the following sampling steps are performed sequentially:

- Sample $u_{g,kl} \sim N(\eta_{0,kl} + \sum_{g=1}^{G} \eta_{g,kl} 1_{g=m}, \sigma^2_{\phi,g,kl} \frac{\pi^2(\phi-2)}{3})$ with $u_{g,kl}$ truncated to 1 if $\delta_{g,kl} = 1$ or to zero if $\delta_{g,kl} = 0$.

- Denote the prior variance for the hth cluster by $\sigma^2_h, h = 1, \ldots, H$. Denote the clus-

ter membership of the klth edge weight by c_{kl}, where $c_{kl} \in \{1, \dots H\}$. The cluster memberships and regression coefficients are updated using a slice sampling technique (Walker, 2007). The update steps are as follows:

1. Sample $u^*_{g,kl} \sim Uniform(0, \nu_{g,kl})$
2. Sample the stick breaking weights as: $v_{g,h} \sim Beta(1 + n_{g,h}, M + \sum_{h'>h} n_{g,h'})$, followed by $\nu_{g,h} = v_{g,h} \prod_{h'<h}(1 - v_{g,h'})$, where $n_{g,h}$ is the number of edges assigned to cluster h in platform g (Sethuraman, 1994).
3. Sample the cluster membership indicators:
   (a) For each edge, calculate the atoms available to it as $A_{kl} = \{h : u^*_{g,kl} < \nu_{g,h}, h = 1, \dots, H, g = 0, \dots, G\}$
   (b) For each edge, calculate the vector of likelihood values $\pi^*_{kl} = [\pi^*_{1,kl} I_{1 \in A_{kl}}, \dots, \pi^*_{H,kl} I_{H \in A_{kl}}]$ where $\pi^*_{h,kl} = \prod_{g=1}^{G} N(u_{g,kl}; \eta_{g,h}, \sigma^2_{\phi,g,kl} \frac{\pi^2(\phi-2)}{3})$
   (c) Calculate the normalized likelihoods $\pi_{h',kl} = \pi^*_{h',kl} / \sum_{h=1}^{H} \pi^*_{h,kl}$
   (d) Draw $r \sim Uniform(0, 1)$ and update $c_{kl} = h'$ where h' is the smallest h' s.t. $r < \sum_{h=1}^{h'} \pi_{h,kl}$
4. Draw the new values for the η terms conditioned on the updated cluster memberships, using a Gaussian posterior distribution.

As a final step, we update the variance terms from gamma posteriors.

Edge Detection: The important network edges (and hence the network structure) can be estimated by either including edges with high marginal inclusion probabilities or those with non-negligible absolute values for the precision off-diagonals, lying above a chosen threshold. We propose a strategy to choose such thresholds in a manner which controls the false discovery rate (FDR). Denoting $\zeta_{g,kl}$ as the marginal posterior exclusion probability for

edge (k, l) in network $\mathcal{G}_g$, one can compute the FDR as in Peterson et al. (2015) as

$$FDR = \frac{\sum_{g=1}^{G}\sum_{k<l}\zeta_{g,kl}1(\zeta_{g,kl} < \kappa)}{\sum_{g=1}^{G}\sum_{k<l}1(\zeta_{g,kl} < \kappa)}, \text{ or } FDR = \frac{\sum_{g=1}^{G}\sum_{k<l}\zeta_{g,kl}1(|\hat{\omega}_{g,kl}| > \kappa^*)}{\sum_{g=1}^{G}\sum_{k<l}1(|\hat{\omega}_{g,kl}| > \kappa^*)}, \quad (2.5)$$

depending on whether the edges are included based on posterior inclusion probabilities or edge strengths. Clearly the FDR increases with κ/κ^*, and one can choose a suitable threshold to control the FDR. In our numerical experiments we found that choosing the edges based on whether the absolute precision off-diagonals were greater than 0.1 results in overall good numerical performance and FDR values which are less than 0.03 across a wide spectrum of scenarios. Hence we recommend this as a default threshold under our approach, and we note that the corresponding threshold for posterior probability for edge selection can be obtained as one which yields similar FDR as computed using (2.5).

Inferring Network Differences: In addition to network estimation, the proposed BJNL provides a natural framework for testing network differences between experimental conditions at multiple scales. In particular, for our Stroop task data analysis, we use MCMC samples under BJNL to obtain the posterior distribution for differences in edge level partial correlations as well as global and local network metrics. At the edge-level, T-tests of the Fisher Z-transformed partial correlation differences for all MCMC samples (after burn-in) were used to infer differences in edge strengths across networks. Similarly, the differences in the graph metrics across conditions were computed at each MCMC iteration, and the central tendency and dispersion of their distributions were statistically assessed by T-tests and Kolmogorov-Smirnov (KS) tests. The p-values of these tests were used to assess significance after controlling for false discoveries (Benjamini and Yekutieli, 2001).

2.3 Simulation

2.3.1 Simulation Setup

We conducted a series of simulations to compare group level network estimation between BJNL and competing methods. These approaches include the graphical horseshoe estimator (HS) (Carvalho et al., 2010; Li et al., 2019) which extends the horseshoe prior in regression settings to graphical model estimation, and the graphical lasso approach (GL) (Friedman et al., 2008) which imposes L_1 penalty on the off-diagonals to impose sparsity, as well as the Joint Graphical Lasso (JGL) (Danaher et al., 2014) which uses a fused lasso penalty to pool information across graphs while encouraging sparsity via a L_1 penalty. While both the HS and GL approaches estimate individual networks separately, the JGL approach is designed to jointly estimate multiple networks. The HS was implemented using Matlab codes provided on the author's website. The JGL and the graphical lasso were implemented using the *JGL* and *glasso* packages in R, respectively. Our method was implemented in Matlab, version 8.3.0.532 (R2014a).

The data for the simulation study was generated under a Gaussian graphical model for n=60 subjects with T=300 time points each and for dimensions $p = 40, 100$. Each subject had data corresponding to two experimental conditions having networks with shared and differential patterns. We considered three different network structures: (a) Erdos-Renyi networks which randomly generate edges with equal probabilities, (b) small-world networks generated under the Watts-Strogatz model (Watts and Strogatz, 1998), and (c) scale-free networks generated using the preferential attachment model (Barabási and Albert, 1999) resulting in a hub network. For each type of network, we obtained an adjacency matrix corresponding to the first experimental condition, and then flipped a proportion of the edges in this adjacency matrix to obtain the second network, adding edges where there were no edges and removing an equal number of edges. The proportion of flipped edges was set to 25%(low), 50%(medium), and 75%(high), which correspond to varying levels of discordance

between the experimental conditions.

After generating the networks, the corresponding precision matrices were constructed as follows. For each edge, we generated the corresponding off-diagonal element from a Uniform(-1,1) distribution and fixed the diagonal elements to be one and the off-diagonals corresponding to absent edges as zero. In order to ensure that the resulting precision matrices were positive definite, we subtracted the minimum of the eigenvalues from each diagonal element of the generated precision matrix. To enable a group level comparison for each scenario, all subjects had the same network across all time points within each experimental condition and the same precision matrices for each network.

Erdos-Renyi Random Networks Truly random graphs in which each edge has equal probability of being present (Erdős and Rényi, 1961).

Small-World Networks Graphs generated using the Watts-Strogatz model (Watts and Strogatz, 1998) in which a small number of hub nodes have a large number of connections, meaning that while most nodes are not directly neighbors, they can easily be reached in a small number of steps via the hub nodes.

Scale-Free Networks Graphs in which the number of connections follows a power law (Barabási and Albert, 1999).

Tuning: We used BJNL with 1000 burn-in iterations and 5000 MCMC iterations. We specified the tuning parameters as follows. We chose $\lambda_0 = 100$ and $\tau_{g,kl} \sim Ga(a_\tau, b_\tau)$ with $a_\tau = 0.1$ and $b_\tau = 1$ to enforce a sharp spike at zero and thick tails for the slab component. The stick breaking weights in the mixture distribution in (2.4) were modeled as $\nu_{g,h} \sim Be(1, M)$, where $M \sim Ga(a_m, b_m)$, and we choose $a_m = 1, b_m = 1$, to encourage a small number of edge clusters for a parsimonious representation. We could increase a_m to encourage a larger number of clusters. However, we have observed that varying a_m has a limited effect on the final estimated network. Our experience in extensive numerical studies suggests that the performance of the approach is not overly sensitive to the choice of λ_0 as long

as it is large enough (> 100); however, extremely large values of λ_0 can result in numerical instability. Moreover, performance is fairly robust to the choice of the hyperparameters in the prior for the precision parameter of the slab component, as long as the ratio $a_\tau/b_\tau < 1$. For the joint graphical lasso that depends on two tuning parameters (a lasso penalty and a fused lasso penalty), we searched a 30×30 grid over $[0.01, 0.1]$ for both parameters to find the best combination using a AIC criteria as recommended in Danaher et. al (2014). The graphical lasso was run independently for each network over a grid of regularization parameter values, and the optimal graph was selected for each network using a BIC criteria as described in Yuan and Lin (2007).

Performance metrics: We assessed the performance of the three algorithms in terms of the ability to estimate the individual networks, as measured by the area under the receiver operating characteristic (ROC) curve (AUC), the accuracy in estimating the strength of connections, as measured by the L_1 error in estimating the precision matrix (L_1 error), the power to detect true differential edges as measured via sensitivity (TPR) and control over false positives for differential edges which is computed as 1-specificity (FPR). For all the metrics, we performed pairwise comparisons using Wilcoxon signed rank tests in order to assess whether one approach performed significantly better than the others. For edge detection, point estimates for the penalized networks were obtained by choosing the threshold for the absolute off-diagonal elements as 0.005, while for BJNL we computed thresholds controlling for false discoveries as described above.

2.3.2 Simulation Results

Figure 2.5 displays the ROC curves for the 100 node simulations, Figures 2.6, 2.7, and 2.8 displays box plots of the reported metrics for the three network generation types, and Tables 2.1 and 2.2 report results for the 40 and 100 node simulations, respectively. The results across the three network types are relatively consistent. First, we note that the degree of dissimilarity between the networks does not appear to have a major effect on the relative

performance of the algorithms, although we conjecture that the differences could be more pronounced for smaller sample sizes. For all settings involving Erdos-Renyi graphs, the proposed BJNL approach outperformed the HS, JGL, and GL uniformly across all metrics under the Wilcoxon signed rank test. Notably, the proposed approach simultaneously achieved a significantly higher TPR and a significantly lower FPR for differential edges, indicating that it was both better able to detect significant differences and less likely to incorrectly classify an edge as differential. These suggest a greater power to detect true differential edges with an adequate control over false positives across all network types, under the BJNL. Further, an increased improvement of the TPR over competing approaches and relative stability of the FPR for differential edges for $p = 100$ versus $p = 40$ indicates a clear advantage of the proposed joint estimation approach for increasing dimensions. For the small-world and scale-free networks, the BJNL also had significantly improved AUC, TPR, and L_1 error metrics, and a comparable or lower FPR, compared to all other considered approaches.

On the other hand, the significantly higher L_1 error under the JGL potentially points to the perils of smoothing over edge strengths across networks under penalized approaches. In particular, assigning similar magnitudes for precision matrix off-diagonals for shared edges may adversely affect the identification of differential edges, as well as the estimation of varying edge strengths for common edges across networks. Moreover while HS has low FPR, it consistently exhibits the lowest AUC and TPR and the highest L_1 error for $p = 100$ across all scenarios, which is concerning. On the other hand, the GL had the highest FPR for both the small-world and scale-free network simulations, but has a reasonable TPR. These results under HS and GL illustrate the difficulties resulting from the separate estimation of individual networks which may result in exceedingly low power to detect true positives (as with HS), or an inflated number of false positives (as with GL).

To examine the sensitivity of the proposed approach with respect to the chosen link function, we performed additional simulation studies by fitting the proposed model to the 100 node data generated as above, but under a probit link. The results in Table 2.3 illustrate

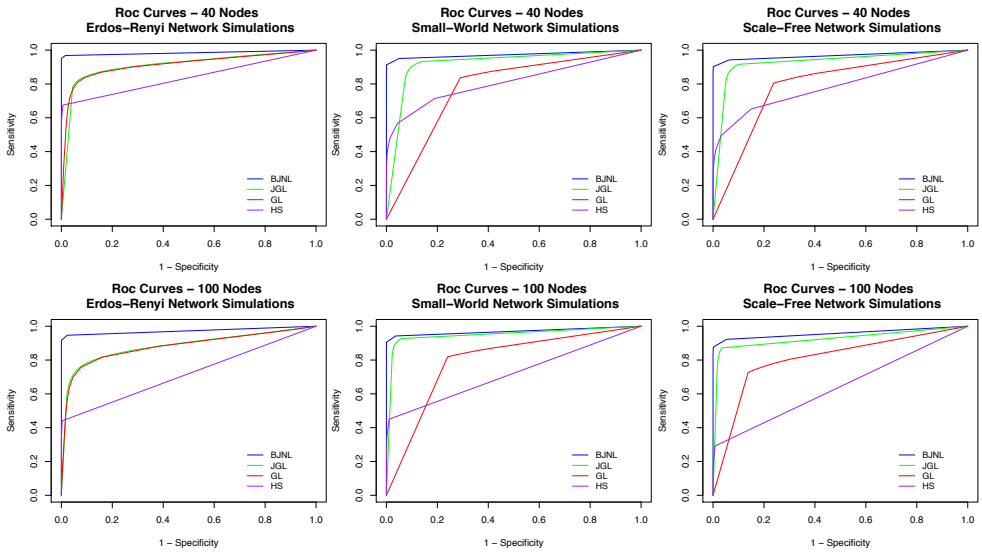


Figure 2.5: ROC curves for edge detection for the 40 and 100 node simulations. The blue, green, red, and purple solid lines correspond to BJNL, JGL, GL, and HS respectively.

non-significant differences in the performance metrics for network estimation across the logit and the probit links, which illustrate the robustness of the proposed approach resulting from the specification of the DP prior on the shared and differential components.

2.4 Stroop Task fMRI Data Analysis

We applied the proposed BJNL to the fMRI Stroop task study to investigate similarities and differences in the brain network under the two experimental conditions and passive fixation (REST). The first analysis was aimed at comparing the mental states of task performance (TASK) and passive fixation (REST), with the hypothesis that the brain networks exhibit major differences between these two grossly different conditions. The TASK data consisted of the subject-wise concatenation of the prewhitened fMRI time courses acquired during the exertion (EXR) and relaxed (RLX) task blocks, while the REST data consisted of the subject-wise concatenation of the prewhitened fMRI time courses acquired during the passive fixation blocks. The second analysis aimed to detect finer differences in connectivity between the mental states of EXR and RLX task performance. The study hypothesized that

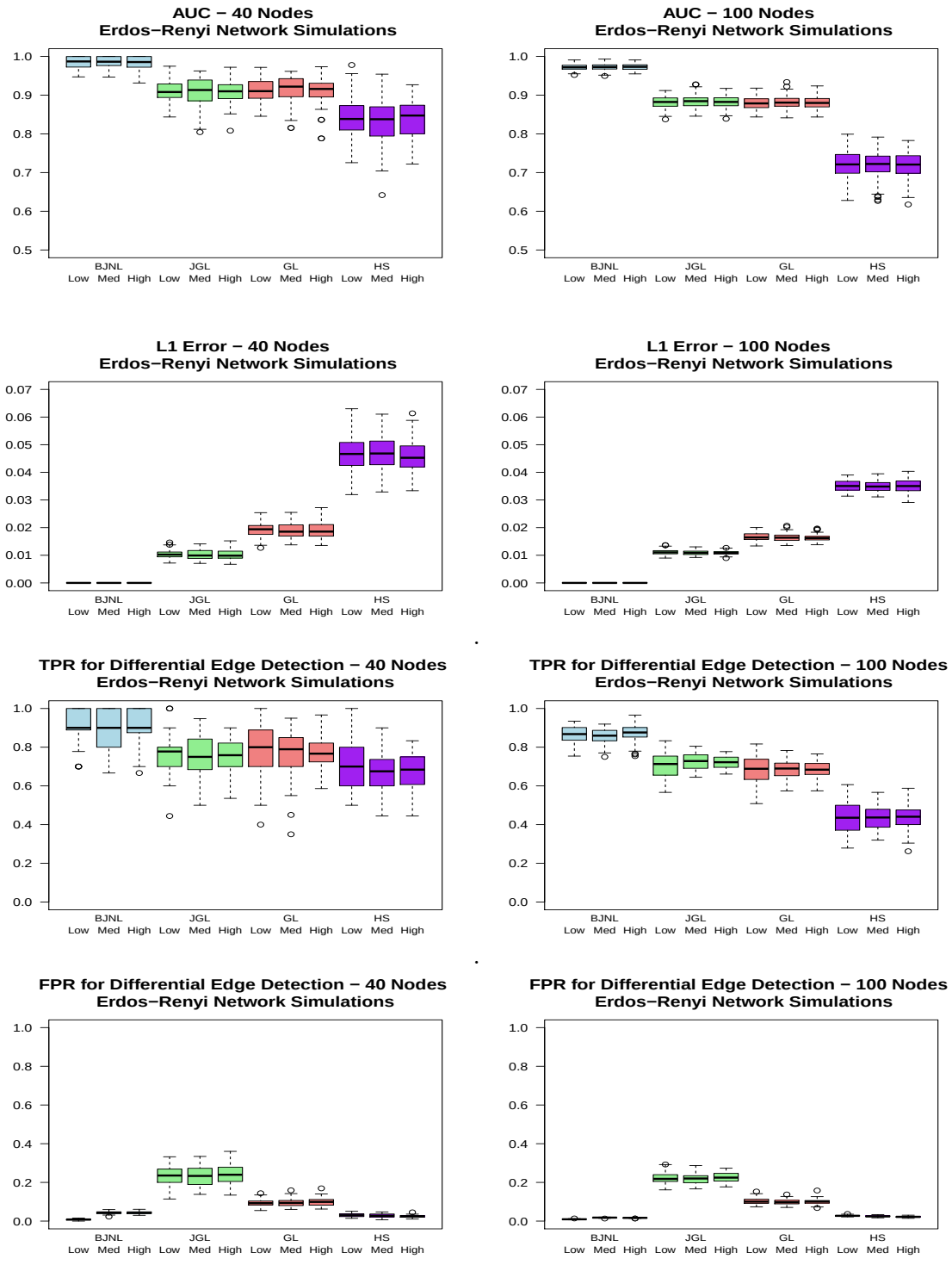


Figure 2.6: Box plots of the AUC, L_1 Error, and TPR/FPR for differential edge detection for the Erdos-Renyi simulations for Bayesian Joint Network Learning (BJNL), the Joint Graphical Lasso (JGL), Graphical Lasso (GL) and the Graphical Horseshoe Estimator (HS). Within each approach, the box plots are organized as: low difference, medium difference, and high difference in edges between experimental conditions, in that order.

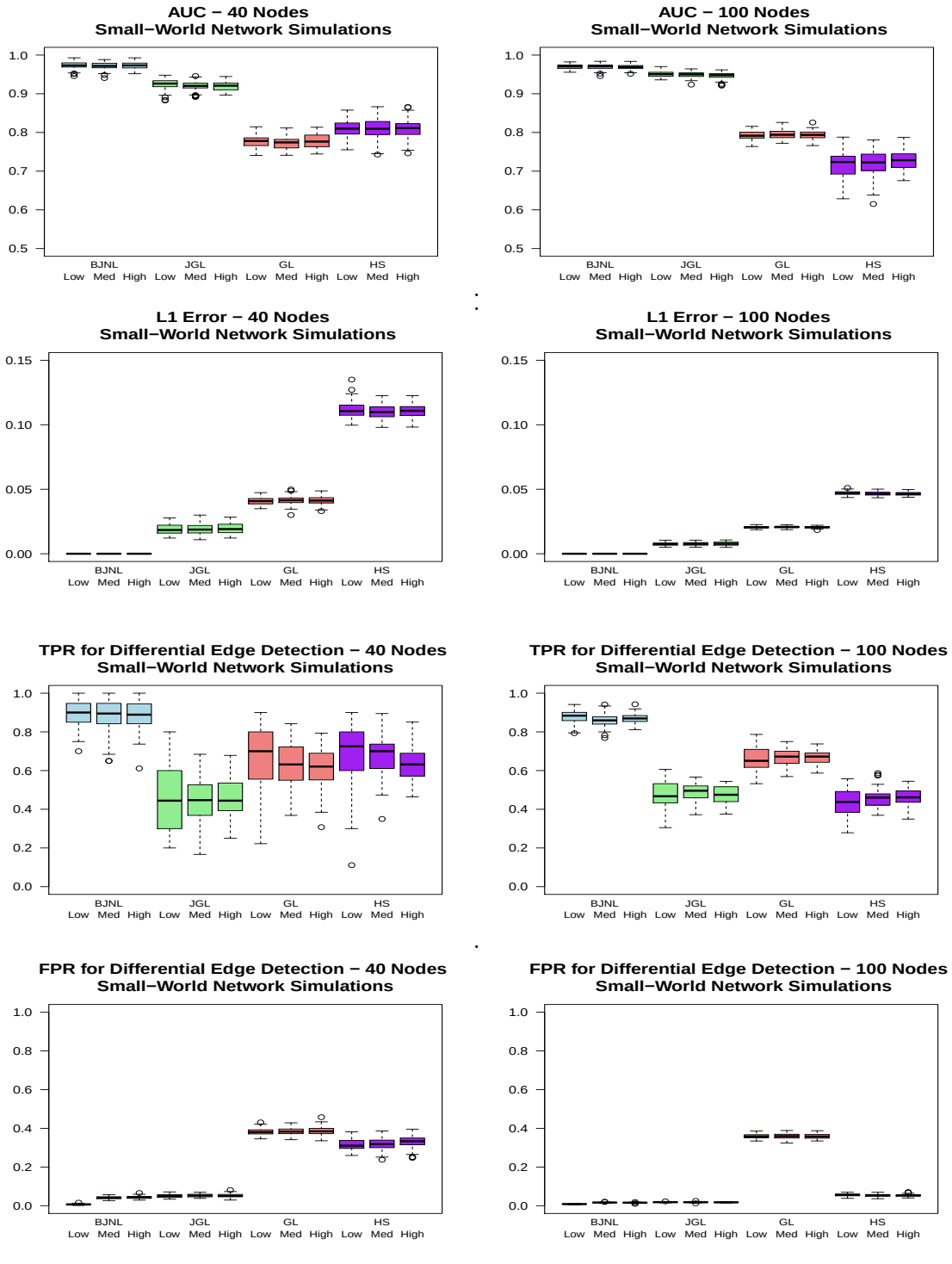


Figure 2.7: Box plots of the AUC, L_1 Error, and TPR/FPR for differential edge detection for the Small-World network simulations. Within each color, the box plots are organized as: low difference, medium difference, and high difference in edges between experimental conditions, in that order.

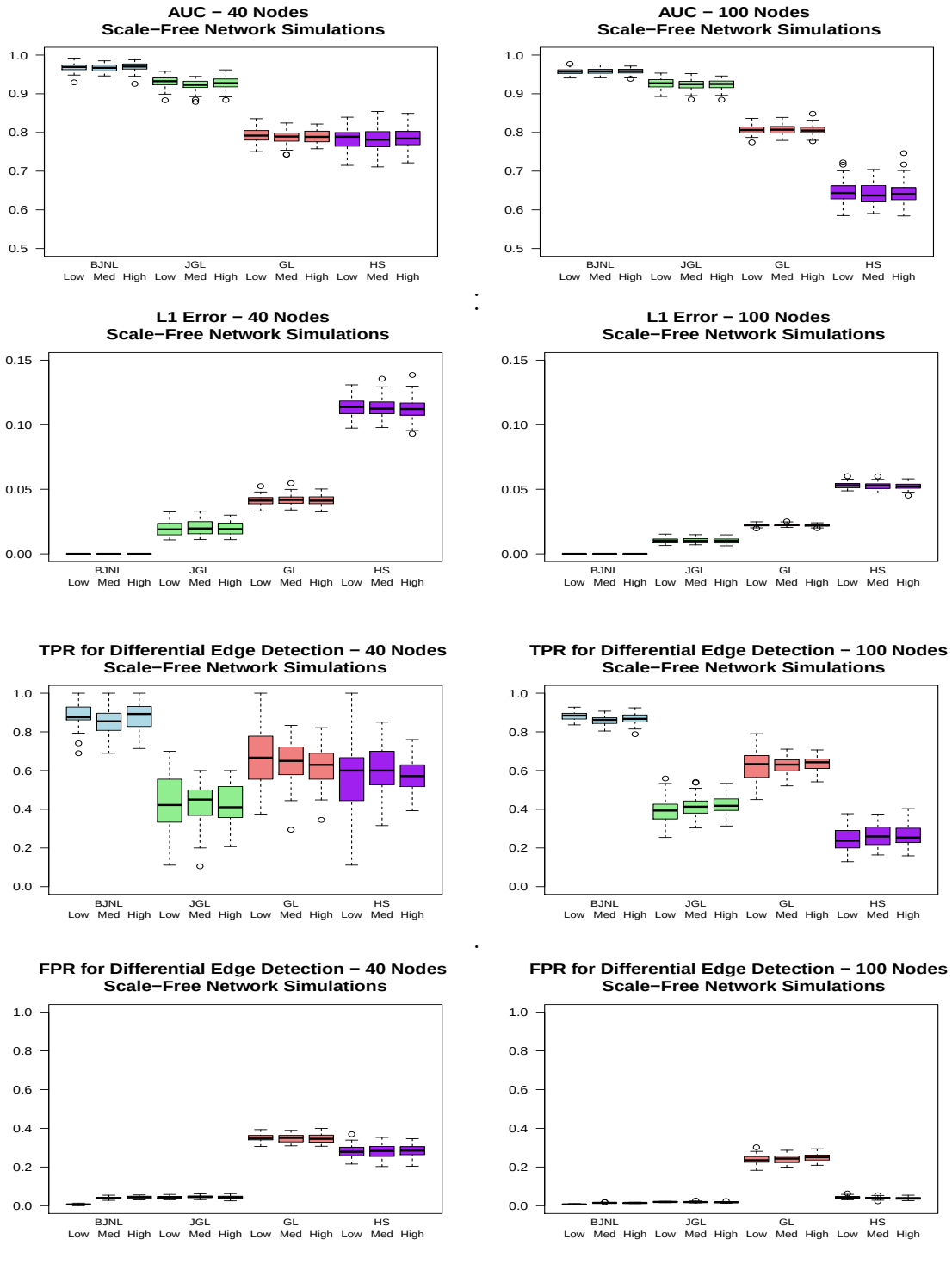


Figure 2.8: Box plots of the AUC, L_1 Error, and TPR/FPR for differential edge detection for the Scale-Free network simulations. Within each color, the box plots are organized as: low difference, medium difference, and high difference in edges between experimental conditions, in that order.

Table 2.1: 40 node simulation results comparing BJNL, JGL, GL, and HS. Text in bold indicates a method was better than both competing methods as assessed through Wilcoxon signed rank tests at $\alpha = 0.05$.

	AUC				L_1 Error $\times$ 100			
	BJNL	JGL	GL	HS	BJNL	JGL	GL	HS
Erdos-Renyi								
low	**0.98 (0.01)**	0.91 (0.03)	0.91 (0.03)	0.84 (0.05)	**0.08 (0.02)**	1.02 (0.15)	1.91 (0.26)	4.66 (0.62)
med	**0.99 (0.01)**	0.91 (0.04)	0.91 (0.04)	0.78 (0.03)	**0.09 (0.02)**	1.02 (0.17)	1.90 (0.27)	11.15 (0.58)
high	**0.98 (0.02)**	0.91 (0.03)	0.91 (0.03)	0.83 (0.06)	**0.09 (0.02)**	1.02 (0.18)	1.90 (0.29)	11.35 (0.78)
Small World								
low	**0.97 (0.01)**	0.92 (0.01)	0.78 (0.02)	0.81 (0.02)	**0.50 (0.05)**	1.89 (0.37)	4.09 (0.27)	4.68 (0.62)
med	**0.97 (0.01)**	0.92 (0.01)	0.78 (0.02)	0.78 (0.03)	**0.50 (0.04)**	1.92 (0.41)	4.16 (0.33)	10.98 (0.53)
high	**0.97 (0.01)**	0.92 (0.01)	0.78 (0.02)	0.78 (0.03)	**0.49 (0.05)**	1.98 (0.41)	4.13 (0.33)	11.33 (0.67)
Scale Free								
low	**0.97 (0.01)**	0.93 (0.01)	0.79 (0.02)	0.84 (0.05)	**0.48 (0.03)**	1.95 (0.50)	4.13 (0.33)	4.59 (0.6)
med	**0.97 (0.01)**	0.92 (0.01)	0.79 (0.02)	0.81 (0.02)	**0.47 (0.02)**	2.01 (0.52)	4.18 (0.36)	11.05 (0.51)
high	**0.97 (0.01)**	0.93 (0.01)	0.79 (0.02)	0.79 (0.03)	**0.47 (0.03)**	1.97 (0.48)	4.13 (0.36)	11.21 (0.81)
	TPR				FPR			
	BJNL	JGL	GL	HS	BJNL	JGL	GL	HS
Erdos-Renyi								
low	**0.92 (0.09)**	0.76 (0.12)	0.78 (0.12)	0.70 (0.12)	**0.01 (0.003)**	0.23 (0.05)	0.10 (0.02)	0.03 (0.01)
med	**0.89 (0.07)**	0.75 (0.10)	0.76 (0.12)	0.69 (0.18)	**0.01 (0.003)**	0.23 (0.05)	0.10 (0.02)	0.32 (0.03)
high	**0.88 (0.06)**	0.75 (0.08)	0.77 (0.08)	0.57 (0.17)	**0.01 (0.003)**	0.24 (0.05)	0.10 (0.02)	0.28 (0.03)
Small World								
low	**0.89 (0.10)**	0.45 (0.16)	0.67 (0.16)	0.66 (0.11)	0.04 (0.01)	0.05 (0.01)	0.38 (0.02)	0.03 (0.01)
med	**0.88 (0.09)**	0.44 (0.11)	0.63 (0.12)	0.67 (0.11)	0.04 (0.01)	0.05 (0.01)	0.38 (0.02)	0.32 (0.03)
high	**0.85 (0.07)**	0.46 (0.09)	0.62 (0.10)	0.60 (0.13)	0.04 (0.01)	0.05 (0.01)	0.39 (0.02)	0.28 (0.03)
Scale Free								
low	**0.89 (0.09)**	0.44 (0.13)	0.66 (0.17)	0.68 (0.09)	0.04 (0.01)	0.04 (0.01)	0.35 (0.02)	0.02 (0.01)
med	**0.88 (0.07)**	0.43 (0.11)	0.65 (0.11)	0.64 (0.08)	0.04 (0.01)	0.05 (0.01)	0.35 (0.02)	0.33 (0.03)
high	**0.88 (0.07)**	0.42 (0.10)	0.62 (0.09)	0.58 (0.09)	0.04 (0.01)	0.04 (0.01)	0.35 (0.02)	0.28 (0.03)

the mental states should be similar between the two task conditions with some fine network differences. In this case, the subject-wise prewhitened fMRI time courses were concatenated for the EXR blocks and also separately for the RLX blocks for analysis.

We performed a brain network analysis based on region of interest (ROI) level data, adopting the 90 node Automated Anatomical Labeling (AAL) cortical parcellation scheme described in Tzourio-Mazoyer et al. (2002). For each ROI, we estimated the representative BOLD time series by performing a singular value decomposition on the time series of the voxels within the ROI and extracting the first principal time series. This resulted in 90 time courses of fMRI measurements, one for each ROI, which were then demeaned. We classified each ROI into one of nine functional modules as defined in Smith et al. (2009). We performed standard pre-processing including slice-timing correction, warping to standard Talairach space, blurring, demeaning, and pre-whitening. The fMRI time series was prewhitened using an ARMA(1,1) model, as is common in imaging toolboxes such as AFNI

Table 2.2: 100 node simulation results comparing Bayesian Joint Network Learning (BJNL), the Joint Graphical Lasso (JGL), Graphical Lasso (GL) and the Graphical Horseshoe Estimator (HS). Text in bold indicates a method was better than all other competing methods as assessed through Wilcoxon signed rank tests at $\alpha = 0.05$.

	AUC				L_1 Error $\times$ 100			
	BJNL	JGL	GL	HS	BJNL	JGL	GL	HS
Erdos-Renyi								
low	**0.97 (0.01)**	0.88 (0.02)	0.88 (0.02)	0.72 (0.03)	**0.11 (0.01)**	1.11 (0.09)	1.66 (0.13)	3.51 (0.19)
med	**0.97 (0.01)**	0.88 (0.02)	0.88 (0.02)	0.72 (0.04)	**0.11 (0.01)**	1.09 (0.09)	1.65 (0.14)	3.50 (0.20)
high	**0.97 (0.01)**	0.88 (0.02)	0.88 (0.02)	0.73 (0.03)	**0.11 (0.01)**	1.09 (0.07)	1.62 (0.11)	3.50 (0.23)
Small World								
low	**0.97 (0.01)**	0.95 (0.01)	0.79 (0.01)	0.72 (0.04)	**0.25 (0.01)**	0.75 (0.12)	2.06 (0.08)	4.70 (0.15)
med	**0.97 (0.01)**	0.95 (0.01)	0.80 (0.01)	0.72 (0.03)	**0.24 (0.01)**	0.77 (0.13)	2.07 (0.08)	4.65 (0.14)
high	**0.97 (0.01)**	0.95 (0.01)	0.79 (0.01)	0.73 (0.03)	**0.24 (0.01)**	0.78 (0.13)	2.06 (0.08)	4.65 (0.14)
Scale Free								
low	**0.96 (0.01)**	0.93 (0.01)	0.81 (0.01)	0.64 (0.03)	**0.20 (0.01)**	1.01 (0.20)	2.23 (0.10)	5.30 (0.23)
med	**0.96 (0.01)**	0.92 (0.01)	0.81 (0.01)	0.64 (0.03)	**0.19 (0.01)**	1.02 (0.21)	2.24 (0.90)	5.26 (0.24)
high	**0.96 (0.01)**	0.92 (0.01)	0.81 (0.01)	0.64 (0.03)	**0.19 (0.01)**	1.00 (0.21)	2.20 (0.08)	5.23 (0.23)
	TPR				FPR			
	BJNL	JGL	GL	HS	BJNL	JGL	GL	HS
Erdos-Renyi								
low	**0.87 (0.05)**	0.71 (0.07)	0.68 (0.07)	0.43 (0.08)	0.01 (0.001)	0.22 (0.03)	0.10 (0.02)	0.03(0.00)
med	**0.88 (0.04)**	0.73 (0.04)	0.69 (0.05)	0.44 (0.06)	0.01 (0.001)	0.22 (0.03)	0.10 (0.01)	0.03(0.00)
high	**0.88 (0.02)**	0.72 (0.03)	0.69 (0.04)	0.44 (0.06)	0.01 (0.001)	0.23 (0.02)	0.10 (0.02)	0.02 (0.00)
Small World								
low	**0.86 (0.04)**	0.47 (0.07)	0.66 (0.06)	0.44 (0.07)	0.02 (0.002)	0.02 (0.00)	0.36 (0.01)	0.06 (0.01)
med	**0.86 (0.04)**	0.49 (0.04)	0.67 (0.04)	0.46 (0.05)	0.02 (0.002)	0.02 (0.00)	0.36 (0.01)	0.05 (0.01)
high	**0.86 (0.02)**	0.48 (0.04)	0.67 (0.03)	0.46 (0.05)	0.01 (0.002)	0.02 (0.00)	0.36 (0.01)	0.05 (0.01)
Scale Free								
low	**0.87 (0.05)**	0.39 (0.06)	0.63 (0.07)	0.25 (0.06)	0.02 (0.002)	0.02 (0.00)	0.24 (0.03)	0.04 (0.01)
med	**0.87 (0.03)**	0.41 (0.05)	0.63 (0.04)	0.26 (0.05)	0.02 (0.002)	0.02 (0.00)	0.24 (0.02)	0.04 (0.01)
high	**0.87 (0.03)**	0.42 (0.04)	0.64 (0.04)	0.27 (0.05)	0.01 (0.002)	0.02 (0.00)	0.25 (0.02)	0.04 (0.01)

Table 2.3: Comparison of the 100 node simulation results using the probit link function to the simulation results using the logit link function.

	Erdos Renyi			Small World			Scale Free		
	AUC	TPR	FPR	AUC	TPR	FPR	AUC	TPR	FPR
Probit	0.97	0.88	0.01	0.96	0.86	0.02	0.97	0.87	0.02
Logit	0.97	0.88	0.01	0.97	0.87	0.02	0.96	0.86	0.02

(Cox, 1996). Further details are provided in Appendix A. The proposed BJNL was run using the same tuning parameters as in the simulations.

Graph metrics: We analyzed the brain's connectivity structure during the different mental states in terms of four graph metrics: global efficiency, local efficiency, clustering coefficient, and characteristic path length. Efficiency measures how effectively information is transmitted from node-to-node in a network. Global efficiency measures information transmission across the entire graph and is calculated by taking the average across all ROIs of the inverse shortest path lengths between ROIs. Thus, large values of global efficiency indicate that, on average, the number of steps required to transmit information from one node to another is small. Local efficiency measures information transmission between an ROI and its neighbors and is calculated for each ROI by taking the average of the inverse shortest path lengths between ROIs in the relevant neighborhood, where the relevant neighborhood is the collection of ROIs with a connection to the selected ROI. The clustering coefficient measures the interconnectedness of the graph and is calculated for each ROI by examining how many of its neighbors are also neighbors to each other. Finally, characteristic path length is the average across ROIs of the shortest path length in the networks, with smaller values indicating a more efficient network.

All metrics were calculated using the Matlab Brain Connectivity Toolbox (Rubinov and Sporns, 2010). In addition, we also examined differences in local graph metrics across experimental conditions corresponding to several brain regions that were found to be differentially activated in a previous study using the same Stroop task experiment (Khachouf et al., 2017). Although distinct from earlier activation analysis, potential connectivity differences in these previously identified brain regions will bolster earlier activation based discoveries and also help illustrate the biological interpretability of the connectivity analysis.

2.4.1 Results

TASK vs REST Conditions: The analysis produced a large contingent of edges with signifi-

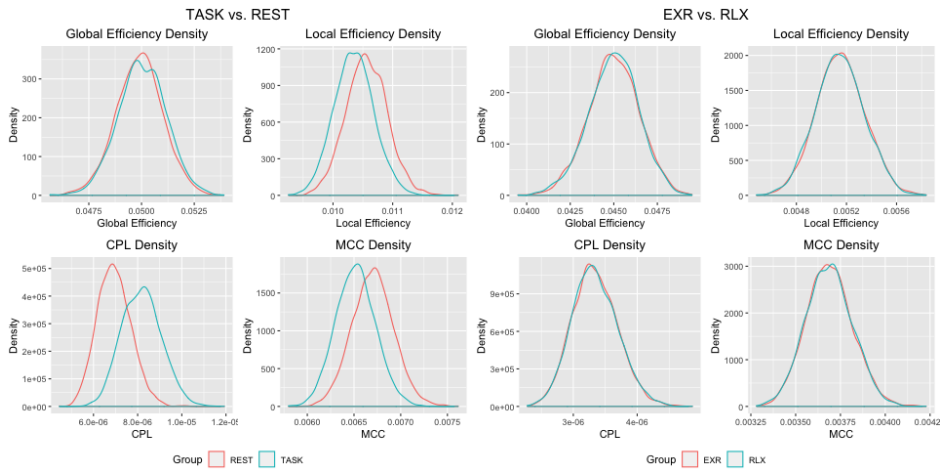


Figure 2.9: Estimated densities of graph metrics for the analysis of task vs. passive fixation and maximum exertion (EXR) vs. relaxed (RLX) task performance.

cantly different edge strengths in the two mental states - Figure 2.10 displays a heatmap of the significant edge counts by functional module. Our analysis revealed 1550 significantly different edges (under T-tests) that provide evidence supporting the study hypothesis that there are major differences in the brain networks due to the manifest phenomenological and procedural dissimilarity of task performance and rest. Moreover, our examination revealed significant differences in the mean (under T-tests) and the posterior distributions (under KS tests) for all network metrics between the two conditions (Figure 2.9). Additional examination of local network differences between task and fixation conditions corresponding to 20 pre-specified regions revealed larger clustering coefficients for REST in all implicated regions, and larger local efficiencies for REST in 18 of the 20 regions (see Table 2.4 for the brain regions and p-values).

EXR vs. RLX conditions of task performance: Compared with the relatively large network differences between TASK and REST, the network structures corresponding to the EXR and RLX task conditions exhibited more nuanced differences. Our analysis revealed 226 significantly different edges between the EXR and RLX conditions - see Figure 2.10 for a heatmap of the significant edge counts by functional module. None of the graph metrics were significantly different between the EXR and RLX conditions, implying that the network dif-

Table 2.4: P-values for testing differences between TASK and REST in local efficiency and clustering coefficient for the nodes identified in the Khachouf (2017) study. Our analysis revealed larger clustering coefficients for REST in all implicated regions, and larger local efficiencies for REST in 18 of the 20 regions

AAL Region	Metric Center (t-test)						Metric Distribution (ks-test)	
	Local Efficiency			Clustering Coefficient			Local Efficiency	Clustering Coefficient
	BJNL	JGL	GL	BJNL	JGL	GL	BJNL	BJNL
Cingulate_Ant_L	**0.01**	0.27	0.52	**<0.01**	0.40	0.53	**0.03**	**<0.01**
Cingulate_Ant_R	**<0.01**	0.75	0.61	**<0.01**	0.87	0.78	**<0.01**	**<0.01**
Cingulate_Mid_L	**0.01**	**<0.01**	0.94	**<0.01**	**<0.01**	0.70	**0.03**	**<0.01**
Cingulate_Mid_R	**<0.01**	0.50	0.58	**<0.01**	0.54	0.45	**<0.01**	**<0.01**
Cingulate_Post_L	**<0.01**	**0.01**	**0.04**	**<0.01**	**0.04**	**0.03**	**0.01**	**<0.01**
Cingulate_Post_R	**<0.01**	0.55	0.12	**<0.01**	0.26	0.08	**<0.01**	**<0.01**
Supp_Motor_Area_L	**<0.01**	**<0.01**	0.99	**<0.01**	**<0.01**	0.85	**<0.01**	**<0.01**
Supp_Motor_Area_R	**<0.01**	0.35	0.94	**<0.01**	0.73	0.86	**<0.01**	**<0.01**
Insula_L	**<0.01**	0.13	0.34	**<0.01**	**0.02**	0.34	0.07	**0.04**
Insula_R	**0.03**	0.31	0.06	**<0.01**	0.06	0.11	**0.02**	**<0.01**
Parietal_Sup_L	**<0.01**	**<0.01**	0.52	**<0.01**	**0.03**	0.71	**<0.01**	**<0.01**
Parietal_Sup_R	**<0.01**	**0.02**	0.64	**<0.01**	**0.05**	0.38	**<0.01**	**<0.01**
Parietal_Inf_L	**<0.01**	0.74	0.64	**<0.01**	0.64	0.69	**0.02**	**<0.01**
Parietal_Inf_R	**<0.01**	**<0.01**	0.14	**<0.01**	0.09	0.37	**<0.01**	**<0.01**
Thalamus_L	**<0.01**	0.70	0.68	**<0.01**	0.54	0.81	**<0.01**	**<0.01**
Thalamus_R	**<0.01**	0.64	**<0.01**	**<0.01**	0.69	**<0.01**	**<0.01**	**<0.01**
Hippocampus_L	**<0.01**	0.28	0.58	**<0.01**	0.71	0.61	**<0.01**	**<0.01**
Hippocampus_R	**0.01**	0.83	0.86	**<0.01**	0.63	0.72	0.13	**<0.01**
Paracentral_Lobule_L	**<0.01**	0.98	0.89	**<0.01**	0.36	0.57	**<0.01**	**<0.01**
Paracentral_Lobule_R	**<0.01**	0.38	0.60	**<0.01**	0.82	0.46	**<0.01**	**<0.01**

ferences did not manifest at a global level (Figure 2.9). However, more localized changes were discovered in the pre-selected regions that were previously shown to be differentially activated between EXR versus RLX (Khachouf et al., 2017). Significant differences were found in terms of mean local efficiency in the right inferior occipital node and the left caudate. Similarly, significant differences were found in mean and distribution for the clustering coefficient for the right inferior occipital node. Several borderline network differences were also identified - see Table 2.5 for the p-values.

Interpretation of Findings: Our BJNL analysis identified strong connectivity differences between Stroop task performance and passive fixation in terms of significantly higher efficiency and clustering, and lower characteristic path length for REST as well as stronger positive connections involving frontoparietal circuits, EC, DMN, sensorimotor, and visual cortices in the TASK condition compared to REST. Our findings also aligned with the widely used the-

Table 2.5: P-values for testing differences between EXR and RLX in local efficiency and clustering coefficient for the nodes identified in the Khachouf (2017) study. Significant differences were found in terms of mean local efficiency in the right inferior occipital node and the left caudate, and borderline significant results were found for the left superior parietal cortex and left insula. Similarly, we found significant differences in the mean and posterior distribution of the clustering coefficient for the right inferior occipital node. We identified borderline significant differences in the clustering coefficient distribution and center for the left anterior cingulate, the mean of the clustering coefficient in the left insula, and the mean and distribution of the clustering coefficient in the left caudate.

AAL Region	Metric Center (t-test)						Metric Distribution (ks-test)	
	Local Efficiency			Clustering Coefficient			Local Efficiency	Clustering Coefficient
	BJNL	JGL	GL	BJNL	JGL	GL	BJNL	BJNL
Parietal_Sup_L	0.10	0.37	0.97	0.33	0.36	0.70	0.36	0.42
Parietal_Sup_R	0.88	0.66	0.48	0.78	0.64	0.31	0.83	0.74
Supp_Motor_Area_L	0.88	0.92	0.14	0.53	0.90	0.21	0.80	0.36
Supp_Motor_Area_R	0.94	0.77	0.73	0.50	0.81	0.71	0.97	0.86
Cingulate_Ant_L	0.47	**0.03**	0.64	0.07	**0.03**	0.62	0.28	0.07
Cingulate_Ant_R	0.19	**<0.01**	0.13	0.15	**<0.01**	0.26	0.52	0.26
Cingulate_Mid_L	0.89	0.34	0.87	0.57	0.38	0.53	0.24	0.67
Cingulate_Mid_R	1.00	**0.02**	0.82	0.72	0.07	0.81	0.39	0.24
Insula_L	0.07	0.21	0.82	0.08	0.21	0.65	0.24	0.15
Insula_R	0.81	0.21	0.10	0.72	0.21	**0.04**	0.97	0.13
Caudate_L	**0.05**	0.35	0.84	0.09	0.34	0.39	0.24	0.07
Caudate_R	0.20	0.49	0.89	0.71	0.48	0.81	0.27	0.76
Occipital_Sup_L	0.31	**0.01**	0.66	0.43	**0.05**	0.63	0.57	0.69
Occipital_Sup_R	0.76	1.00	0.69	0.75	**<0.01**	0.77	0.71	0.93
Occipital_Mid_L	0.22	1.00	0.39	0.86	**<0.01**	0.56	0.51	0.80
Occipital_Mid_R	0.82	1.00	0.56	0.86	**<0.01**	0.46	0.91	0.93
Occipital_Inf_L	0.64	1.00	0.73	0.80	**<0.01**	0.63	0.57	0.56
Occipital_Inf_R	**0.01**	0.35	0.13	**0.03**	0.35	0.14	**0.01**	**0.04**
Precuneus_L	0.27	0.56	0.30	0.19	0.53	0.51	0.15	0.36
Precuneus_R	0.27	0.47	0.14	0.84	0.46	0.23	0.43	0.97
Paracentral_Lobule_L	0.70	0.54	0.93	0.68	0.56	0.98	0.54	0.21
Paracentral_Lobule_R	0.64	0.25	**0.02**	0.61	0.24	**0.03**	0.35	0.39
Fusiform_L	0.38	**0.01**	0.11	0.39	**0.01**	0.13	0.31	0.52

ory of global workspace where more difficult tasks are associated with increased connection distance, as well as reduced clustering (GeiBerg et. al, 2013). More localized associations were also discovered in all regions identified as differentially activated in previous studies (Khachouf et al., 2017), which highlights the biological interpretability of our connectivity findings. Our analysis provides exciting new insights into the connectivity differences between passive fixation and the task-related network that requires a rearrangement of connections in order to perform the task.

On the other hand, fewer connectivity differences were discovered between EXR and RLX task performance, as expected due to the only difference between conditions being the level of voluntary effort invested in the task. While no global topological differences between the EXR and RLX conditions were discovered, the BJNL analysis did reveal some fine differences in the functional modules including the EC and FPL that are involved in high level cognitive function, as well as some limited localized connectivity differences in 23 pre-specified brain regions that previously showed major activation differences in (Khachouf et al., 2017). In general compared to TASK versus REST, the RLX task performance condition featured significantly more negative connections between regions compared to EXR, and there were fewer connectivity differences between nodes within the EC. Compared to a much larger number of connectivity differences in EC and other functional modules in TASK versus REST, the limited connectivity differences between EXR and RLX implies a restricted rearrangement of the network between EXR and RLX.

Comparison with penalized approaches: We are also interested in comparing the network differences under BJNL with those obtained under penalized methods. Hence we performed an illustrative analysis for the Stroop task data using the GL and JGL approaches that involved permutation testing to infer significant network differences between experimental conditions. A permuted sample for two experimental conditions was generated by randomly switching the labels across conditions multiple times. Then, the networks corresponding to these permuted samples were computed using JGL and GL. Subsequently, the network

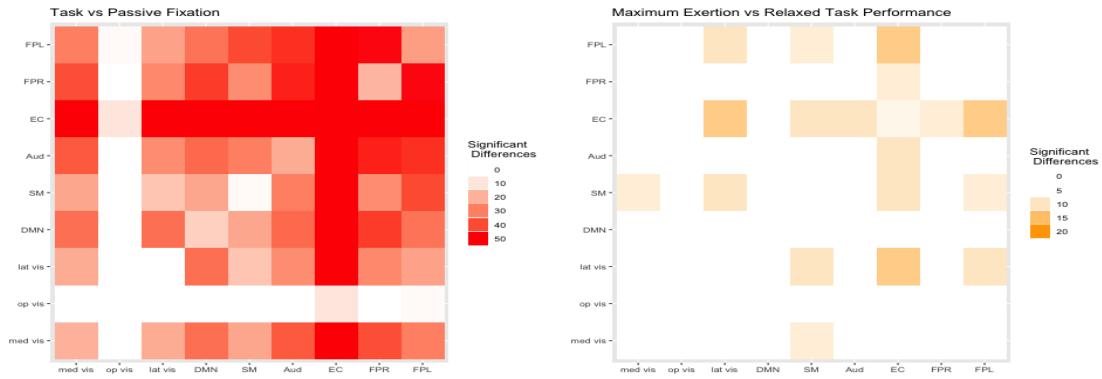


Figure 2.10: Heatmaps of the number of differential edges between conditions. The heatmap on the left corresponds to the analysis of task vs. passive fixation, and the heatmap on the right corresponds to the analysis of maximum exertion (EXR) vs. relaxed task performance (RLX).

differences corresponding to distinct experimental conditions were computed. The above process was repeated 10,000 times, and the permutation distributions for between-network differences were constructed to compute p-values to test for significant differences.

Note that it was computationally infeasible to use AIC to select the tuning parameters for JGL for all 10,000 permutations since the run time for the best tuning parameter search over a grid took one hour per permutation. Hence, the starting values for the tuning parameters for JGL were selected as those values used for the JGL analysis for the original data without permutation. The tuning parameters were then adaptively searched on a permutation-by-permutation basis until the resulting edge density was within 20% of the edge density for the network corresponding to the original samples. While the process was required to make testing under the JGL computationally feasible, it could potentially result in misleading results under JGL due to possible mis-specification of network densities.

The analysis revealed that only one of the resulting edges for the EXR versus RLX network comparisons was significant under the GL, whereas only 62 edges were significant for the JGL. Similarly, for the analysis of TASK versus REST, the JGL identified 476 edges with differential strengths and 3873 common edges (versus 1550 differential edges and 1565 common edges under BJNL). In this case, GL was able to identify 51 edges with differential strengths, and 552 common edges. We believe that the low number of differential edges

between EXR versus RLX conditions under the penalized approaches is unrealistic, and that more differences are to be expected between TASK and REST since it involves significant differences in brain activation across the brain (Khachouf et al., 2017). Further, only 5 of the 20 pre-specified brain regions which were shown to be differentially activated had significant network differences between TASK and REST under the penalized approaches. These results suggests the proposed BNJL method has much better statistical power to detect differences in brain networks under different cognitive states compared to penalized approaches for modeling networks.

2.5 Discussion

In Topic 1 we introduced a novel Bayesian approach to joint estimation of multiple group level brain networks that pools information across networks to estimate shared and differential patterns in brain functional networks formed under different cognitive conditions. The proposed BJNL approach naturally enables a systematic inferential framework for comparing networks, which is a central question of interest in many connectome studies including our Stroop task application where the focus is to investigate connectivity differences between passive fixation and relaxed and exertion modes of Stroop task. Our analysis of Stroop task data revealed important dissimilarities between the task and rest conditions, but more subdued differences between the two task conditions, which aligns with the scientific hypotheses of the study. Moreover the connectivity differences were found to be concentrated in brain regions shown to be differentially activated for Stroop task in previous studies, which signifies that the connectivity differences are biological interpretability. In contrast, a separate estimation of networks using penalized approaches identified negligible or limited connectivity differences between varying modes of mental effort that seem biologically implausible. In addition, the joint estimation of multiple networks under a penalized approach is not naturally conducive for comparing networks and hence one had to use computationally prohibitive

permutation tests that tend to give sub-optimal results in terms of network accuracy and inferring between-network differences.

In this topic, we demonstrated BJNL for estimating networks using fMRI data because they are the most prevalent type of functional images. However, our method can also be generalized to data from other imaging modalities in a straightforward manner. One advantage of our proposed approach for clustering the edge weights is that it allows for unsupervised estimation of the number of clusters. This means that in generalizing the method to other modalities, we do not have to laboriously tune the clustering parameters to each individual problem. Going beyond multiple experimental conditions, our approach can also be used to jointly model networks across multiple cohorts, such as healthy individuals, subjects with mild cognitive disorder, and those with Alzheimer's disease (Kundu et al., 2019). Future work should investigate the scalability of BJNL to larger numbers of conditions while taking into account the dynamic nature of the brain networks over time.

Chapter 3

Sparse Bayesian Independent Component Analysis

3.1 Introduction

3.1.1 ICA - Introduction

Neuroimaging studies have revealed that human brain function is composed of activation from various "functional brain networks" that consist of spatially disjoint brain regions that exhibit coherent temporal dynamics (Friston et al., 1993). These networks are most commonly studied using functional magnetic resonance imaging (fMRI), which captures the blood-oxygen-level-dependent (BOLD) signal series across hundreds of thousands of brain locations (known as voxels). The brain functional networks have well-known neurobiological interpretations.

Independent component analysis (ICA) is one of the most popular techniques for studying brain functional networks. The standard spatial ICA approach decomposes the data for a single subject into a set of spatially independent ICs and corresponding mixing matrix as,

$$\mathbf{Y}_i = \mathbf{A}_i\mathbf{S}_i + \mathbf{e}_i, \tag{3.1}$$

where $\mathbf{A}_i$ is the subject specific mixing matrix for the ith subject and $\mathbf{S}_i$ is their corresponding set of spatial maps (independent components). This type of spatial ICA model is known as a probabilistic spatial ICA model due to the inclusion of the error term $\mathbf{e}_i$ (Beckmann and Smith, 2004).

While the ICA decomposition for a single subject is straightforward, in the majority of studies researchers will be interested in aggregating data across subjects to obtain more reliable estimates of brain functional networks. In such cases, performing a separate decomposition for each subject is undesirable, as it is difficult to match and compare the ICs from the different subjects. Instead, group ICA approaches are used to decompose the data in a manner that ensures the spatial components are comparable across subjects. These methods include temporal-concatenation group ICA (Calhoun et al., 2001, 2009; Erhardt et al., 2011;

Du and Fan, 2013) and tensor ICA (Beckmann and Smith, 2005).

The above data aggregation approaches are suitable for identifying population-level brain functional networks. However considerable differences in brain functional networks also exist across individuals and sub-populations. These individual level differences are associated with different clinical (Greicius et al., 2004; Liu et al., 2008) or demographic characteristics. Under TC-GICA approaches, an ad-hoc two-stage technique must be used to answer such questions. First, an initial TC-GICA decomposition is performed, then the back-projected subject-specific IC maps are regressed against the covariates of interest to obtain covariate effect maps. This comes with the disadvantage of failing to account for variability in the subject-level ICA decomposition, which in turn can result in less accurate effect estimates and inefficient hypothesis testing.

Recently, Shi and Guo (2016) introduced a hierarchical covariate-adjusted ICA method (hc-ICA) that provides a formal statistical framework for estimating covariate effects and testing differences in brain functional networks. This approach also enables model-based estimation of sub-population networks such as those of a diseased group versus a control group. Their hierarchical approach models subject-specific spatial ICs as a function of population-level source signals modulated by covariate effects and additional between-subject random variabilities. The approach has been shown to result in more accurate estimates of brain networks at the individual and group levels, and improved power to detect covariate effects as compared to TC-GICA approaches such as dual-regression (Shi and Guo, 2016; Wang and Guo, 2019; Lukemire, Wang, Verma and Guo, 2020).

While the hc-ICA framework offers significant improvements over two-stage approaches, it faces several limitations which can significantly hamper its performance. In particular, hc-ICA relies on a mixture of Gaussians (MoG) approach to model the population level source signals. In the hc-ICA model, voxel-specific activation in different components corresponds to different configurations of a latent state variable underlying the MoG. Though this approach offers nice interpretability, it is computationally expensive due to the cardinality of the latent

space increasing exponentially with the number of components. Proposed solutions to this issue include a subspace EM (Shi and Guo, 2016) or template ICA methods (Mejia et al., 2020).

Furthermore, hypothesis testing under hc-ICA (or TC-GICA with dual-regression) involves performing tests at hundreds of thousands of voxels for each covariate effect under investigation. The influence of clinical covariates on the spatial extent of functional brain networks should be small in practice. That is, most brain function should be common across subjects, and clinical covariates such as disease status will likely only influence small regions of the brain network (see, for example, Rytty et al. (2013); Smith et al. (2014); Reineberg et al. (2015)). Because of this, uncorrected voxel-level hypothesis testing will likely have an unacceptably large false discovery rate. Current hc-ICA and dual-regression approaches must account for this after estimation via some method of FDR control (Genovese et al., 2002). However, this can result in less power to detect true covariate effects.

Motivated by these limitations, we introduce a Sparse Bayesian Independent Component Analysis (SparseBayes ICA) approach. This is a fully Bayesian hc-ICA method that uses shrinkage priors to model the covariate effects and places Bayesian non-parametric priors on the shared source signals. To the best of our knowledge, this is the first Bayesian hc-ICA method. Our approach addresses the limitations of current group-level ICA techniques. In particular, our approach uses a Dirichlet process mixture of Gaussians to model the population-level source signals as an infinite mixture of Gaussians. This resolves the MoG cardinality issues discussed above, and comes with the added advantage that it does not require us to specify a number of Gaussian components. To address the large scale hypothesis testing required testing for covariate effects in ICA, we propose to use horseshoe shrinkage priors on the covariate effects, allowing us to build the sparsity assumption into our model. To the best of our knowledge, this is the first approach to incorporate shrinkage estimates for covariate effects directly into the ICA decomposition. Finally, current hc-ICA approaches are very sensitive to the choice of initial parameter values due to their use of an EM algorithm in

estimation. Our approach allows us to perform posterior computation using Markov chain Monte Carlo (MCMC), and as such it is able to better explore the parameter space. Using MCMC also allows us to construct credible intervals for any quantities of interest, which avoids some of the difficulties with constructing standard error estimates for performing hypothesis testing in the typical hc-ICA case. This also enables hypothesis testing for other quantities, such as the population-level spatial maps.

We apply our proposed SparseBayes ICA approach to a resting-state fMRI study of Zen meditators. We show that our approach is able to identify the default mode network in a way that is more consistent with the literature than competing methods. We also are able to identify several regions exhibiting differences related to meditation training on the default mode network where other methods find fewer regions, or none at all.

This chapter proceeds as follows. In Section 2 we introduce the proposed method, and in Section 3 we discuss the steps required by the posterior computation. Then, in Section 4 we carry out extensive simulation studies to demonstrate some of the potential benefits of our approach. In Section 5, we use our approach to carry out an ICA decomposition of the motivating Zen meditation resting-state study. We also compare our findings to those from hc-ICA and TC-GICA. Finally, in Section 6 we provide some concluding remarks.

3.2 SparseBayes ICA

In this section, we present the SparseBayes ICA framework for investigating individual differences in brain networks.

3.2.1 Preprocessing prior to ICA

Following standard ICA procedures, we perform some preprocessing steps prior to ICA decomposition which include centering, dimension reduction, and whitening of the fMRI data (Hyvärinen and Oja, 2000). Let T be the number of fMRI scans are acquired during

the imaging session and let V be the total number of voxels. Let $\widetilde{\mathbf{Y}}_i(v) \in \mathbb{R}^T$ be the centered time series recorded for subject i at voxel v and $\widetilde{\mathbf{Y}}_i = [\widetilde{\mathbf{Y}}_i(1), ..., \widetilde{\mathbf{Y}}_i(V)]$ is the $T \times V$ fMRI data matrix for subject i. We perform the following dimension reduction and prewhitening of the data (Shi and Guo, 2016):

$$\mathbf{Y}_i = (\mathbf{\Lambda}_{i,Q} - \tilde{\sigma}^2_{i,Q}\mathbf{I}_Q)^{-\frac{1}{2}}\mathbf{U}'_{i,Q}\widetilde{\mathbf{Y}}_i, \tag{3.2}$$

where $\mathbf{U}_{i,Q}$ and $\mathbf{\Lambda}_{i,Q}$ contain the first Q eigenvectors and eigenvalues from the singular value decomposition of $\widetilde{\mathbf{Y}}_i$. $\tilde{\sigma}^2_{i,Q}$ is the residual variance, representing the variability in $\tilde{\mathbf{Y}}_i$ that is not accounted by the first Q components; it is estimated by the average of the smallest $T-Q$ eigenvalues that are not included in $\mathbf{\Lambda}_{i,Q}$. The parameter Q is the number of ICs and can be determined by quantitative methods such as the Laplace approximation method (Minka, 2001; Shi and Guo, 2016). In neuroimaging studies, the selection of Q can also be guided by the scale of brain networks investigators are interested in identifying. The literature (Smith et al., 2013) has established that low model orders lead to large-scale networks responsible for broad sets of similar functions and high model orders give rise to small networks with more specific functions under the same umbrella. The ICA decomposition is then performed after the prior-ICA pre-processing. In what follows, we present the methods for the prewhitened data, $\mathbf{Y}_i = [\mathbf{Y}_i(1), ..., \mathbf{Y}_i(V)]$ $(i = 1, ..., N)$, which is of dimension $Q \times V$.

3.2.2 SparseBayes ICA Method

In this section, we present a SparseBayes ICA method for investigating individual differences in brain networks using multi-subject fMRI data. The SparseBayes ICA is a hierarchical group ICA model. The first-level decomposes a subject's fMRI signals into a product of an individual-level mixing matrix and individual-level ICs,

$$\mathbf{Y}_i(v) = \mathbf{A}_i\mathbf{S}_i(v) + \mathbf{E}_i(v), \tag{3.3}$$

where $\mathbf{S}_i(v) = [s_{i1}(v), ..., s_{iQ}(v)]'$ is a $Q \times 1$ vector with $s_{iq}(v)$ representing the spatial source signal of the qth IC (i.e., brain functional network) at voxel v for subject i. The Q elements of $\mathbf{S}_i(v)$ are assumed to be independent and non-Gaussian. $\mathbf{A}_i$ is the ith subject's $Q \times Q$ mixing matrix, and $\mathbf{E}_i(v) = [e_{i1}(v), ..., e_{iQ}(v)]'$ is a $Q \times 1$ noise term. Following previous work (Beckmann and Smith, 2004; Guo, 2011; Shi and Guo, 2016), we assume that the noise term is independent across voxels and follows a Gaussian distribution with isotropic variance, i.e. $\mathbf{E}_i(v) \sim N(0, \sigma_e^2 \mathbf{I}_Q)$. Since $\mathbf{Y}_i$ is whitened, we have that the mixing matrix, $\mathbf{A}_i$, is orthonormal (Hyvärinen and Oja, 2000).

At the second-level of hc-ICA, we model subject-specific spatial source signals $\mathbf{S}_i(v)$ in terms of population-level source signals modulated by covariate effects and additional between-subject random variations:

$$\mathbf{S}_i(v) = \mathbf{S}_0(v) + \sum_{p=1}^{P} x_{ip} \boldsymbol{\beta}_p(v) + \mathbf{r}_i(v), \tag{3.4}$$

where $\mathbf{S}_0(v)$ models the population-level spatial source signals of the Q statistically independent and non-Gaussian ICs; $x_{ip}, p = 1, \ldots, P$, are covariates containing subject-specific characteristics such as disease groups or demographic and biological traits; $\boldsymbol{\beta}_p(v)$ is the $Q \times 1$ covariate effects coefficients capturing the effects of the pth covariate on the Q brain networks at voxel v, and $\mathbf{r}_i(v)$ is an error term describing subject-specific variability for the IC. We assume that $\mathbf{r}_i(v) \sim N_Q(\mathbf{0}, \text{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_Q^2\})$, which allows for different levels of between-subject variability across the brain networks.

3.2.3 Source Distribution Model

To conduct inference for the hierarchical ICA model, we model the unobserved population-level spatial source signals $\mathbf{S}_0(v)$ with a source distribution. Various probability distributions have been proposed to model source signals in ICA (Hyvärinen and Oja, 2000). For example, mixtures of Gaussians (MoG) has been applied in previous work (Guo, 2011; Guo and Tang,

2013; Shi and Guo, 2016). In these works, mixtures of two or three Gaussians were typically used to model the fMRI spatial source signals to capture the background fluctuations, activation and deactivation across the voxels in the brain. Though these MoG models offer nice interpretation and tractable likelihood, they have several drawbacks including limited capacity to capture the some complex distribution patterns in the source signals and significant increase in computational load with the increase of ICs. In this paper, we propose to use a Dirichlet process mixture of Gaussians (DPM) to model the spatial source signals as an infinite mixture of Gaussians.

$$\begin{aligned} S_{0,q}(v)|K_q(v), \mu_{K_q(v)}, \phi^2_{K_q(v)}, \gamma^2_q &\sim N(\mu_{K_q(v)}, \phi^2_{K_q(v)}\gamma^2_q) \\ \mu_{K_q(v)}, \phi^2_{K_q(v)}|G &\sim G \\ G &\sim DP(MG_0), \end{aligned} \tag{3.5}$$

where $\gamma^2_q = \sigma^2_e + \sigma^2_q$, and $DP(MG_0)$ denotes a Dirichlet process with total mass parameter M and centering measure G_0 (Ferguson, 1983; Müller et al., 1996). For our purposes we use $G_0 = N\text{-}\Gamma^{-1}(0, 1, \alpha_G, \beta_G)$, a normal inverse-gamma distribution with shape α_G and scale β_G. The DPM approach performs non-parametric clustering of the source signals. Here, $\mathbf{K}$ is a matrix indicating cluster memberships; $K_q(v) = h$ if the vth voxel in the qth IC belongs to cluster h. There are several major advantages of the DPM source distribution model. It is extremely flexible and can thus capture different patterns in source intensities across the brain. Additionally, voxels in different ICs can share the same cluster membership, which eliminate the need in the previous MoG models to determine the voxel memberships over a set whose cardinality increases exponentially with the number of ICs, a primary difficulty in the current hc-ICA with MoG source modeling (Shi and Guo, 2016).

3.2.4 Modeling Covariate Effects

The neuroscience literature has shown evidence that brain networks can vary considerably due to subjects' clinical, biological and demographic characteristics. For example, neuroimaging studies have shown that neural activity and connectivity in specific functional networks are significantly associated with mental disorders and their responses to treatment regimes (Anand et al., 2005; Greicius et al., 2007; Chen et al., 2007; Sheline et al., 2009). Other studies have found activity patterns in major functional networks vary with demographic factors including age and gender (Quiton and Greenspan, 2007; Cole, Farrell, Gibson and Egan, 2010). There is a strong interest in statistical methods that reliably identify and quantify the differences in brain functional networks that are related to subjects' clinical and demographic characteristics. Neuroscience literature has shown that the network differences are generally found at limited spatial locations in the brain instead of broadly across large brain regions. A limitation of the existing methods for studying differences on ICA-derived brain networks (Shi and Guo, 2016; Wang and Guo, 2019) is that they don't take into account the sparsity in the covariate effects on brain networks. Motivated by the neuroscience findings, we propose to incorporate sparsity regularization in the modeling of covariate effects in our hierarchical ICA. We model the covariate effect maps via a horseshoe+ prior which provides a robust method at handling unknown sparsity (Carvalho et al., 2010; Bhadra et al., 2017),

$$\beta_{p,q}(v) \sim N(0, \lambda_{p,q}(v)^2 \tau^2 \gamma_q^2),$$

where τ^2 and $\lambda_{p,q}(v)^2$ are global and local shrinkage parameters, respectively. The overall level of shrinkage is controlled by τ^2, and $\lambda_{p,q}(v)^2$ applies shrinkage to individual elements

of $\boldsymbol{\beta}$. Under the horseshoe+, we place half-Cauchy priors on the shrinkage terms,

$$\begin{aligned} \lambda_{p,q}(v) &\sim \mathcal{C}^+(0, \eta_{p,q}(v)) \\ \tau &\sim \mathcal{C}^+(0,1) \\ \eta_{p,q}(v) &\sim \mathcal{C}^+(0,1), \end{aligned} \tag{3.6}$$

where $\mathcal{C}^+(0,1)$ is a half-Cauchy distribution. This approach results in a significant amount of shrinkage on non-existent/small covariate effects, and little-to-no shrinkage of large covariate effects. Compared to alternative shrinkage approaches such as a two-group model using a spike-and-slab prior (George and McCulloch, 1993), the horseshoe prior, as a one-group model, is computationally more efficient when modeling high dimensional data such as brain networks (Polson and Scott, 2010).

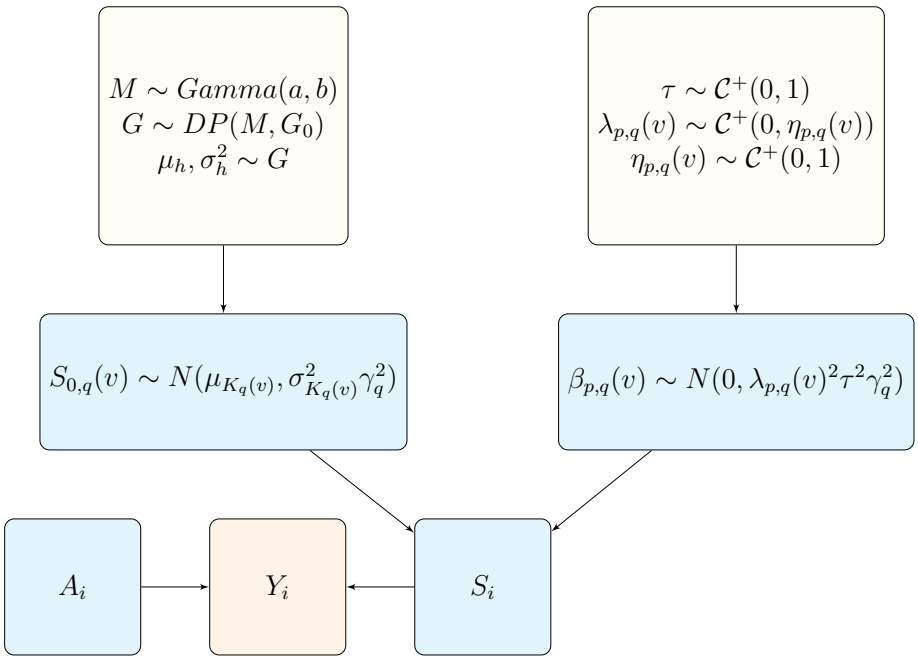


Figure 3.1: Directed graph illustrating the relationships between the model parameters. Yellow nodes correspond to hyperparameters which are updated or tuned, blue nodes correspond to parameters of interest in the ICA model, and the orange node corresponds to the observed data.

3.3 Posterior Computation

To facilitate posterior computation for SparseBayes ICA, we rewrite the model in a non-hierarchical form by collapsing the two-level models in (3.3) and (3.4) and then multiplying the orthogonal mixing matrix $\mathbf{A}_i$ on both sides to obtain the following,

$$\mathbf{A}_i'\mathbf{Y}_i(v) = \mathbf{S}_0(v) + \sum_{p=1}^{P} x_{ip}\boldsymbol{\beta}_p(v) + \boldsymbol{\epsilon}_i(v),$$

where $\boldsymbol{\epsilon}_i(v) = \mathbf{r}_i(v) + \mathbf{A}_i'\mathbf{e}_i(v)$ is the combined error across the two levels of models and we have $\boldsymbol{\epsilon}_i(v) \sim N_Q(\mathbf{0}, \text{diag}\{\gamma_1^2, \gamma_2^2, \ldots, \gamma_Q^2\})$ with $\gamma_q^2 = \sigma_e^2 + \sigma_q^2$. We specify a non-informative prior on these variance terms, i.e. $\gamma_q^2 \sim \gamma_q^{-2} d\gamma_q^2$. Furthermore, we specify a uniform distribution over a sphere as the prior distribution for each participant's mixing matrix $\mathbf{A}_i$ (Hoff, 2009).

3.3.1 MCMC algorithm

Posterior computation for the SparseBayes ICA model is carried out via MCMC and all updates can be performed via Gibbs sampling. The detailed updating steps are presented in the following.

3.3.1.1 Updating the Mixing Matrix $\mathbf{A}_i$

The posterior distribution for the ith subject's mixing matrix is given by,

$$\mathbf{A}_i | \mathbf{Y}_i, \mathbf{S}_i, \{\gamma_q\} \sim \text{BMF}(\mathbf{W}, \mathbf{Z}, \mathbf{C}), \tag{3.7}$$

where BMF denotes a Matrix-Bingham-von-Mises-Fisher Distribution with $\mathbf{W} = \mathbf{Y}_i\mathbf{Y}_i'$, $\mathbf{C} = \mathbf{Y}_i\mathbf{S}_i'$, and $\mathbf{Z} = -\frac{1}{2}\text{diag}\{\gamma_1^{-2}, \gamma_2^{-2}, \ldots, \gamma_Q^{-2}\}$. Sampling from this distribution can be carried out using the approach developed in Hoff (2009) by cycling through all pairs of columns of $\mathbf{A}_i$ indexed here by q and q' with $q \neq q'$, $q, q' = 1, \ldots Q$ and performing the following steps:

1. For the pair q, q' let $\mathbf{N}_{qq'}$ denote null space of $\mathbf{A}_{i,\cdot-[q,q']}$, the mixing matrix with columns q and q' removed. $\mathbf{N}_{qq'}$ has dimension $Q \times 2$, and the sample for the q and q'th columns of the mixing matrix can be found by finding an orthonormal rotation of this space.

2. Sample a 2×2 rotation matrix, $\mathbf{R} \sim BMF(\tilde{\mathbf{W}},\tilde{\mathbf{Z}},\tilde{\mathbf{C}})$. Here, $\tilde{\mathbf{W}} = \mathbf{N}'_{qq'}\mathbf{W}\mathbf{N}_{qq'}$, $\tilde{\mathbf{Z}} = \text{diag}\{Z_{qq}, Z_{q'q'}\}$, and $\tilde{\mathbf{C}} = \mathbf{N}'_{qq'}\mathbf{C}_{\cdot[qq']}$. The corresponding sampling density is $\mathbf{R} \propto \text{etr}\{\tilde{\mathbf{C}}'\mathbf{R} + \tilde{Z}\mathbf{R}'\tilde{\mathbf{W}}\mathbf{R}\}$, a 2-dimensional BMF distribution. We can sample from this by defining $\mathbf{R}$ in terms of a sign and an angle. Further details are provided in Hoff (2009).

3. Set $\mathbf{A}_{i,qq'} = \mathbf{N}_{qq'}\mathbf{R}$.

3.3.1.2 Updating the DPM Cluster Memberships K

Under the stick breaking representation of the DP (Sethuraman, 1994), we can write G as,

$$G = \sum_{h=1}^{\infty} w_h \delta_{\mu_h, \phi_h^2}, \tag{3.8}$$

where $w_h = \nu_h \prod_{h'=1}^{h-1}(1-\nu_{h'})$, $\nu_h \sim Beta(1, M)$, and $(\mu_h, \phi_h^2) \sim G_0$. Under this representation, the density for the $S_{0,qv}$ can be written,

$$p(S_{0,q}(v)|-) = \sum_{h=1}^{\infty} w_h f(S_{0,q}(v); \mu_h, \phi_h^2\gamma_q^2), \tag{3.9}$$

where $f(\cdot; \mu_h, \phi_h^2\gamma_q^2)$ denotes the density of a normal distribution with mean μ_h and variance $\phi_h^2\gamma_q^2$.

The parameters of the DPM can be updated using the slice sampling approach of Walker (2007), which introduces latent variables that allow us to reduce the sum in (3.9) to a finite number of terms. Following this approach, we augment the model with two steps. The first step is to introduce a latent variable for each element, $u_q(v) \sim U(0, w_{K_q(v)})$, where U denotes

a uniform distribution. The joint density for $S_{0,q}(v)$ and $u_q(v)$ can be written:

$$p(S_{0,q}(v), u_q(v)|\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\phi}^2, \gamma_q^2) = \sum_{h=1}^{h^*} I_{u_q(v)<w_h} f(S_{0,q}(v); \mu_h, \phi_h^2\gamma_q^2), \tag{3.10}$$

where $h^* = \min\{h : u_q(v) > 1 - \sum_{h'=1}^{h} w_{h'} \text{ for all } q, v\}$. The second step is to add the cluster membership variable for each voxel and IC, $K_q(v)$, where $p(K_q(v) = h|u_q(v), w_h) \propto I_{u_q(v)<w_h}$, yielding,

$$p(S_{0,q}(v), u_q(v), K_q(v)|\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\phi}^2, \gamma_q^2) = I_{u_q(v)<w_{K_q(v)}} f(S_{0,q}(v); \mu_{K_q(v)}, \phi_{K_q(v)}^2\gamma_q^2), \tag{3.11}$$

Note that this augmented model simplifies computation dramatically, since each element of S_0 belongs to a single cluster, and thus once we know the cluster memberships $\{K_q(v)\}$ the typical posterior inference for Normal distributions can be applied.

Updating the latent terms is straightforward. First, we sample the stick breaking weights from their conditional posterior $\nu_h|\mathbf{K}, M \sim Beta(1 + n_h, M + \sum_{h'=h+1}^{\infty} n_{h'})$, where n_h is the number of elements of $\mathbf{K}$ equal to h. Then we sample the $u_q(v)$ from $u_q(v)|\mathbf{w}, K_q(v) \sim U(0, w_{K_q(v)})$. Finally, the cluster memberships $\mathbf{K}$ can be sampled from their posterior distributions, which are given by:

$$p(K_q(v) = h|\mathbf{u}, \boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\phi}^2) = \frac{I_{u_q(v)<w_h} f(S_{0,q}(v); \mu_h, \phi_h^2\gamma_q^2)}{\sum_{h'=1}^{h^*} I_{u_q(v)<w'_h} f(S_{0,q}(v); \mu_{h'}, \phi_{h'}^2\gamma_q^2)}. \tag{3.12}$$

3.3.1.3 Updating Population-level IC maps, $\mathbf{S}_0$, and covariate effects, $\boldsymbol{\beta}$

Conditioned on the cluster memberships $\mathbf{K}$, the posterior distribution for the population-level IC maps and covariate effects is multivariate normal. Let $\mathbf{X}^* = [\mathbf{x}_1', \mathbf{x}_2', \ldots, \mathbf{x}_N']'$, $\boldsymbol{\beta}_q^*(v) = [S_{0,q}(v), \beta_{1,q}(v), \ldots, \beta_{P,q}(v)]'$, and $\mathbf{Y}_q^*(v) = [\{\mathbf{A}_1'\mathbf{Y}_1(v)\}_q, \{\mathbf{A}_2'\mathbf{Y}_2(v)\}_q, \ldots, \{\mathbf{A}_N'\mathbf{Y}_N(v)\}_q]'$

then,

$$\boldsymbol{\beta}_q^*(v)|\mathbf{X}^*, \mathbf{Y}_q(v)^*, \gamma_q^2, \tau^{-2}, \boldsymbol{\lambda}, \boldsymbol{\sigma^2}, \boldsymbol{\mu}, \boldsymbol{\phi}^2, K_q(v) \sim N_{P+1}(\boldsymbol{\mu}_{\beta^*,q}(v), \boldsymbol{\Sigma}_{\beta^*,q}(v)), \tag{3.13}$$

where

$$\begin{aligned}
\boldsymbol{\Sigma}_{\beta^*,q}(v) &= \left(\mathbf{X}^{*'}\mathbf{X}^*\gamma_q^{-2} + \gamma_q^{-2}\text{diag}\{\sigma_{K_q(v)}^{-2}, \tau^{-2}\lambda_{1,q}(v)^{-2}, \ldots, \tau^{-2}\lambda_{P,q}(v)^{-2}\}\right)^{-1}, \\
\boldsymbol{\mu}_{\beta^*,q}(v) &= \boldsymbol{\Sigma}_{\beta^*,q}(v)\left(\mathbf{X}^{*'}\mathbf{Y}_q^*(v)\gamma_q^{-2} + \left[\frac{\mu_{K_q(v)}}{\phi_{K_q(v)}^2}, \mathbf{0}_P'\right]'\right).
\end{aligned} \tag{3.14}$$

3.3.1.4 DPM Cluster Parameters

Conditioned on the cluster memberships, $\mathbf{K}$, and the population-level spatial maps $\mathbf{S}_0$, we can sample the mean and variance of each DPM cluster from their posterior distribution. First, the posterior for the cluster variances is an inverse gamma,

$$\phi_h^2 \sim IG(a_h, b_h), \tag{3.15}$$

where $a_h = \alpha_G + \frac{n_h}{2} - 1$, $b_h = \beta_G + \frac{1}{2}\sum_{q=1}^{Q}\sum_{v=1}^{V}\frac{(S_{0,q}(v)-\bar{S}_{h,q})^2+S_{0,q}(v)^2}{\gamma_q^2}I_{K_q(v)=h}$, and $\bar{S}_{h,q}$ is the mean of all elements of $\mathbf{S}_{0,q}$ belonging to cluster h.

The cluster means have a normal posterior distribution,

$$\mu_h|\phi_h \sim N(\mu_h^*, \phi_h^{2^*}), \tag{3.16}$$

$$\text{with} \quad \mu_h^* = \frac{\sum_{q=1}^{Q}\frac{1}{\gamma_q^2}\sum_{v=1}^{V}S_{0,q}(v)I_{K_q(v)=h}}{1+\sum_{q=1}^{Q}\frac{n_{h,q}}{\gamma_q^2}} \quad \text{and} \quad \phi_h^{2^*} = \frac{\phi_h^2}{1+\sum_{q=1}^{Q}\frac{n_{h,q}}{\gamma_q^2}},$$

where $n_{h,q}$ is the number of voxels in IC q belonging to cluster h. Note that the posterior distribution (3.16) is essentially a weighted version of the typical normal posterior update, where the spatial source signals from ICs with lower variance are weighted more heavily. This is an advantage by having the IC-specific variance in our proposed DPM source distribution

model (3.5).

3.3.1.5 Total Mass Parameter

The total mass, M, controls how many clusters are formed. With a Gamma(a, b) prior on M, the posterior for M can be found following Escobar and West (1995). First, draw a random uniform variable $z \sim U(0, 1)$. Then,

$$M|a, h, b, z, p^* \sim \begin{cases} Gamma(a + h^*, b - \log(z)) & \text{w.p. } p^* \\ Gamma(a + h^* - 1, b - \log(z)) & \text{w.p. } 1 - p^* \end{cases}, \tag{3.17}$$

where

$$p^* = \frac{\frac{a+h^*-1}{QV[b-\log(z)]}}{\frac{a+h^*-1}{QV[b-\log(z)]} + 1}.$$

3.3.1.6 Horseshoe Hyperparameters

We update the horseshoe hyperparameters using the approach developed in Makalic and Schmidt (2015). In particular, using the representation of the half-Cauchy distribution as a scale mixture of inverse-gamma distributions we can re-write (3.6) as

$$
\begin{aligned}
\beta_{p,q}(v) &\sim N(0, \lambda_{p,q}(v)^2\tau^2\gamma_q^2) \\
\lambda_{p,q}(v)^2|\eta_{p,q}(v)\nu_{p,q}(v) &\sim IG(\frac{1}{2}, \frac{1}{\nu_{p,q}(v)}) \\
\nu_{p,q}(v) &\sim IG(\frac{1}{2}, \eta_{p,q}(v)^2) \\
\eta_{p,q}(v)^2 &\sim IG(\frac{1}{2}, \frac{1}{c_{p,q}(v)}), \\
c_{p,q}(v) &\sim IG(\frac{1}{2}, 1) \\
\tau^2|\xi &\sim IG(\frac{1}{2}, \frac{1}{\xi}) \\
\xi &\sim IG(\frac{1}{2}, 1).
\end{aligned} \tag{3.18}
$$

The corresponding posterior distributions are:

$$
\begin{aligned}
\tau^2|\xi, \{\lambda_{p,q}(v)\}, \{\gamma_q\}, \boldsymbol{\beta} &\sim IG(\frac{1+PQV}{2}, \frac{1}{\xi} + \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{v=1}^{V}\frac{1}{2\lambda_{p,q}(v)^2\gamma_q^2}\beta_{p,q}(v)^2) \\
\xi|\tau^2 &\sim IG(\frac{1}{2}, \frac{1}{\tau^2}) \\
\lambda_{p,q}(v)^2|\nu_{p,q}(v), \tau, \gamma_q, \beta_{p,q}(v) &\sim IG(1, \frac{1}{\nu_{p,q}(v)} + \frac{1}{2\tau^2\gamma_q^2}\beta_{p,q}(v)^2) \\
\nu_{p,q}(v)|\lambda_{p,q}(v), \eta_{p,q}(v) &\sim IG(1, \frac{1}{\lambda_{p,q}(v)^2} + \frac{1}{\eta_{p,q}(v)^2}) \\
\eta_{p,q}(v)^2|c_{p,q}(v), \nu_{p,q}(v) &\sim IG(1, \frac{1}{c_{p,q}(v)} + \frac{1}{\nu_{p,q}(v)}) \\
c_{p,q}(v)|\eta_{p,q}(v) &\sim IG(1, 1 + \frac{1}{\eta_{p,q}(v)^2}).
\end{aligned} \tag{3.19}
$$

3.3.1.7 IC-specific Variance Terms

The posterior for each variance term γ_q^2, $q = 1, \ldots, Q$ is inverse gamma:

$$\gamma_q^2 \sim IG\Bigg(\frac{(N+P+1)V}{2}, \sum_{i=1}^{N}\sum_{v=1}^{V}\left[\frac{\left(Y_{i,q}(v)-\hat{Y}_{i,q}(v)\right)^2}{2}\right] + \sum_{v=1}^{V}\left[\frac{(S_{0,q}(v)-\mu_{K_q(v)})^2}{2\phi^2_{K_q(v)}} + \sum_{p=1}^{P}\frac{\beta_{p,q}(v)^2}{\tau^2\lambda_{p,q}(v)^2}\right]\Bigg), \tag{3.20}$$

where $\hat{Y}_{i,q}(v)$ is the q,vth element of $\mathbf{A}_i\mathbf{S}_i$.

3.3.2 Inference for Covariate Effects

We focus here on posterior inference for the population-level spatial source signals $\mathbf{S}_0(v)$ and the covariate effects $\boldsymbol{\beta}$. Denote the jth sample of $\mathbf{S}_0$ from its posterior distribution by $\mathbf{S}_0(v)^{(j)}$. Similarly let $\beta_{p,q}(v)^{(j)}$ be the jth sample for the pth covariate effect at IC q and voxel v. The point estimates are given by $\hat{\mathbf{S}}_0(v) = \frac{1}{J}\sum_{j=1}^{J}\mathbf{S}_0(v)^{(j)}$ and $\hat{\beta}_{p,q}(v) = \frac{1}{J}\sum_{j=1}^{J}\beta_{p,q}(v)^{(j)}$ where J is the number of MCMC samples. We use Bayesian credible intervals to determine significance of the covariate effects. The proportion of the samples on either side of zero can be used to construct a p-value as follows. Let $I(\beta_{p,q}(v)^{(j)} > 0) = 1$ if the jth sample from the posterior is positive, 0 otherwise. Denote the credible interval-based p-value for the pth covariate at the qth IC and vth voxel by $CIP_{p,q}(v)$. Then,

$$CIP_{p,q}(v) = 2 \times \min\left\{\frac{1}{J}\sum_{j=1}^{J} I(\beta_{p,q}(v)^{(j)} > 0), 1 - \frac{1}{J}\sum_{j=1}^{J} I(\beta_{p,q}(v)^{(j)} > 0)\right\}. \tag{3.21}$$

Throughout the rest of this manuscript, we determine the significant voxels to be those for which $CIP_{p,q}(v) < 0.05$.

3.4 Simulation Studies

In this section, we evaluate the performance of the proposed SparseBayes ICA approach through extensive simulation studies. We conducted a full factorial experiment in which we varied the sample size, number of ICs, and variability in the data, where 50 replications were generated for each combination of factor levels.

We designed the spatial IC maps and covariate effects to mimic real brain networks. Specifically, for each IC source signal, we created 3D spatial maps using the MNI space brain re-sampled to have dimension $36 \times 46 \times 26$. We then generated two covariates for each subject from independent distributions. The spatial distribution of the ICs, i.e. S_0, and covariate effects maps for the first three networks are illustrated in Figure 3.2 for illustration purposes. For the spatial distribution of the qth IC ($q = 1, \ldots, Q$), we first generated a set of voxel-specific intensity values from normal distributions with IC-specific parameters (see Appendix B for detailed settings). We then spatially distributed the intensity values in the activated region of the IC in the way that values in neighboring voxels tend to be similar to each other and that spatial source signals are higher in the central regions in an IC map and gradually decreases towards the outer regions in an IC. This procedure aims to generate IC maps that mimic the activation patterns observed in real-world brain networks. Similarly, the intensities for the covariate effects were drawn from independent normal distributions with covariate-specific parameters (see Appendix) and spatially distributed in the IC maps following a similar procedure as described above.

We constructed the subject-specific spatial components, $\mathbf{S}_i$, as a combination of the population-level signals, covariate effects and subject-specific random effects. The subject-specific random effects, i.e., $\mathbf{r}_i(v)$, were generated from a zero-mean Gaussian distribution with the variance $\sigma_q^2 \sim Gamma(b, 1)$. We considered three levels of between-subject variability: low ($b = 1.5$), medium ($b = 3$) and high ($b = 5$). Subject-specific mixing matrices were generated based on time courses from real imaging studies. Finally, we generated observed fMRI series based on the subject-specific spatial maps and mixing matrices and Gaussian

random noise with a standard deviation of σ_e^2 were added to the mixtures. We considered two noise levels with a low ($\sigma_e^2 = 1.5$) and a high setting ($\sigma_e^2 = 3.0$).

We compare to two advanced TC-GICA approaches, as well as the current state-of-the-art hc-ICA approach. The first TC-GICA approach is dual regression or spatio-temporal regression (STR), which is one of most commonly used methods in group ICA. The second TC-GICA approach is the GICA3 method (Erhardt et al., 2011), which has been shown to be a top TC-GICA method for estimating subject-specific ICs. The two TC-GICA approaches estimate covariate effects by performing post-hoc regression analysis on the back-reconstructed individual-level ICs maps (Beckmann et al., 2009). In comparison to the two-step ad-hoc procedure taken by the TC-GICA methods, the hc-ICA approach (Shi and Guo, 2016) is the current state-of-art ICA method for testing covariate effects by formally incorporating and estimating covariate effects in group ICA decomposition. To obtain multiplicity corrected p-values for the TC-GICA approaches and hc-ICA, we then perform FDR correction using Benjamini and Yekutieli method (Benjamini and Yekutieli, 2001; Genovese et al., 2002). For the SparseBayes ICA approach, we apply the proposed credible intervals described in Section 3.3.2 to assess significance of covariate effects. The proposed horseshoe priors provides sparsity regularization for estimating covariate effects which helps achieve adequate FDR control (van der Pas et al., 2017).

We calculate the precision-recall curves for identifying true covariate effects for each replicate over a range of significance thresholds. Figures 3.3 and 3.4 display boxplots of the AUC for each method and combination of simulation settings for $Q = 3$ and $Q = 6$, respectively. From the plots, it is clear that the proposed SparseBayes ICA approach consistently outperforms the other two methods in accurately detecting the covariate effects across various simulation settings. In particular, SparseBayes ICA has considerable advantages over the other methods in small sample sizes. The SparseBayes ICA method has more power to detect active covariate effects while still maintaining a lower rate of false positives.

Figure 3.2: Spatial extent for the S_0 maps and covariate effects used in the simulation experiment.

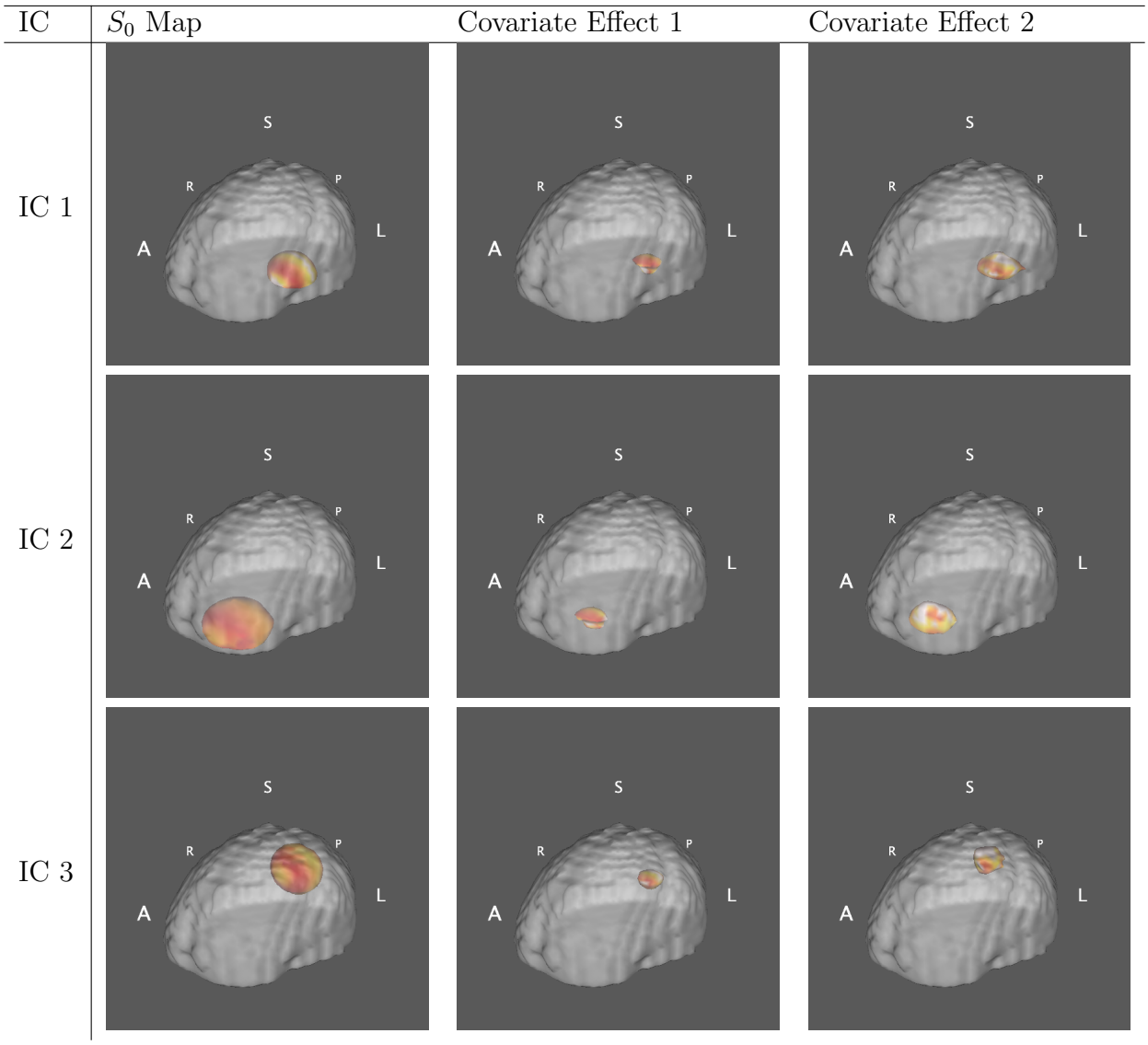


Figure 3.3: Area under the precision recall curve for identifying significant covariate effects 3 components.

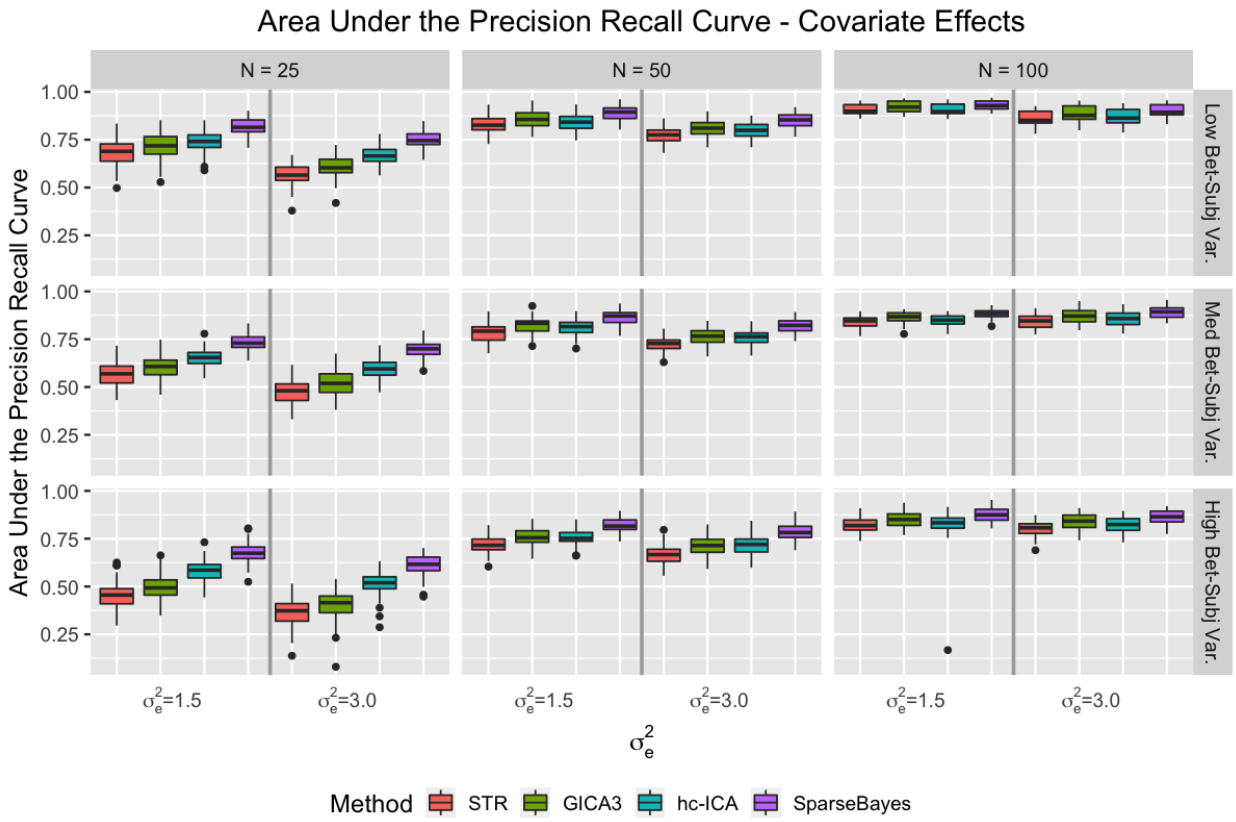
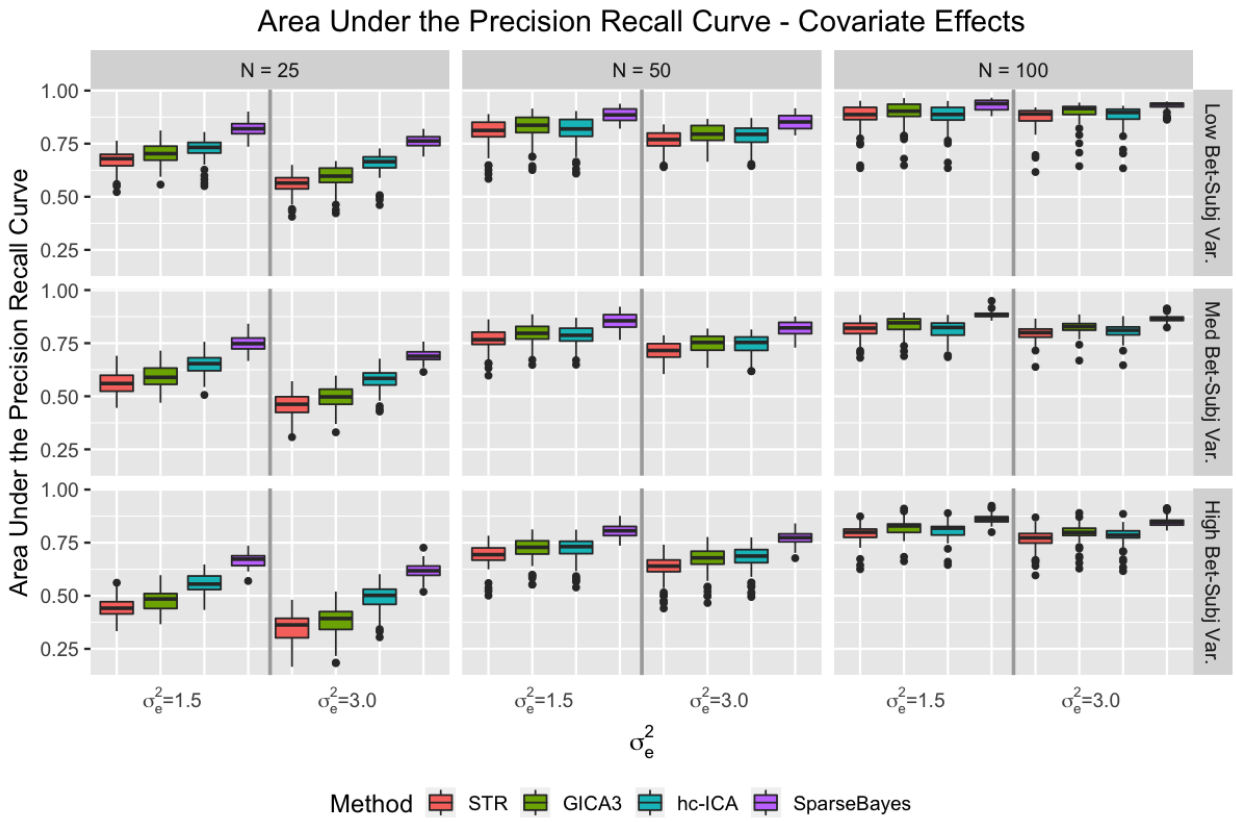


Figure 3.4: Area under the precision recall curve for identifying significant covariate effects, 6 components.



3.5 Zen Meditation Study

We apply the SparseBayes ICA to investigate differences in brain networks related to Zen meditation practice.

3.5.1 Data and Preprocessing

In a Zen meditation study (Pagnoni, 2012), 24 right-handed (one ambidextrous), English speaking participants were recruited for the study. The meditation (MEDT) group consisted of 12 participants with at least 3 years of daily meditation practice in the Zen objectless tradition (mean years of practice = 8.7, SD = 6.5 years). The control group (CTRL) consisted of 12 gender, age, and education matched controls.

T1-weighted high resolution anatomical scans (MPRAGE, 176 sagittal slices, 1mm isotropic voxel size) and resting-state fMRI images (echo-planar, 200 volumes, 35 axial slices, voxel size: 3 mm isotropic, TR = 2.35 s, TE = 30 ms) were collected with a 3.0 Tesla Siemens Magnetom Trio Scanner. The resting-state run lasted for approximately 8 minutes. During the scan, participants were instructed to fixate on a cross directly in front of their view. While fixating, participants were to focus on their breathing and avoid allowing their mind to wander to any thoughts or physical sensations. Standard fMRI preprocessing pipelines were applied, including slice time correction, motion correction, image alignment and registration to a common MNI space, smoothing the functional images with an 8mm kernel, and removal of low-frequency drifts. Further details are provided in Pagnoni (2012) and Kemmer et al. (2015).

One week before the scanning session, participants completed a sustained attention CANTAB computerized psychological test (Sahakian and Owen, 1992). Performance was measured in terms of sensitivity to a target sequences, quantified as A-prime score. We control for this score when investigating differences between meditators and controls by including it in our model.

3.5.2 SparseBayes ICA Modeling Procedure

We applied the proposed SparseBayes ICA method to analyze the Zen meditation data. We decomposed the resting-state data into 15 ICs, and included two covariates in the model: MEDT status (Zen vs. Control) and A-prime score (transformed to Z-scores). We used 3000 burn-in iterations and 5000 MCMC iterations. We also fit this model using TC-GICA with GICA3 back reconstruction and hc-ICA. FDR method was applied for TC-GICA and hc-ICA, while the credible interval-based procedure from Section 3.3.2 was used for SparseBayes ICA.

Previous work with populations of meditators has identified the default mode network (DMN) as a primary network of interest (Brewer et al., 2011; Garrison et al., 2015). In the results that follow, we focus on the DMN network for investigating the effect of meditation training.

3.5.3 Results

Each of the three approaches was able to identify a network corresponding to the DMN. Figure 3.5 displays the population level spatial map obtained using each method. Compared to the hc-ICA and TC-GICA results, the SparseBayes ICA derived DMN better complies with the spatial distribution of the network established in the neuroscience literature. Specifically the SparseBayes ICA DMN network includes all the key areas in DMN including the posterior cingulate cortex (PCC), angular gyri and the prefrontal cortex (Alves et al., 2019). In comparison, the DMN networks estimated by the other two methods mainly include the PCC regions with weaker signals for the angular gyri and no inclusion of the prefrontal cortex.

Figure 3.6 shows the regions where the methods identify significant differences between meditation and control participants. SparseBayes ICA finds significant MEDT differences in the ventral posterior cingulate, the angular gyrus, the visuo-motor cortex, and the dorso-lateral PFC. In comparison, hc-ICA only finds significant effects in the visuo-motor cortex

and TC-GICA is unable to find any significant effects after performing FDR control.

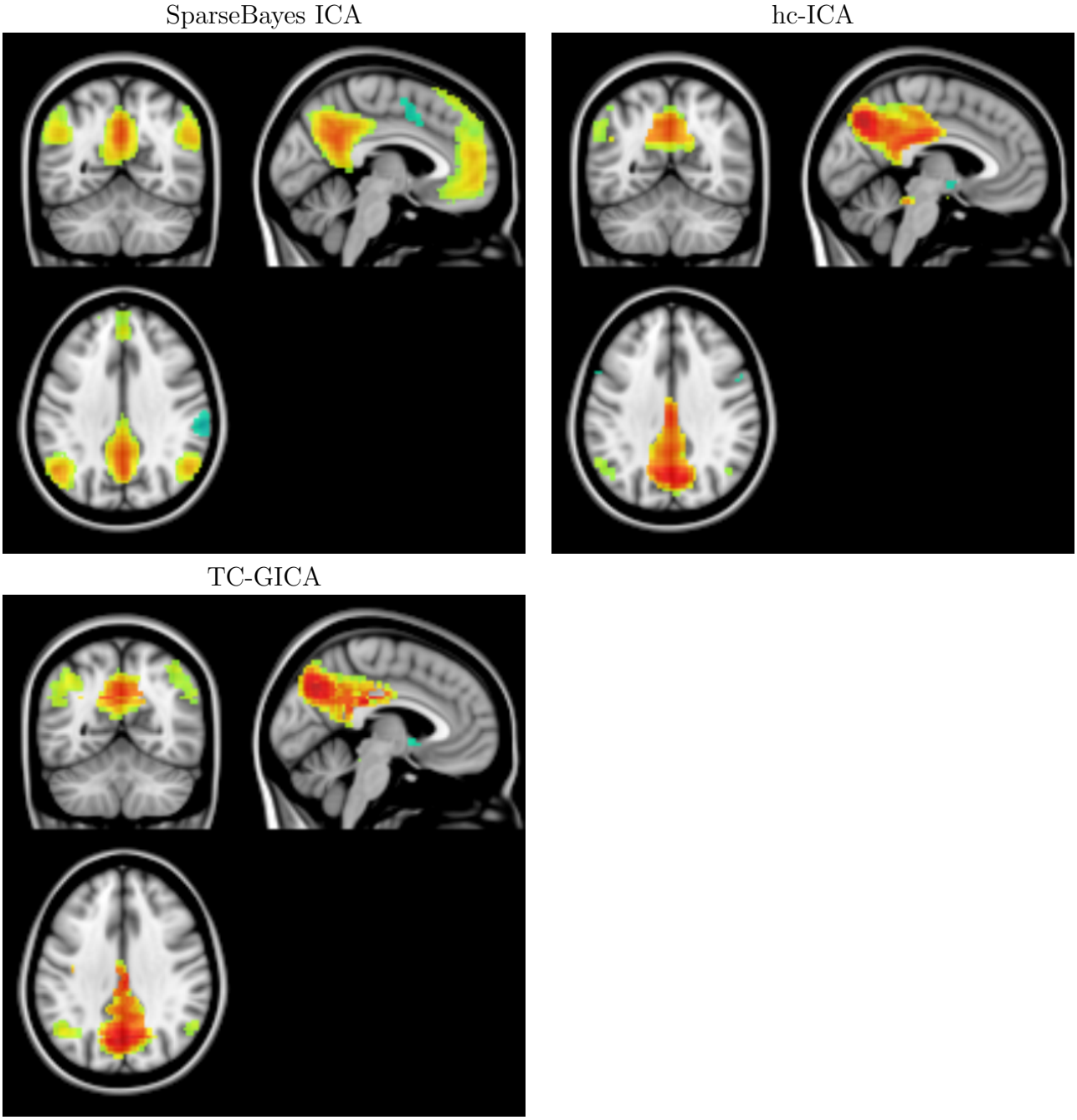
The regions identified using SparseBayes ICA are consistent with prior research into meditation and the DMN. In particular, Brewer et al. (2011) showed reductions in DMN activity at rest in meditators relative to healthy controls in the PCC and and dorsolateral PFC. Reductions in connectivity in these two primary DMN regions have also been found in meditators relative to healthy controls during meditation tasks, as opposed to the resting state task investigated here (Brewer et al., 2011; Garrison et al., 2015). It has been suggested that the lowered DMN activation in meditators is consistent with a reduction in mind-wandering and self-referential thought.

Our findings also suggest that meditators exhibit stronger DMN connectivity in several regions relative to controls. The largest meditation effect in terms of spatial extent found by SparseBayes ICA is a positive region in the superior parietal lobule. Taren et al. (2017) also reported an increase in activation in the superior parietal lobule after mindfulness meditation training, specifically an increase in the connectivity between the superior parietal lobule and the dlPFC. A smaller spatial effect was also found in this region using hc-ICA. Moreover, we identified a second, smaller area of increased activation in meditators located in the inferior parietal lobule, which is consistent with the findings in Taylor et al. (2013).

3.6 Discussion

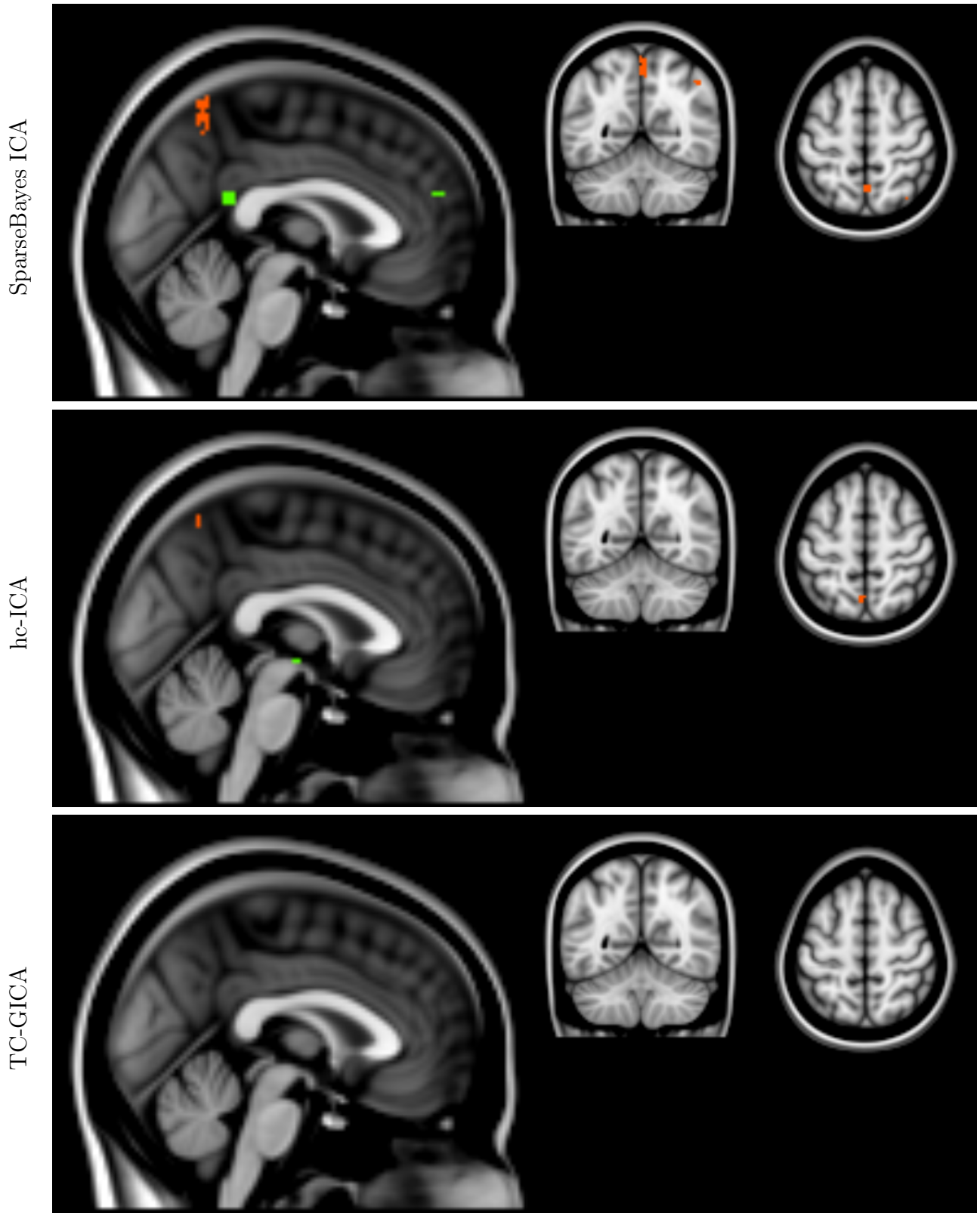
In this work, we introduced a fully Bayesian approach to hierarchical independent component analysis. This method offers several significant improvements over existing group ICA approaches. In particular, it provides a solution to modeling the population-level spatial maps and incorporates sparsity regularization in estimating covariate effects in the ICA decomposition. Through simulation, we showed that the method has better performance in detecting significant covariate effects. We used our SparseBayes ICA approach to study individual- and group-level differences in the DMN between meditators and healthy controls,

Figure 3.5: Population-level spatial maps corresponding to the DMN as identified by Sparse-Bayes ICA, hc-ICA, and TC-GICA with GICA3.



The figure displays Z-transformed source signal intensity. A threshold of $|Z| \geq 1.64$ was used to mask the spatial maps.

Figure 3.6: The differences between meditation and control subjects in the DMN.



Orange voxels denote effects with a positive sign and green voxels denote effects with a negative sign.

and showed that our proposed method was able to find biologically meaningful differences while other approaches could not.

Chapter 4

Repeated Measures Sparse Bayesian Independent Component Analysis

4.1 Introduction

Human brain function can be decomposed into collections of spatially distinct regions exhibiting highly correlated neural activation patterns. These collections of regions are known as brain functional networks, and they have well-understood biological meanings. These networks are known to exhibit differences in spatial extent (the size of the network) and intensity across individuals. This naturally has led to great interest in understanding how different clinical and demographic variables, i.e. covariates, can influence the brain networks. Many studies have been conducted to assess this in various populations, most of which are cross-sectional studies conducted at a single site.

In recent years there has been an increase in large-scale neuroimaging studies. These studies involve scanning sessions from multiple data collection sites, and often involve repeated imaging sessions over the course of several years. Some notable example include the Adolecent Brain Cognitive Development study (Jernigan et al., 2018), the Alzheimer's Disease Neuroimaging Initiative (ADNI) (Mueller et al., 2005), and the Human Connectome Project (Van Essen et al., 2013). Such studies offer exciting opportunities to study the brain network in different populations. In particular, these studies are well positioned to examine how covariates such as disease status influence brain function. With the increase in the number of multi-center, longitudinal studies, statistical techniques are needed that are tailored to the unique challenges posed by these data.

One primary question in multi-center, longitudinal imaging studies is how to best accommodate additional variability due to center and longitudinal effects. Different data collection sites may have different scanner brands, head coils, or any number of other features which can affect the fMRI data that is collected. Recent work has shown that these center effects can influence the resulting findings of the study in non-trivial ways. For example, (Yu et al., 2018) identified significant site effects on several brain connectivity measures. Similarly, the brain networks from repeated scans on the same individuals are clearly not independent, and it is important to incorporate this into any subsequent data analysis. In this work, we

will focus on how to solve these problems in the context of independent component analysis (ICA).

ICA is the most commonly used approach for studying the brain network. ICA separates the observed fMRI time courses into a set of brain maps (spatial components) and corresponding temporal patterns. However, to date there has been little work on how to incorporate multi-center longitudinal measurements in the ICA model, which is the focus of this work.

The standard approach for group-level ICA analysis is temporal concatenation group ICA (TC-GICA). In TC-GICA, each subject's data is stacked along the time domain, and then an ICA analysis is performed on a dimension-reduced version of the stacked data (Calhoun et al., 2001). Subject-specific brain networks can be obtained by performing back-reconstruction from the group level maps. When covariate effects are of interest, this leads to an ad-hoc two stage dual regression approach where the estimated subject maps are used as the outcome in a regression model with the covariates as predictors. Such an approach could potentially be extended to multi-center and longitudinal data by implementing some form of data harmonization.

One of the leading techniques for data harmonization is ComBat (Johnson et al., 2007). ComBat techniques have recently been growing in popularity in the neuroimaging field. For example, (Yu et al., 2018) used ComBat to correct data collection site effects in connectivity matrices and graph metrics. Similarly, Fortin et al. (2017) used ComBat to harmonize DTI measurements across different sites. Recently, Beer et al. (2020) developed a longitudinal form of ComBat termed longComBat, which could potentially be applied to harmonize data in longitudinal multi-center data sets.

While ComBat approaches have proven very effective in several imaging situations, they rely on performing correction of derived features such as connectivity matrices or cortical thickness measures. In many neuroimaging studies this is not an issue, since the derived features will be calculated on a subject-by-subject level and then harmonized. However, for

group ICA purposes, these derived features would be the estimates for the subject-level brain networks, which are back-reconstructed from a group level ICA analysis. This means that we will be unable to correct for batch effects *during* the ICA decomposition. Two-stage dual regression methods have already been shown to suffer from reduced power in many cases due to the lack of information about variation in the subject-level ICA decomposition when estimating the covariate effects (Shi and Guo, 2016; Lukemire, Wang, Verma and Guo, 2020). Similarly, data harmonization will be limited in what it can achieve due to the dependency on the back-reconstructed spatial maps. In light of these issues, we propose the use of a hierarchical model to incorporate multi-center and longitudinal measurements.

In recent years there has been significant development of hierarchical covariate-adjusted ICA (hc-ICA) techniques for identifying individual differences in brain networks due to clinical and demographic covariates (Shi and Guo, 2016; Wang and Guo, 2019; Lukemire, Wang, Verma and Guo, 2020). These hc-ICA methods model the individual-level brain networks as a function of a shared component among all subjects and individual-level covariate effects. This provides a principled statistical approach to estimating covariate effects, performing hypothesis testing, and examining the brain network for different sub-populations such as healthy and diseased groups.

In Chapter 3, we proposed a fully Bayesian variant of the hc-ICA model that incorporates sparsity into the covariate effect estimates termed SparseBayes ICA. In this work, we introduce a general framework of repeated measures Sparse Bayesian ICA (RM-SparseBayes ICA) that can perform ICA decomposition of multi-center and/or longitudinal data. To the best of our knowledge, this is the first such ICA approach. Our method models the multi-center, longitudinal nature of the data as follows.

First, our approach uses random intercept terms in the hierarchical model to introduce dependencies among individuals' brain networks from the same data collection center. These terms allow us to capture additional variation in the spatial components that is due to data collection site effects. We include fixed effects in the model for known sources of potential

variability such as scanner model. This has the advantage of pooling formation across centers with the same values for these covariates (e.g. same scanner model) when estimating effects on the brain network.

To incorporate longitudinal measurements we introduce subject specific random terms. In particular, we model the covariate effects as fixed, with investigator-determined random slopes for quantities that vary at the subject level. This allows us to introduce a significant amount of flexibility into the model describing the spatial components over time. Moreover, by adopting a fixed set of effects for the covariates, we are able to pool information across visits to estimate the covariate effects.

Our proposed RM-SparseBayes ICA approach provides a flexible, general framework for ICA decomposition of data from multiple centers, longitudinal data, or both. Taken together, the inclusion of these two sets of random effects allows us to incorporate information about longitudinal measurements and/or multiple centers directly into the ICA decomposition, improving our ability to estimate the subject-level brain networks of interest. Additionally, our proposed RM-SparseBayes ICA approach uses shrinkage priors to model the fixed individual- and center-effects, which results in improved ability to detect significant effects on the brain network, as we will show in simulations.

This paper proceeds as follows. In Section 2, we introduce the RM-SparseBayes ICA model and discuss how we model the center and individual effects in a hierarchical model. Section 3 describes the steps in the posterior computation. In Section 4 we conduct a simulation study to evaluate the performance of our proposed approach. We then apply the RM-SparseBayes ICA approach to the ADNI2 longitudinal fMRI data in Section 5. Finally, in Section 6 we provide some concluding remarks.

4.2 Methods

4.2.1 Preprocessing

We perform standard preprocessing steps prior to the ICA decomposition including centering, dimension reduction, and whitening of the fMRI data (Hyvärinen and Oja, 2000). Let $\mathbf{Y}_{i,c,j}^{(centered)}(v) \in \mathbb{R}^T$ be the centered time series recorded for subject i from center c at visit j at voxel v, where T is the total number of fMRI scans during the imaging session and V is the number of voxels in the brain. Let $\mathbf{Y}_{i,c,j}^{(centered.)} = [\mathbf{Y}_{i,c,j}^{(centered.)}(1), ..., \mathbf{Y}_{i,c,j}^{(centered.)}(V)]$ be the corresponding $T \times V$ fMRI data matrix encompassing the data from the entire brain. We perform the following dimension reduction and prewhitening of the data (Shi and Guo, 2016):

$$\mathbf{Y}_{i,c,j} = (\mathbf{\Lambda}_{i,c,j,Q} - \tilde{\sigma}^2_{i,c,j,Q}\mathbf{I}_Q)^{-\frac{1}{2}}\mathbf{U}'_{i,c,j,Q}\mathbf{Y}_{i,c,j}^{(centered.)}, \tag{4.1}$$

where $\mathbf{U}_{i,c,j,Q}$ and $\mathbf{\Lambda}_{i,c,j,q,Q}$ contain the first Q eigenvectors and eigenvalues from the singular value decomposition of $\mathbf{Y}_{i,c,j}^{(centered.)}$. $\tilde{\sigma}^2_{i,c,j,Q}$ is the residual variance and is estimated by the average of the smallest $T - Q$ eigenvalues that are not included in $\mathbf{\Lambda}_{i,c,j,Q}$. Here, Q is the number of independent components in the model. The number of components should be determined based on the level of granularity desired by the investigator – low model orders tend to result in large-scale brain networks, while high model orders tend to split the large-scale brain networks into component pieces with specific functionality. Throughout this work, we present the methods for the prewhitened data, $\mathbf{Y}_{i,c,j} = [\mathbf{Y}_{i,c,j}(1), ..., \mathbf{Y}_{i,c,j}(V)]$ ($i = 1, ..., n_c$, $c = 1, \ldots, C$, $j = 1, \ldots, J$), which is of dimension $Q \times V$.

4.2.2 RM-SparseBayes ICA Model

We next present our proposed RM-SparseBayes ICA method for investigating individual differences in brain networks using longitudinal data collected across multiple subjects and

centers. Our approach builds on the hierarchical ICA models developed in Shi and Guo (2016), Wang and Guo (2019), and our Chapter 3. For an analysis with C centers and J visits, let $\mathbf{y}_{i,c,j}(v)$ be the preprocessed data for subject i, center c, visit j, and voxel v. The first level of the model is a spatial ICA model:

$$\mathbf{y}_{i,c,j}(v) = \mathbf{A}_{i,c,j}\mathbf{S}_{i,c,j}(v) + \mathbf{r}_{i,c,j}(v), \tag{4.2}$$

where $\mathbf{A}_{i,c,j}$ is the ith subject from center c's $Q \times Q$ mixing matrix at visit j, which is orthonormal due to the prewhitening step (Hyvärinen and Oja, 2000). The spatial source signals at the vth brain voxel for subject i and center c at visit j are contained in $\mathbf{S}_{i,c,j}(v) = [s_{i,c,j,1}(v), ..., s_{i,c,j,Q}(v)]'$. The Q elements of $\mathbf{S}_{i,c,j}(v)$ are assumed to be independent and non-Gaussian. $\mathbf{r}_{i,c,j}(v) = [r_{i,c,j,1}(v), ..., r_{i,c,j,Q}(v)]'$ is a $Q \times 1$ noise term. Following previous work (Beckmann and Smith, 2004; Guo, 2011; Shi and Guo, 2016), we assume that the noise term is independent across voxels and follows a Gaussian distribution with isotropic variance, i.e. $\mathbf{r}_{i,c,j}(v) \sim N(0, \sigma_y^2\mathbf{I}_Q)$.

At the second level, our proposed approach models each subject's spatial maps as a linear combination of a shared component across all subjects, random intercepts corresponding to center, a set of potentially time-varying covariate effects, and subject specific random effects. We write the model for IC q, $q = 1, \ldots, Q$, across visits as:

$$\begin{aligned}\mathbf{S}_{i,c,q}(v) = {} & S_{0,q}(v)\mathbf{1}_J + L_{q,c}(v)\mathbf{1}_J + \mathbf{X}_{i,c}^{(center)}\boldsymbol{\beta}_q^{(center)}(v) \\ & + \mathbf{X}_{i,c}^{(ind.)}\boldsymbol{\beta}_q^{(ind.)}(v) + \mathbf{Z}_{i,c}\mathbf{B}_{i,c,q}(v) + \mathbf{E}_{i,c,q}(v),\end{aligned} \tag{4.3}$$

where $S_{0,q}(v)$ is the part of brain network q that is common across all subjects at voxel v, $L_{q,c}(v)\mathbf{1}_J$ is the center specific random intercept for center c and component q, $\mathbf{X}_{i,c}^{(ind.)}$ is a $J \times P_1$ matrix of covariate values where each row contains the covariate settings for subject i in center c at visit j. Similarly, $\mathbf{X}_{i,c}^{(center)}$ is a $J \times P_2$ matrix of center effects. $\mathbf{B}_{i,c,q}(v)$ is a

vector of random effects for subject i at center c and voxel v. Finally, $\mathbf{E}_{i,c,q}(v)$ is the error in estimating the subject and visit specific spatial maps. We assume that the errors across different components are independent, with $\mathbf{E}_{i,c,q}(v) \sim N(0, \sigma_q^2)$.

Following our earlier work, we model the source distribution under a Dirichlet process mixture of Gaussians as,

$$\begin{aligned} S_{0,q}(v)|K_q(v), \mu_{K_q(v)}, \phi^2_{K_q(v)}, \sigma_q^2 &\sim N(\mu_{K_q(v)}, \phi^2_{K_q(v)}\sigma_q^2) \\ \mu_{K_q(v)}, \phi^2_{K_q(v)}|G &\sim G \\ G &\sim DP(MG_0), \end{aligned} \tag{4.4}$$

where $\gamma_q^2 = \sigma_y^2 + \sigma_q^2$, $K_q(v)$ is a cluster membership indicator, and $DP(MG_0)$ denotes a Dirichlet process with total mass parameter M and centering measure G_0 (Ferguson, 1983; Müller et al., 1996). We take $G_0 = N\text{-}\Gamma^{-1}(0, 1, \alpha_G, \beta_G)$, a normal inverse-gamma distribution with shape α_G and scale β_G. This approach allows us to flexibly model the source signals, without having to assume a fixed number of Gaussian components. This additional flexibility has been shown to improve estimation of the S_0 maps in simulations and in fMRI data analysis in our Chapter 3.

The covariate and center fixed effects are modeled under a Horseshoe+ prior (Carvalho et al., 2010; Bhadra et al., 2017). In particular, let $\mathbf{X}_{i,c}$ be a $J \times P$ matrix with rows corresponding to the individual and center covariates for subject i in center c at visit j, with $P = P_1 + P_2$. Similarly, let $\boldsymbol{\beta}_q(v) = [\beta_{1,q}^{(ind.)}(v), \ldots, \beta_{P_1,q}^{(ind.)}(v), \beta_{1,q}^{(center)}(v), \ldots, \beta_{P_2,q}^{(center)}(v)]'$ be the stacked individual and center covariate effects. Then under the Horseshoe+ prior we have,

$$\beta_{p,q}(v) \sim N(0, \lambda_{p,q}(v)^2\tau^2\gamma_q^2),$$

for $p = 1, \ldots, P$ and $q = 1, \ldots, Q$. Here, τ^2 is a global shrinkage parameter controlling the overall sparsity level among the covariate effects. $\lambda_{p,q}(v)^2$ is a local shrinkage term that is different across voxels, components, and effects that allows different levels of shrinkage in different spatial locations in the brain. The Horseshoe+ prior models the shrinkage terms under half-Cauchy priors as,

$$\begin{aligned} \lambda_{p,q}(v) &\sim \mathcal{C}^+(0, \eta_{p,q}(v)) \\ \tau &\sim \mathcal{C}^+(0, 1) \\ \eta_{p,q}(v) &\sim \mathcal{C}^+(0, 1), \end{aligned} \tag{4.5}$$

where $\mathcal{C}^+(0, 1)$ is a half-Cauchy distribution. This approach allows us to achieve the desired sparsity in estimating the individual and center level covariate effects.

4.2.3 Modeling the relationship between scans

A key innovation of our approach is how we model the dependency between different scanning sessions. There are several potential sources of within-group correlation in multi-center longitudinal studies.

First, we include a random intercept corresponding to data collection center to account for possible random variation in the brain network due simply to location. We place Normal priors on these terms,

$$L_{q,c}(v) \sim N(0, \sigma^2_{center}) \text{ for } c = 1, \ldots C$$

We place inverse gamma priors on the variance terms for the random intercepts. Specifically $\sigma^2_{center} \sim IG(\alpha_{center}, \beta_{center})$.

Second, to address the longitudinal nature of the data collection, we model within subject

correlation using a random subject-specific intercept, as well as random slopes for some research-defined subset of the covariates of interest. The random intercept accounts for differential variation in the source signals, while the random slopes allow us to account for potentially varying covariate effects across different subjects within the population. For example, in a sample of subjects with neurodegenerative diseases, we would potentially expect different rates of changes in the brain network in different subjects. The subject specific effects, $\mathbf{B}_{i,c,q}(v)$, are modeled under a multivariate normal distribution,

$$\mathbf{B}_{i,c,q}(v) \sim N(\mathbf{0}, \mathbf{R}_q(v)). \tag{4.6}$$

Note that under this setup we allow the relationship between the subject-specific random terms to be voxel and component specific. This is important, as the correlation structure between subject-effects is not likely to be the same across the entire brain and across all brain networks. For example, the correlation structure is likely different in active and inactive regions within a network. While this approach is quite flexible, it is not possible to estimate a separate covariance matrix at each voxel and network. Instead, we use a Bayesian non-parametric prior to cluster these covariance terms. In particular we use the DPM to flexibly model the subject-specific covariance terms,

$$\begin{aligned}
\mathbf{B}_{i,c,q}(v) &\sim N(\mathbf{0}, \mathbf{R}_q(v)) \text{ for } i = 1, \ldots N \\
\mathbf{R}_q(v) &\sim D \\
D &\sim DP(M_D, D_0),
\end{aligned}$$

where D_0 is an inverse Wishart distribution and M_D is the concentration parameter. By clustering the covariance terms under the DPM we are able to introduce flexibility in the dependency structure among the subject-level random effects without compromising the

ability to estimate the model parameters.

4.3 Posterior Computation

The posterior computation is simplified by collapsing the models in Equations (4.2) and (4.3) and premultiplying by the inverse of the orthogonal mixing matrix $\mathbf{A}_{i,c,j}$ to obtain a model for the unmixed preprocessed fMRI time courses:

$$\tilde{\mathbf{Y}}_{i,c,q}(v) = \mathbf{A}'_{i,c,j}\mathbf{Y}_{i,c,j}(v) = S_{0,q}(v)\mathbf{1}_J + L_{q,c}(v)\mathbf{1}_J + \mathbf{X}_{i,c}\boldsymbol{\beta}_q(v) + \mathbf{Z}_{i,c}\mathbf{B}_{i,c,q}(v) + \boldsymbol{\epsilon}_{i,c,q}(v),$$

where $\boldsymbol{\epsilon}_{i,c,j}(v) = \mathbf{E}_{i,c,j}(v) + \mathbf{A}'_{i,c,j}\mathbf{r}_{i,c,j}(v)$ is the combined error for the unmixed data. Here, $\boldsymbol{\epsilon}_{i,c,j}(v) \sim N_Q(\mathbf{0}, \text{diag}\{\gamma_1^2, \gamma_2^2, \ldots, \gamma_Q^2\})$ with $\gamma_q^2 = \sigma_y^2 + \sigma_q^2$. We model each combined error term using a non-informative prior, $\gamma_q^2 \sim \gamma_q^{-2} d\gamma_q^2$. We specify a uniform distribution over a sphere as the prior distribution for each participant's visit-specific mixing matrix, $\mathbf{A}_{i,c,j}$ (Hoff, 2009).

4.3.1 MCMC algorithm

The detailed updating steps for the posterior computation for the RM-SparseBayes ICA model are listed below.

4.3.1.1 Subject-Visit Level Mixing Matrix

The posterior distribution for the ith subject from center c's mixing matrix at visit j is given by,

$$\mathbf{A}_{i,c,j}|\mathbf{Y}_{i,c,j}, \mathbf{S}_{i,c,j}, \{\gamma_q^2\} \sim \text{BMF}\left(\mathbf{Y}_{i,c,j}\mathbf{Y}'_{i,c,j}, -\frac{1}{2}\text{diag}\{\gamma_1^{-2}, \gamma_2^{-2}, \ldots, \gamma_Q^{-2}\}, \mathbf{Y}_{i,c,j}\mathbf{S}'_{i,c,j}\right), \quad (4.7)$$

where BMF denotes a Matrix-Bingham-von-Mises-Fisher Distribution, which can be sampled from one pair of columns at a time as described in Hoff (2009). Detailed steps specific to

sampling a mixing matrix can be found in the posterior computation section of Chapter 3.

4.3.1.2 Center Random Effects

Let $\mathbf{Y}^*_{i,c,q}(v) = \tilde{\mathbf{Y}}_{i,c,q}(v) - S_{0,q}(v)\mathbf{1}_J - \mathbf{X}_{i,c}\boldsymbol{\beta}_q(v) - \mathbf{Z}_{i,c}\mathbf{B}_{i,c,q}(v)$. Then,

$$\mathbf{Y}^*_{i,c,q}(v)|L_{q,c}(v)\mathbf{1}_J, \gamma_q^2, \sim N(L_{q,c}(v)\mathbf{1}_J, \gamma_q^2 I_Q). \tag{4.8}$$

Let $\mathbf{Y}^*_q(v) = [\mathbf{Y}^*_{1,q}(v)', \ldots, \mathbf{Y}^*_{N,q}(v)']'$ be the stacked subject-level data for IC q and voxel v and let $\mathbf{W}_i$ be a $J \times C$ matrix with $w_{j,c} = 1$ if subject i belongs to center c and $w_{j,c} = 0$ otherwise. We can write the model across all subjects in matrix notation,

$$\mathbf{Y}^*_q(v)|\mathbf{W}, \mathbf{L}_q(v) \sim N\left(\mathbf{W}[L_{q,1}(v), L_{q,2}(v), \ldots, L_{q,C}(v)]', \boldsymbol{\Lambda}_q(v)\right), \tag{4.9}$$

where $\mathbf{W} = [[\mathbf{W}'_1, \ldots, \mathbf{W}'_N]']$ is a $NJ \times C$ matrix with rows corresponding to the center and visit membership of a subject and $\boldsymbol{\Lambda}_q(v) = \left(\gamma_q^2 I_{NJ}\right)$ captures the variation in $\mathbf{Y}^*_q(v)$.

The priors on the center-specific random intercepts are independent Normal with zero-means and variances of σ^2_{center}. The posterior for the center-specific random intercepts is also Normal, i.e.,

$$\begin{aligned}
[L_{q,1}(v), \ldots, L_{q,C}(v)]'|\mathbf{Y}^*_q(v), \mathbf{W}, \boldsymbol{\Lambda}_q(v), \sigma^2_{center} &\propto N\left(\mu_q^{(L)}(v), \Sigma_q^{(L)}(v)\right) \\
\Sigma_q^{(L)}(v) &= \left(\mathbf{W}'\boldsymbol{\Lambda}_q(v)\mathbf{W} + \sigma^2_{center} I\}\right)^{-1} \\
\mu_q^{(L)}(v) &= \Sigma_q^{(L)}(v)\mathbf{W}'\boldsymbol{\Lambda}_q(v)\mathbf{Y}^*_q(v).
\end{aligned}$$

Conditioned on the draws for the random intercepts, the corresponding posterior samples

for the variance term can be drawn an inverse gamma distribution:

$$\sigma^2_{center}|a_{center}, b_{center}, \mathbf{L} \propto IG\left(a_{center} + \frac{1}{2}QVC, b_{center} + \frac{1}{2}\sum_{q=1}^{Q}\sum_{v=1}^{V}\sum_{c=1}^{C} L_{q,c}(v)^2\right). \quad (4.10)$$

4.3.1.3 Subject-Specific Random Effects

Let $\mathbf{e}^*_{i,q}(v) = \tilde{\mathbf{Y}}_{i,c,q}(v) - L_{q,c}(v)\mathbf{1}_J - S_{0,q}(v)\mathbf{1}_J - \mathbf{X}_{i,c}\boldsymbol{\beta}_q(v)$ be the residual unmixed data after removing the random intercept for center, the fixed covariate effects, and the common set of spatial maps. Then we have the hierarchy,

$$\begin{aligned} \mathbf{e}^*_{i,q}(v)|\gamma^2_q, \mathbf{Z}_i, \mathbf{B}_{i,q}(v) &\sim N(\mathbf{Z}_i\mathbf{B}_{i,q}(v), \gamma^2_q I_J) \\ \mathbf{B}_{i,q}(v) &\sim N(\mathbf{0}, \mathbf{R}_q(v)), \end{aligned} \quad (4.11)$$

and thus,

$$\begin{aligned} \mathbf{B}_{i,q}(v)|\mathbf{e}^*_{i,q}(v), \gamma^2_q, \mathbf{Z}_i &\sim N\left(\boldsymbol{\mu}^{(B)}_{i,q}(v), \boldsymbol{\Sigma}^{(B)}_{i,q}(v)\right) \\ \boldsymbol{\Sigma}^{(B)}_{i,q}(v) &= \left(\frac{1}{\gamma^2_q}\mathbf{Z}'_i\mathbf{Z}_i + \mathbf{R}_q(v)^{-1}\right)^{-1} \\ \boldsymbol{\mu}^{(B)}_{i,q}(v) &= \boldsymbol{\Sigma}^{(B)}_{i,q}(v)\left(\frac{1}{\gamma^2_q}\mathbf{Z}'_i\mathbf{e}^*_{i,q}(v)\right). \end{aligned} \quad (4.12)$$

Conditioned on the subject-specific random effects it is straightforward to sample the random effects variance terms. This sampling is facilitated using the stick breaking representation of the Dirichlet process mixture (Sethuraman, 1994). In particular we can write D as,

$$D = \sum_{d=1}^{\infty} w^{(R)}_d \delta_{\mathbf{R}_d}, \quad (4.13)$$

where $w^{(R)}_d = \nu^{(R)}_d \prod_{d'=1}^{d-1}(1-\nu^{(R)}_{d'})$, $\nu^{(R)}_d \sim Beta(1, M)$, and $\mathbf{R}_d \sim D_0$. Under this representation, the density for the $\mathbf{B}_{i,c,q}(v)$ is written,

$$p(\mathbf{B}_{i,c,q}(v)|-) = \sum_{d=1}^{\infty} w_d^{(R)} g\left(\mathbf{B}_{i,c,q}(v); \mathbf{0}, \mathbf{R}_q(v)\right), \tag{4.14}$$

where $g(\cdot; \mathbf{0}, \mathbf{R}_q(v))$ denotes a multivariate normal density with mean $\mathbf{0}$ and variance $\mathbf{R}_q(v)$.

We update the DPM parameters using a slice sampling procedure (Walker, 2007). First, we introduce a latent variable for each element, $\ell_q(v) \sim Uniform(0, w_{H_q(v)})$. The joint density for the $\mathbf{B}_{i,c,q}(v)$ and $\ell_q(v)$ can be written:

$$p\left(\{\mathbf{B}_{i,c,q}(v)\}, \ell_q(v)|\boldsymbol{w}^{(R)}, \{\mathbf{R}_d\}\right) = \sum_{d=1}^{d^*} I_{\ell_q(v)<w_d^{(R)}} \prod_{i=1}^{N} g\left(\mathbf{B}_{i,c,q}(v); \mathbf{0}, \mathbf{R}_d\right), \tag{4.15}$$

where $d^* = \min\{d : \ell_q(v) > 1 - \sum_{d'=1}^{d} w_{d'}^{(R)} \text{ for all } q, v\}$. We then add the cluster membership variable for each voxel and IC, $H_q(v)$, where $p(H_q(v) = d|\ell_q(v), w_h^{(R)}) \propto I_{\ell_q(v)<w_d^{(R)}}$, yielding,

$$p\left(\{\mathbf{B}_{i,c,q}(v)\}, \ell_q(v), H_q(v)|\boldsymbol{w}^{(R)}, \{\mathbf{R}_d\}\right) = I_{\ell_q(v)<w_{H_q(v)}} \prod_{i=1}^{N} g(\mathbf{B}_{i,c,q}(v); \mathbf{0}, \mathbf{R}_{H_q(v)}), \tag{4.16}$$

Once we obtain the cluster memberships, $\{H_q(v)\}$, the typical posterior inference for Normal distributions can be applied by pooling information across voxels and ICs for all members of the same cluster.

The final step is to update the latent terms involved in the slice sampler. We sample the stick breaking weights from their conditional posterior $\nu_d^{(R)}|\mathbf{H}, M \sim Beta(1 + n_d, M_D + \sum_{d'=d+1}^{\infty} n_{d'})$, where n_d is the number of elements of $\mathbf{H}$ equal to d. We sample the $\ell_q(v)$ from their posterior: $\ell_q(v)|\mathbf{w}^{(R)}, H_q(v) \sim Uniform(0, w_{H_q(v)}^{(R)})$. Finally, the cluster memberships $\mathbf{H}$ can be sampled from their posterior distributions, which are given by:

$$p(H_q(v) = d|\boldsymbol{\ell}, \boldsymbol{w}^{(R)}, \mathbf{R}) = \frac{I_{\ell_q(v)<w_d^{(R)}} \prod_{i=1}^{N} g(\mathbf{B}_{i,c,q}(v); \mathbf{0}, \mathbf{R}_d)}{\sum_{d'=1}^{d^*} I_{\ell_q(v)<w_{d'}^{(R)}} \prod_{i=1}^{N} g(\mathbf{B}_{i,c,q}(v); \mathbf{0}, \mathbf{R}_{d'})}. \tag{4.17}$$

4.3.1.4 Population level maps, S_0, and Individual and Center Effects, $\boldsymbol{\beta}$

The population level maps, individual effects, and center effects can be updated as described in Chapter 3 after removing the random effects from the unmixed data, $\tilde{\mathbf{Y}}_{i,c,q}(v)$. Let $\mathbf{Y}^{(r)}_{i,c,q}(v) = \tilde{\mathbf{Y}}_{i,c,q}(v) - L_{q,c}(v)\mathbf{1}_J - \mathbf{Z}_{i,c}\mathbf{B}_{i,c,q}(v)$, and let $\mathbf{Y}^{(r)}_q(v)$ be the $NJ \times 1$ vector of the $\mathbf{Y}^{(r)}_{i,c,q}(v)$ stacked across all subjects and visits. Let $\mathbf{X}^* = [\mathbf{X_{1,1}}; \ldots; \mathbf{X}_{n_C,C}]$ be the corresponding $NJ \times P$ matrix of individual and center covariates. Then,

$$\boldsymbol{\beta}^*_q(v)|\mathbf{X}^*, \mathbf{Y}_q(v)^*, \gamma^2_q, \tau^{-2}, \boldsymbol{\lambda}, \boldsymbol{\sigma^2}, \boldsymbol{\mu}, \boldsymbol{\phi}^2, K_q(v) \sim N_{P+1}(\boldsymbol{\mu}_{\beta^*,q}(v), \boldsymbol{\Sigma}_{\beta^*,q}(v)), \tag{4.18}$$

where

$$\begin{aligned}
\boldsymbol{\Sigma}_{\beta^*,q}(v) &= \left(\mathbf{X}^{*'}\mathbf{X}^*\gamma_q^{-2} + \gamma_q^{-2}\text{diag}\{\sigma^{-2}_{K_q(v)}, \tau^{-2}\lambda_{1,q}(v)^{-2}, \ldots, \tau^{-2}\lambda_{P,q}(v)^{-2}\}\right)^{-1}, \\
\boldsymbol{\mu}_{\beta^*,q}(v) &= \boldsymbol{\Sigma}_{\beta^*,q}(v)\left(\mathbf{X}^{*'}\mathbf{Y}^*_q(v)\gamma_q^{-2} + \left[\frac{\mu_{K_q(v)}}{\phi^2_{K_q(v)}}, \mathbf{0}'_P\right]'\right).
\end{aligned} \tag{4.19}$$

4.3.2 Inference for Covariate Effects

Let $\mathbf{S}_0(v)^{(k)}$ be the kth sample of $\mathbf{S}_0$ from its posterior distribution and let $\beta_{p,q}(v)^{(k)}$ be the kth sample for the pth fixed effect at IC q and voxel v. Point estimates are given by $\hat{\mathbf{S}}_0(v) = \frac{1}{K}\sum_{k=1}^K \mathbf{S}_0(v)^{(k)}$ and $\hat{\beta}_{p,q}(v) = \frac{1}{K}\sum_{k=1}^K \beta_{p,q}(v)^{(k)}$ where K is the number of MCMC samples. We use Bayesian credible intervals to determine significance of the covariate effects. Let $I(\beta_{p,q}(v)^{(k)} > 0) = 1$ if the kth sample from the posterior is positive, 0 otherwise. Denote the credible interval-based p-value for the pth covariate at the qth IC and vth voxel by $CIP_{p,q}(v)$. Then,

$$CIP_{p,q}(v) = 2 \times \min\left\{\frac{1}{K}\sum_{k=1}^K I(\beta_{p,q}(v)^{(k)} > 0), 1 - \frac{1}{K}\sum_{k=1}^K I(\beta_{p,q}(v)^{(k)} > 0)\right\}. \tag{4.20}$$

Throughout the following simulations and data analysis, we determine the significant voxels to be those for which $CIP_{p,q}(v) < 0.05$.

4.4 Simulation

We next use a simulation study to investigate RM-SparseBayes ICA's performance under different data generation scenarios. To the best of our knowledge, there are no other approaches developed for this type of multi-center longitudinal data. Thus, we consider two reasonable alternative approaches to compare with our proposed RM-SparseBayes ICA. First, we consider a simple dual regression procedure that fits a linear model at each voxel, ignoring the random variation due to center and visit. We refer to this method as Naive DR. It is not expected to perform well, but provides some baseline idea of expected performance. Next, we consider using a data harmonization method known as ComBat (Johnson et al., 2007). Recently a version of ComBat, termed longComBat, was developed for longitudinal multi-center data (Beer et al., 2020). We implement a three stage approach using longComBat as follows. First, a TC-GICA approach is used with back reconstruction to obtain subject-specific independent component estimates. Then, in a second stage, these estimates are harmonized using the longComBat approach in order to remove variation due to center. Finally, a harmonized component data are used as the outcome in voxel-specific linear mixed effects models that include fixed individual and center effects, as well as random subject-specific intercepts and slopes. The estimates from this third stage are our final estimates under this approach, and the corresponding p-values from the mixed effects models can be used for hypothesis testing.

4.4.1 Data Generation

We created the spatial layout of the brain maps by first selecting three slices of a real brain ($V = 5080$). We created three square regions corresponding to the spatial activation

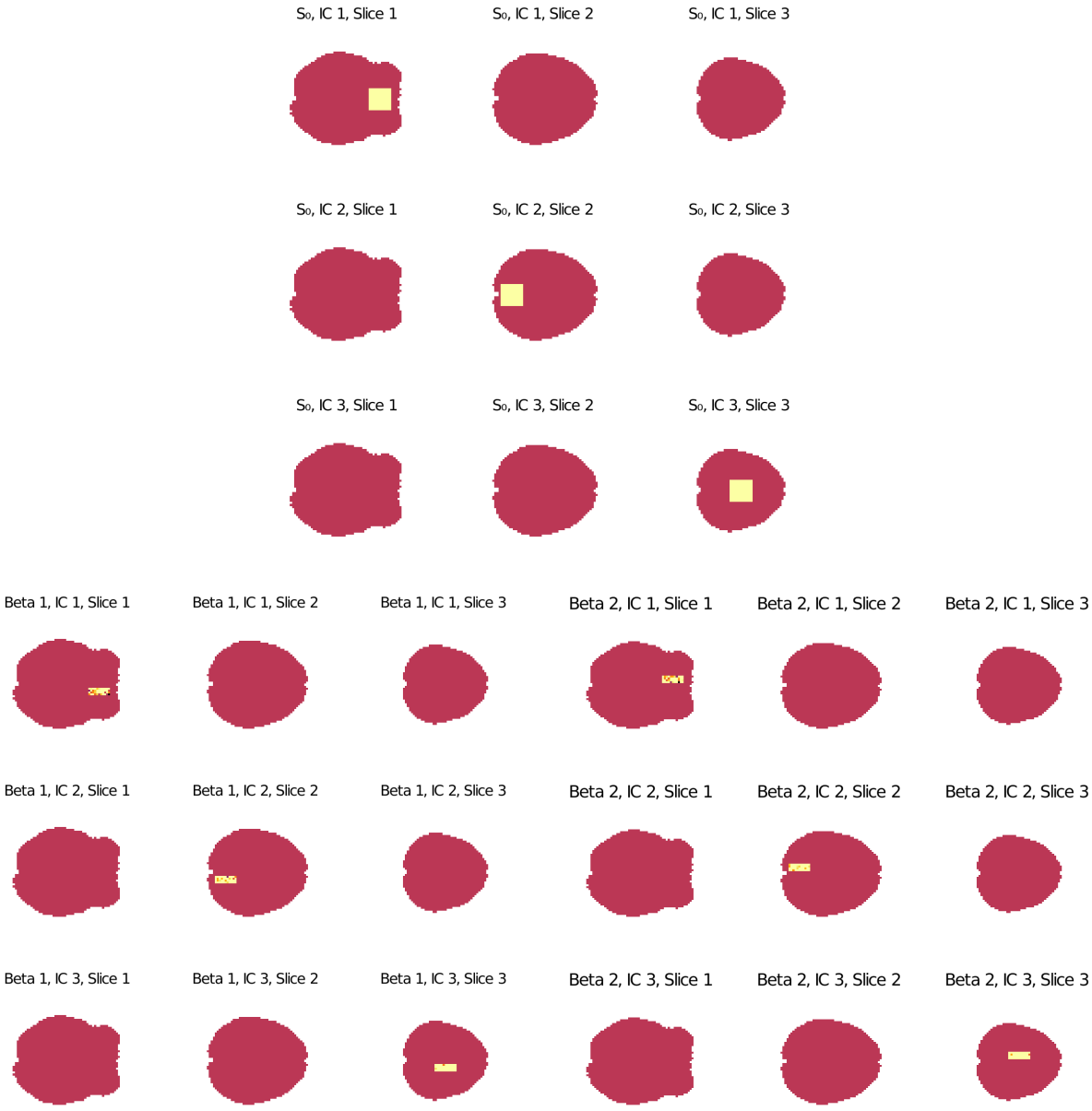


Figure 4.1: True spatial maps used for the simulation study.

of three brain networks (Figure 4.1, top).

Next, we generated covariate effect maps for two individual-level covariates – one continuous and one categorical. These effects were contained entirely within the activation map and were shaped as rectangular bars (Figure 4.1, bottom). Finally, we generated a center effect to represent scanner type. The activation region for this effect was circular. The intensities for the covariate effects were drawn from independent normal distributions with parameters specific to each covariate and IC. The covariate effects in all inactive regions were set to 0.

We considered data from 4 centers, and varied the sample size within each center at $N = 10, 20$, and 30. Each subject had three visits. Each subject's continuous covariate value at baseline was generated from a standard Normal distribution. The value for the continuous covariate at subsequent visits was incremented by a random uniform amount between 0 and 1. The categorical covariate value was drawn from a Bernoulii distribution. The subject-visit level spatial maps were created by adding the population level spatial maps and the covariate and center effects. Then, random variation was added due to visit and center. A subject-specific random intercept was added, as well as additional random variation in the slope for the continuous covariate (i.e. a random slope). Finally, independent noise was generated for each component from a normal distribution with mean 0 and variance σ_q^2, where $\sigma_q^2 \sim Gamma(b, 1)$. We considered 2 levels of variability: low ($b = 1.5$) and high ($b = 3.0$). The random intercept for each center was created by sampling a normal distribution with mean 0 and variance 0.1 at each voxel. For the subject-effect covariance, $\mathbf{R}_{qv}$, three different settings were used throughout the brain.

Finally, the subject and visit level time courses were generated to mimic the time courses of real fMRI studies by multiplying the subject-visit specific spatial maps with an orthonormal mixing matrix, adding $N(0, 0.1)$ noise, and then coloring the data to have 100 time points.

We assess the performance of the estimation procedures by examining the correlation between the estimated and true spatial maps for the source signals and fixed effects. We

also examine the ability to identify significant voxels via the area under the precision recall curve.

4.4.2 Results

Figure 4.2 displays the correlation between the true and estimated S_0 maps under each method. From the plots, our proposed approach shows clear advantages over the two-stage DR approaches, even for larger sample sizes. This indicates that our proposed approach has a better ability to estimate the underlying spatial components that are common across all subjects.

The advantages of our approach are even more pronounced when examining the correlation between the true and estimated individual covariate effects in Figure 4.3. Several randomly selected replicates are plotted in Figures 4.5 and 4.6, allowing us to see why the correlation under our approach is so much higher. From the plots, it is clear that our approach does a much better job of differentiating the active from the inactive regions. Figure 4.4 displays the corresponding area under the precision-recall curve results for identification of non-zero individual effects. From the plots, our approach has a clear advantage in finding significant regions. This advantage persists even as the sample size increases.

4.5 ADNI2 Study

4.5.1 Data and Preprocessing

We apply the RM-SparseBayes ICA approach to investigate brain networks in the Alzheimer's disease neuroimaging initiative 2 (ADNI2) data. This data contains longitudinal fMRI measurements on healthy controls, subjects with mild cognitive impairment (MCI) and Alzheimer's disease (AD). For our analysis, we used resting-state data from baseline screening, one year follow up, and two year follow up. All data were acquired from the ADNI website (`http://adni.loni.usc.edu`). Details on data acquisition are also available there.

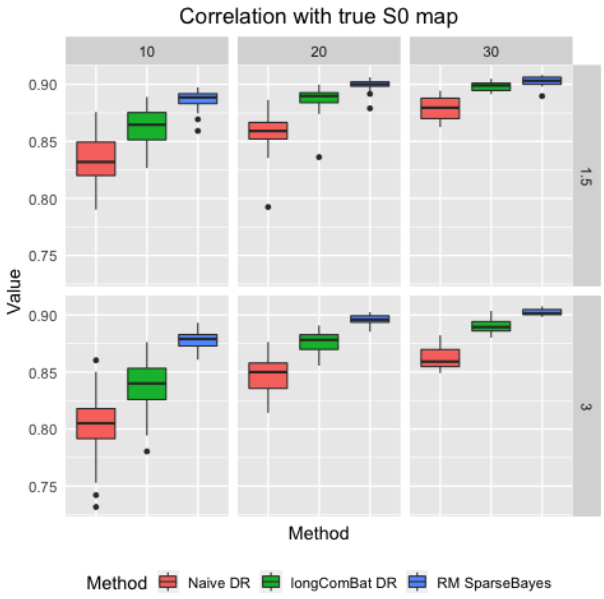


Figure 4.2: Correlation between the estimated and true S_0 maps for each method and each combination of simulation settings.

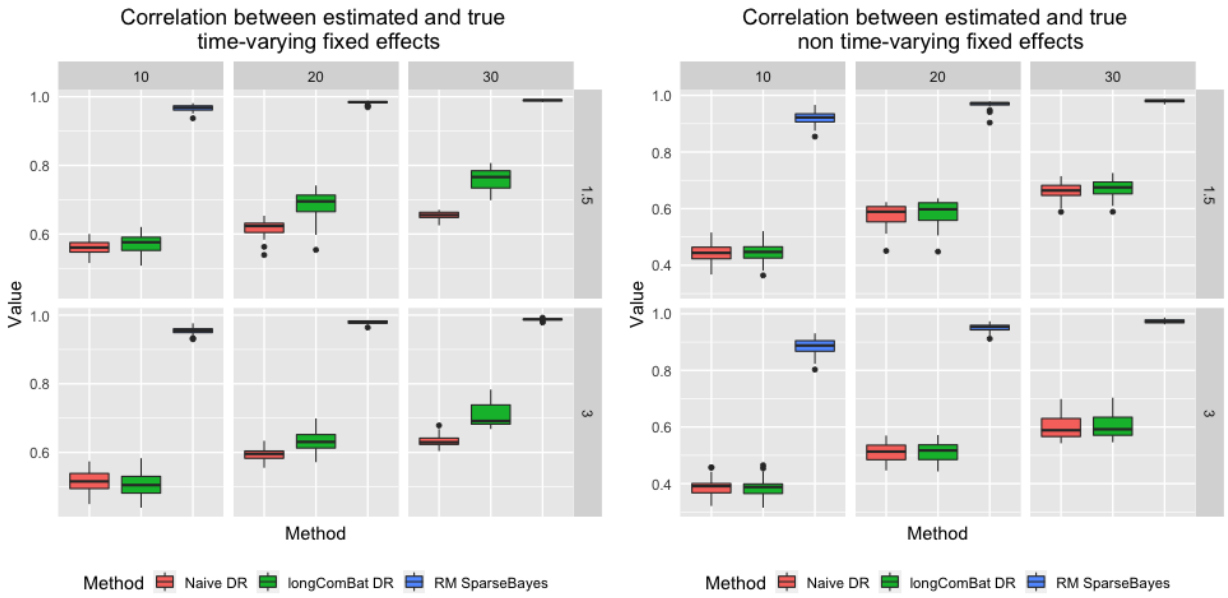


Figure 4.3: Correlation between the estimated and true S_0 maps for each method and each combination of simulation settings.

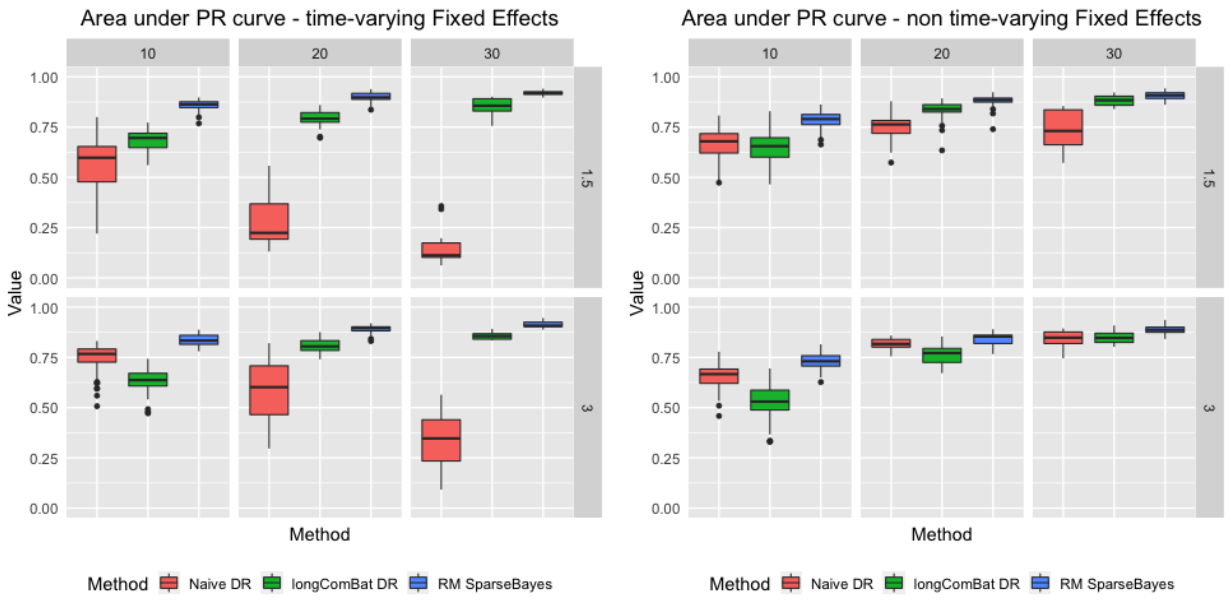


Figure 4.4: Area under the precision-recall curve for identifying non-zero covariate effects under each method.

We included 51 subjects in our analysis, each with 3 visits. At baseline, 16 subjects were classified as healthy controls, 29 as MCI patients, and 6 as AD patients. 24 (47%) of the subjects were female. The mean age at baseline was 73.6. The data were collected from 8 different centers, with 3 different types of scanner model (Achieva, Ingenia, and Intera). 33 scans were conducted on Intera scanners, 21 on Ingenia, and 99 on Achieva.

4.5.2 Modeling Procedure

We decomposed the ADNI2 resting-state data into 14 ICs. At the second level of the model, we included covariates for gender, age, time since baseline scan, MCI (yes/no), AD (yes/no), and time $\times$ diagnosis interactions. In the random component, we specific random intercepts for center and subject. We also included a random slope corresponding to the time since baseline to account for potentially varying subject-specific trajectories. We included two fixed center effects corresponding to indicator variables for Achieva and Ingenia scanners.

We fit the model using 3000 burn in iterations and 5000 MCMC iterations. We examined significance using credible intervals as described in Section 4.3.2.

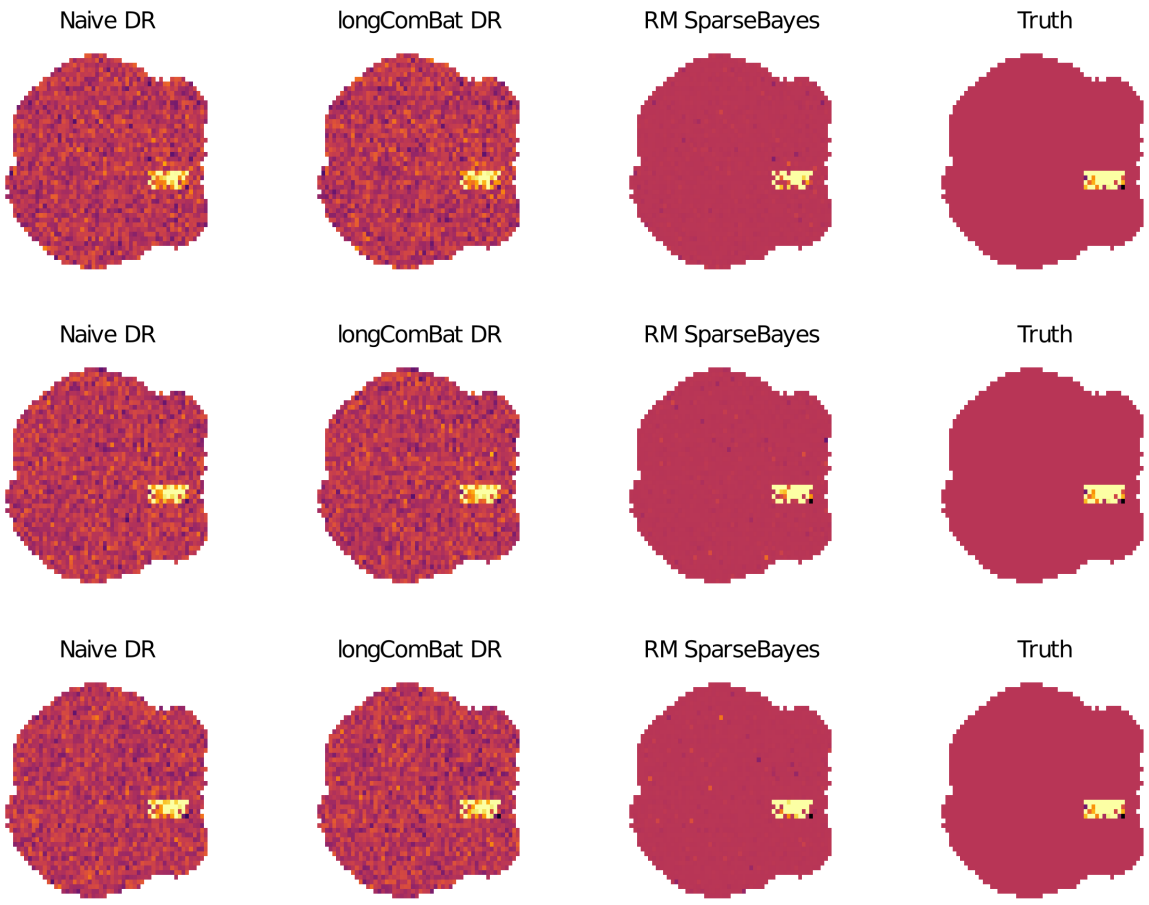


Figure 4.5: Estimates under each method for the individual time-varying effect from 3 randomly selected replicates. The true effect map for the individual time-varying effects is plotted on the right-hand side.

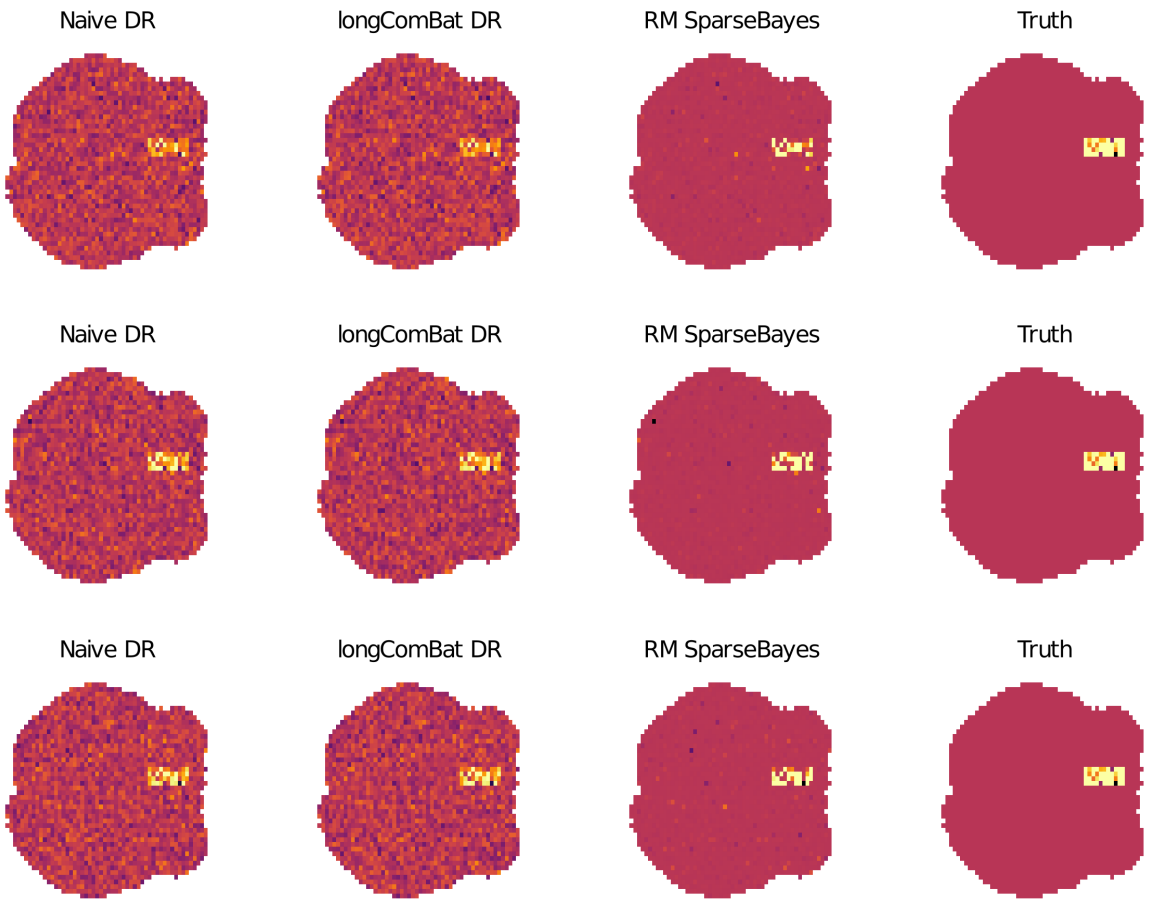


Figure 4.6: Estimates under each method for the individual non-time-varying effect from 3 randomly selected replicates. The true effect map for the individual non-time-varying effects is plotted on the right-hand side.

4.5.3 Results

The overall pattern of our results is similar to previous work (Wang and Guo, 2019). In particular, we see a large number of significant differences in the brain network between healthy controls and MCI and AD patients. Overall, there are more significant differences between healthy controls and AD patients than there are between healthy controls and MCI patients. In what follows, we highlight some of the interesting brain network results.

4.5.4 Sensorimotor Network

We found a cluster of significant voxels corresponding to AD within the sensorimotor network. This cluster of voxels was significant both in the AD main effect and the interaction between AD and time. The main effect for AD is negative, indicating less activation in the sensorimotor network. However, the interaction effect is positive, indicating that over time there is less of a disease effect in this region. There was no corresponding effect for the MCI group. Previous work has also identified AD related differences in sensorimotor network (Agosta et al., 2010; Salustri et al., 2013).

4.5.5 Default Mode Network

We found a significant negative MCI $\times$ time interaction in the left angular gyrus, indicating that this region in the DMN is less active in MCI patients. Interestingly, there was not a corresponding effect for AD patients, possibly due to the smaller AD sample size.

4.6 Discussion

In this work we introduced a general framework for repeated measures Sparse Bayesian ICA. This method can be applied to data collected from multiple centers, longitudinal data, or both. To the best of our knowledge, this is the first such ICA tool. Through simulation, we showed that our proposed approach has significantly better ability to identify individual

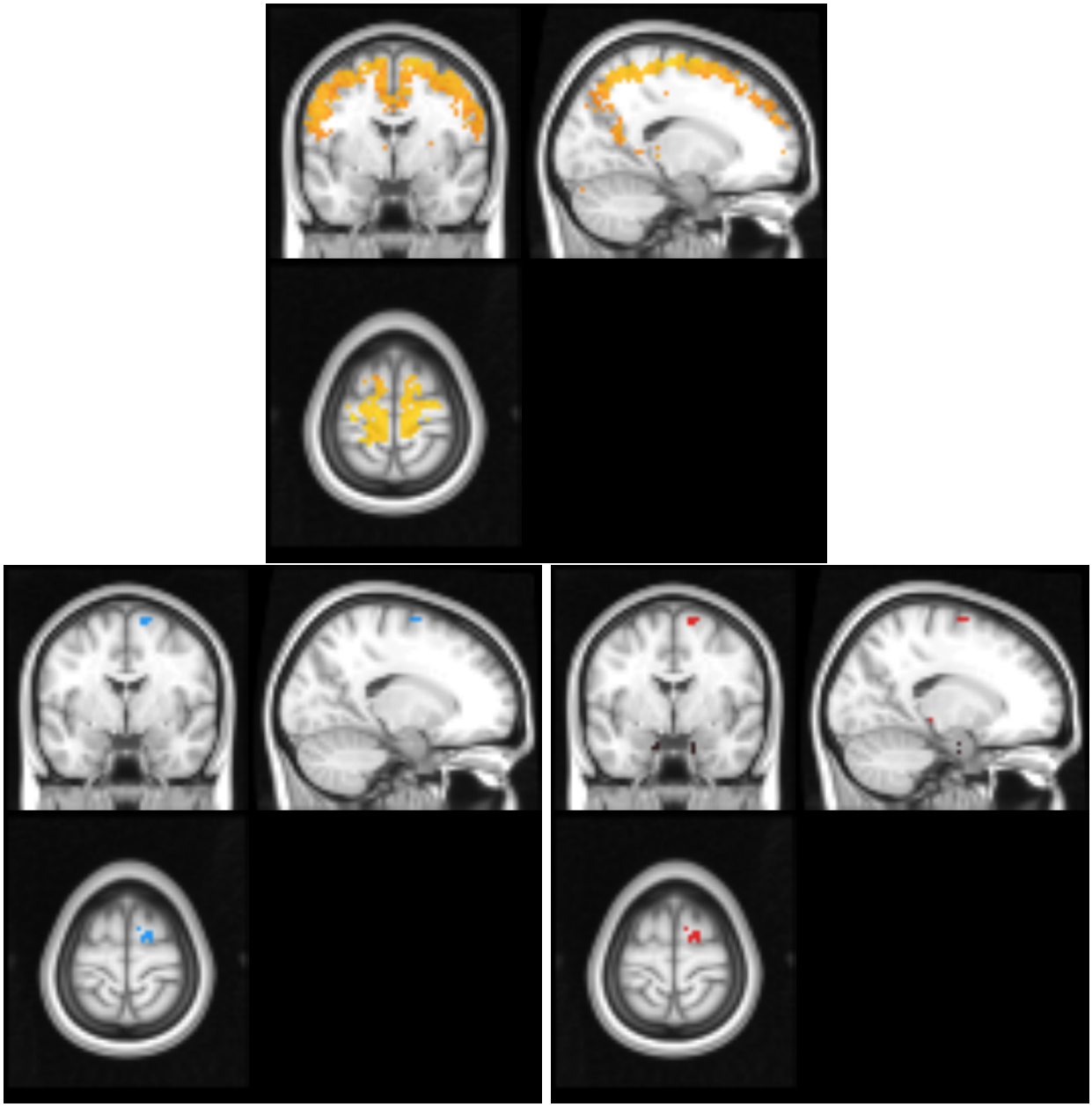


Figure 4.7: Estimated sensorimotor S_0 map using RM-SparseBayes ICA (top). Significant AD effects on the sensorimotor network. The left image shows the main effect for AD in blue (negative effect), and the right side displays the AD $\times$ time interaction in red (positive effect).

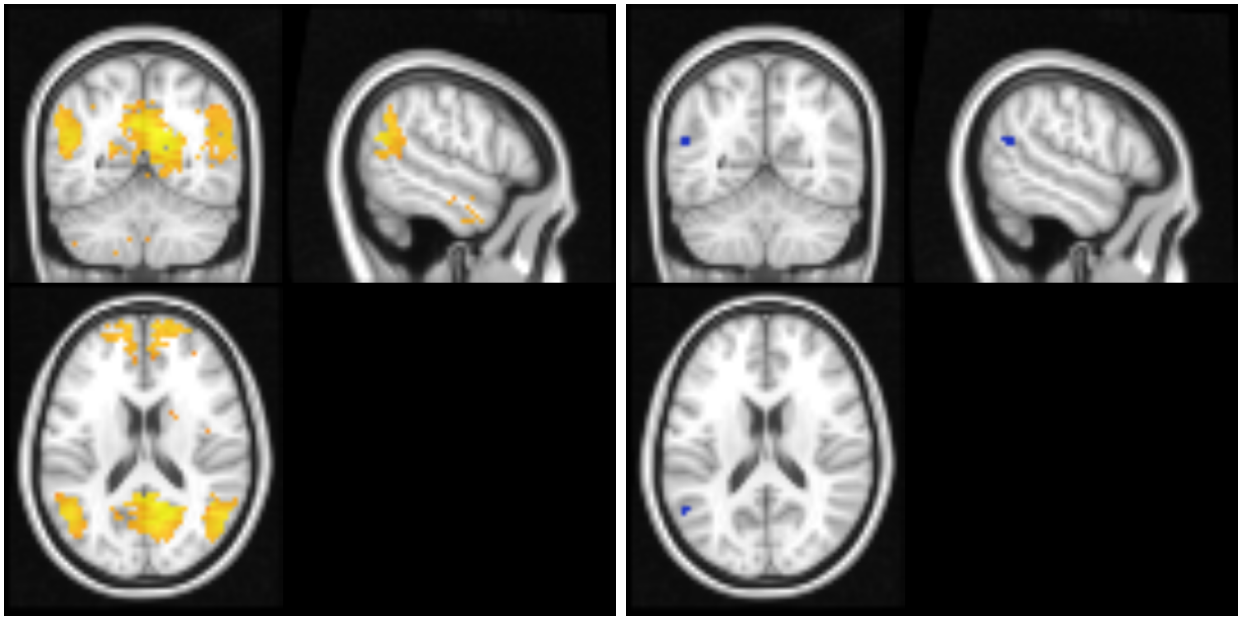


Figure 4.8: RM-SparseBayes ICA estimates for the DMN S_0 map (left). Significant negative effect of MCI $\times$ time interaction (right).

effects on the brain network. Application to a Alzheimer's disease data set found differences in the brain network due to MCI and AD status.

Chapter 5

Summary and Future Directions

In this dissertation, we have developed novel statistical methods for estimating and studying brain functional networks. Our approaches are relevant to both researchers interested in connectivity (Topic 1) and researchers interested in the overall spatial extent of brain networks (Topics 2 and 3).

In Topic 1, we introduced a novel Bayesian Joint Network Learning approach (BJNL) (Lukemire, Kundu, Pagnoni and Guo, 2020). This method examines functional connectivity using a graph theoretical approach, which allows the use of multi-scale analyses using graph metrics and connection strengths. The key innovation of this approach is that it pools information across different groups to model the edge probabilities, while estimating the edge strengths separately for each group conditioned on the edge probabilities. Through simulation, we showed that this approach can result in better graph estimation. In an application to a Stroop task fMRI study we found widespread, meaningful differences between executive control and frontoparietal networks. One potential future direction for the BJNL approach is an extension to longitudinal imaging data studies, which have become more common in recent years. Such an approach could model how brain connectivity changes over time in different groups.

In Topic 2, we developed a fully Bayesian approach to hierarchical covariate-adjusted independent component analysis, which we termed SparseBayes ICA. This approach offers several advantages over current state-of-the-art ICA techniques. In particular, it uses a flexible non-parametric approach to model the source signal distribution and incorporates shrinkage assumptions directly into the ICA decomposition. We show using simulations that this results in significantly improved power to detect true individual effects on the brain network. We then applied the SpareBayes ICA method to a Zen meditation study and showed that our proposed approach could find biologically meaningful differences in the brain network that competing methods failed to find.

In Topic 3, we introduced a general method of repeated measures SparseBayes ICA for the analysis of multi-center and longitudinal data. This uses random and fixed effects for

the data-collection centers to account for potential site effects. The RM-SparseBayes ICA procedure also uses random subject effects to incorporate longitudinal measurements. We conducted simulations and showed that our proposed approach offers significantly improved performance over other approaches. We applied the RM-SparseBayes ICA procedure to estimate the brain network in a longitudinal study of Alzheimer's disease and found regions in the sensorimotor and default mode networks exhibiting time-dependent changes in Alzhemier's and MCI patients, respectively. Future directions for both SparseBayes ICA and RM-SparseBayes ICA include developing computationally efficient forms applicable to large-scale imaging data, as well as incorporation into the Hierarchical Independent Component Analysis Toolbox (Lukemire, Wang, Verma and Guo, 2020).

Appendix A

Appendix for Chapter 2

A.1 fMRI Data Preprocessing

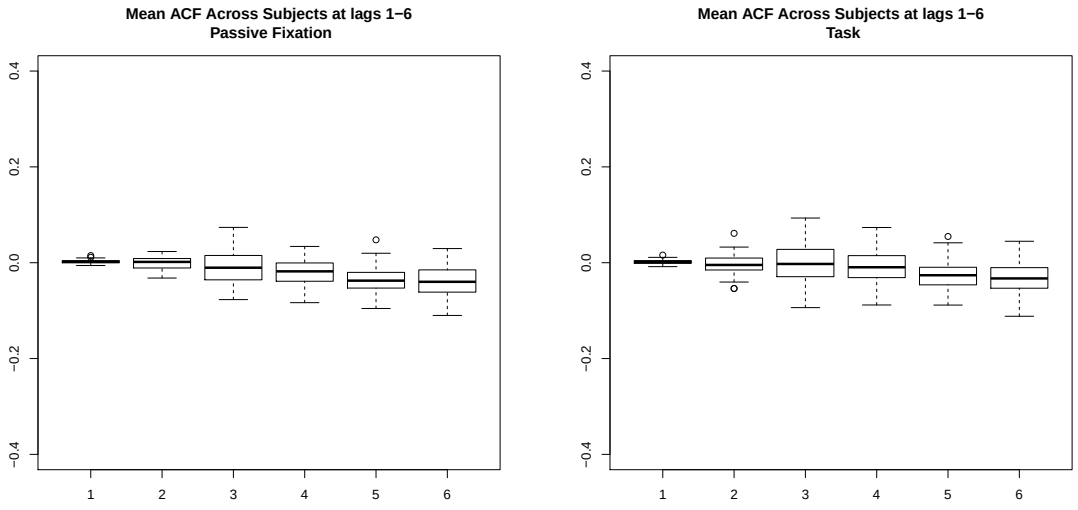
Image preprocessing was performed using the AFNI software package (Cox, 1996) and included slice-timing and motion correction, warping of brain volumes to standard Talairach space, Gaussian spatial blurring (6mm FWHM), and signal percent scaling.

In order to minimize the effect of spurious sources of temporal correlation in the fMRI time series, which could have negatively affected our network analysis, we performed the following additional pre-processing steps for each subject. First, a regression model was fit using a maximum likelihood approach with a baseline portion accounting for noise (a second-order Legendre polynomial, modeling slow signal drifts from biological or scanner-related confounds, plus 6 head-motion parameters) and a set of event-related regressors representing the expected BOLD response for the following classes of stimuli: (a) correctly-responded congruent trials, (b) correctly-responded incongruent trials, (c) commission error trials corresponding to incorrect responses, (d) omission error trials corresponding to omitted responses. Each event was modeled as a mini boxcar starting at the stimulus onset and ending at the time of the subject's response, followed by convolution with a gamma function accounting for the BOLD hemodynamic properties; if a response was not issued within the

allotted 2 second window, the trial was marked as an omission error and the duration of the corresponding boxcar was set at 2 seconds. The residual time-series from the regression analysis were collected, and task-related regressors, multiplied by their corresponding regression coefficients, were added back to them. This procedure aimed to remove the confounding components of the BOLD time series (slow scanner-related drifts, task-unrelated physiological variance, head motion) without also removing the component of interest, that is, task-related variance.

As a final preprocessing step prior to the analysis, we prewhitened the time series for each node. We fit an ARMA(1, 1) model at each node for each subject's data and used the residuals from these models as the corresponding node time series for the analysis. Boxplots of the autocorrelation of the resulting prewhitened data at lags 1−6 are provided in Figure A.1 for both the passive fixation and task conditions. The values at all lags are close to zero, which confirms that the temporal autocorrelation has been largely removed.

Figure A.1: Boxplots of the mean ACF across subjects for the passive fixation condition (left) and the task condition (right).



Appendix B

Appendix for Chapter 3

B.1 Additional Simulation Information

The factors for the simulation study are presented in Table B.1.

The intensities for the covariate effects were drawn from independent normal distributions $N(\mu_{\beta_p}, \sigma^2_{\beta_p})$ $(p = 1, 2)$ Here, $\mu_{\beta_1} = 1.5$, $\mu_{\beta_2} = 3.0$ and $\sigma^2_{\beta_p} = 1.5$ Covariate effects in all inactive regions were set to 0.

Table B.1: Factors in the simulation experiment. For each combination of factors, we repeated the experiment 50 times.

Factor	Levels	Description
Q	3, 6	Number of independent components
N	25, 50, 100	Number of subjects
σ^2_e	1.5, 3	Subject-level variance
b	1.5, 3.0, 5.0	between-subject variability in ICs

Bibliography

Agosta, F., Rocca, M. A., Pagani, E., Absinta, M., Magnani, G., Marcone, A., Falautano, M., Comi, G., Gorno-Tempini, M. L. and Filippi, M. (2010), 'Sensorimotor network rewiring in mild cognitive impairment and alzheimer's disease', Human brain mapping **31**(4), 515–525.

Alves, P. N., Foulon, C., Karolis, V., Bzdok, D., Margulies, D. S., Volle, E. and de Schotten, M. T. (2019), 'An improved neuroanatomical model of the default-mode network reconciles previous neuroimaging and neuropathological findings', Communications biology **2**(1), 1–14.

Amarai, S., Cichoki, A. and Chen, T. (1996), 'A new learning algorithm for blind source separation', Advances in Neural Information Processing Systems **8**, 757–763.

Anand, A., Li, Y., Wang, Y., Wu, J., Gao, S., Bukhari, L., Mathews, V. P., Kalnin, A. and Lowe, M. J. (2005), 'Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study', Biological psychiatry **57**(10), 1079–1088.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R. and Buckner, R. L. (2010), 'Functional-anatomic fractionation of the brain's default network', Neuron **65**(4), 550–562.

Antoniak, C. E. (1974), 'Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems', The Annals of Statistics pp. 1152–1174.

Bandettini, P. A. (2012), 'Twenty years of functional mri: the science and the stories', Neuroimage **62**(2), 575–588.

Barabási, A.-L. and Albert, R. (1999), 'Emergence of scaling in random networks', science **286**(5439), 509–512.

Bassett, D. S., Nelson, B. G., Mueller, B. A., Camchong, J. and Lim, K. O. (2012), 'Altered resting state complexity in schizophrenia', Neuroimage **59**(3), 2196–2207.

Beckmann, C. F., DeLuca, M., Devlin, J. T. and Smith, S. M. (2005), 'Investigations into resting-state connectivity using independent component analysis', Philosophical Transactions of the Royal Society B: Biological Sciences **360**(1457), 1001–1013.

Beckmann, C. F., Mackay, C. E., Filippini, N. and Smith, S. M. (2009), 'Group comparison of resting-state fmri data using multi-subject ica and dual regression', Neuroimage **47**(Suppl 1), S148.

Beckmann, C. F. and Smith, S. M. (2004), 'Probabilistic independent component analysis for functional magnetic resonance imaging', IEEE transactions on medical imaging **23**(2), 137–152.

Beckmann, C. F. and Smith, S. M. (2005), 'Tensorial extensions of independent component analysis for multisubject fmri analysis', Neuroimage **25**(1), 294–311.

Beer, J. C., Tustison, N. J., Cook, P. A., Davatzikos, C., Sheline, Y. I., Shinohara, R. T., Linn, K. A., Initiative, A. D. N. et al. (2020), 'Longitudinal combat: A method for harmonizing longitudinal multi-scanner imaging data', Neuroimage **220**, 117129.

Belliveau, J., Kennedy, D., McKinstry, R., Buchbinder, B., Weisskoff, R., Cohen, M., Vevea, J., Brady, T. and Rosen, B. (1991), 'Functional mapping of the human visual cortex by magnetic resonance imaging', Science **254**(5032), 716–719.

Benjamini, Y. and Yekutieli, D. (2001), 'The control of the false discovery rate in multiple testing under dependency', Annals of statistics pp. 1165–1188.

Bhadra, A., Datta, J., Polson, N. G., Willard, B. et al. (2017), 'The horseshoe+ estimator of ultra-sparse signals', Bayesian Analysis **12**(4), 1105–1131.

Biswal, B., Zerrin Yetkin, F., Haughton, V. M. and Hyde, J. S. (1995), 'Functional connectivity in the motor cortex of resting human brain using echo-planar MRI', Magnetic Resonance in Medicine **34**(4), 537–541.

Brewer, J. A., Worhunsky, P. D., Gray, J. R., Tang, Y.-Y., Weber, J. and Kober, H. (2011), 'Meditation experience is associated with differences in default mode network activity and connectivity', Proceedings of the National Academy of Sciences **108**(50), 20254–20259.

Bullmore, E. and Sporns, O. (2009), 'Complex brain networks: graph theoretical analysis of structural and functional systems', Nature reviews neuroscience **10**(3), 186.

Buxton, R. B. (2013), 'The physics of functional magnetic resonance imaging (fmri)', Reports on Progress in Physics **76**(9), 096601.

Calhoun, V. D., Adali, T., Pearlson, G. D. and Pekar, J. J. (2001), 'A method for making group inferences from functional mri data using independent component analysis', Human brain mapping **14**(3), 140–151.

Calhoun, V. D., Kiehl, K. A. and Pearlson, G. D. (2008), 'Modulation of temporally coherent brain networks estimated using ica at rest and during cognitive tasks', Human brain mapping **29**(7), 828–838.

Calhoun, V. D., Liu, J. and Adalı, T. (2009), 'A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data', Neuroimage **45**(1), S163–S172.

Carvalho, C. M., Polson, N. G. and Scott, J. G. (2010), 'The horseshoe estimator for sparse signals', Biometrika **97**(2), 465–480.

Chen, C.-H., Ridler, K., Suckling, J., Williams, S., Fu, C. H., Merlo-Pich, E. and Bullmore, E. (2007), 'Brain imaging correlates of depressive symptom severity and predictors of symptom improvement after antidepressant treatment', Biological psychiatry **62**(5), 407–414.

Cisler, J. M., Steele, J. S., Smitherman, S., Lenow, J. K. and Kilts, C. D. (2013), 'Neural processing correlates of assaultive violence exposure and ptsd symptoms during implicit threat processing: a network-level analysis among adolescent girls', Psychiatry Research: Neuroimaging **214**(3), 238–246.

cois Cardoso, J.-F. (1997), 'Infomax and maximum likelihood for blind source separation', IEEE Signal processing letters .

Cole, D. M., Smith, S. M. and Beckmann, C. F. (2010), 'Advances and pitfalls in the analysis and interpretation of resting-state fmri data', Frontiers in systems neuroscience **4**, 8.

Cole, L. J., Farrell, M. J., Gibson, S. J. and Egan, G. F. (2010), 'Age-related differences in pain sensitivity and regional brain activity evoked by noxious pressure', Neurobiology of aging **31**(3), 494–503.

Collett, D. (2002), Modelling binary data, Chapman and Hall/CRC.

Cox, R. W. (1996), 'AFNI: software for analysis and visualization of functional magnetic resonance neuroimages', Computers and Biomedical Research **29**(3), 162–173.

Danaher, P., Wang, P. and Witten, D. M. (2014), 'The joint graphical lasso for inverse covariance estimation across multiple classes', Journal of the Royal Statistical Society: Series B (Statistical Methodology) **76**(2), 373–397.

Daniels, J. K., Frewen, P., McKinnon, M. C. and Lanius, R. A. (2011), 'Default mode alterations in posttraumatic stress disorder related to early-life trauma: a developmental perspective', Journal of psychiatry & neuroscience: JPN **36**(1), 56.

Donnat, C., Tozzi, L. and Holmes, S. (2019), 'Constrained bayesian ica for brain connectome inference', arXiv preprint arXiv:1911.05770 .

Du, Y. and Fan, Y. (2013), 'Group information guided ica for fmri data analysis', Neuroimage **69**, 157–197.

Erdős, P. and Rényi, A. (1961), 'On the strength of connectedness of a random graph', Acta Mathematica Hungarica **12**(1-2), 261–267.

Erhardt, E. B., Rachakonda, S., Bedrick, E. J., Allen, E. A., Adali, T. and Calhoun, V. D. (2011), 'Comparison of multi-subject ica methods for analysis of fmri data', Human brain mapping **32**(12), 2075–2095.

Escobar, M. D. and West, M. (1995), 'Bayesian density estimation and inference using mixtures', Journal of the american statistical association **90**(430), 577–588.

Ferguson, T. S. (1973), 'A bayesian analysis of some nonparametric problems', The annals of statistics pp. 209–230.

Ferguson, T. S. (1983), Bayesian density estimation by mixtures of normal distributions, in 'Recent advances in statistics', Elsevier, pp. 287–302.

Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E. et al. (2017), 'Harmonization of multi-site diffusion tensor imaging data', Neuroimage **161**, 149–170.

Fox, P. T. and Raichle, M. E. (1986), 'Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects', Proceedings of the National Academy of Sciences **83**(4), 1140–1144.

Friedman, J., Hastie, T. and Tibshirani, R. (2008), 'Sparse inverse covariance estimation with the graphical lasso', Biostatistics **9**(3), 432–441.

Friston, K. J., Frith, C. D., Liddle, P. F. and Frackowiak, R. S. J. (1993), 'Functional connectivity: the principal-component analysis of large (PET) data sets', Journal of cerebral blood flow and metabolism **13**, 5.

Garrison, K. A., Zeffiro, T. A., Scheinost, D., Constable, R. T. and Brewer, J. A. (2015), 'Meditation leads to reduced default mode network activity beyond an active task', Cognitive, Affective, & Behavioral Neuroscience **15**(3), 712–720.

Genovese, C. R., Lazar, N. A. and Nichols, T. (2002), 'Thresholding of statistical maps in functional neuroimaging using the false discovery rate', Neuroimage **15**(4), 870–878.

George, E. I. and McCulloch, R. E. (1993), 'Variable selection via gibbs sampling', Journal of the American Statistical Association **88**(423), 881–889.

Gianaros, P. J., Derbtshire, S. W., May, J. C., Siegle, G. J., Gamalo, M. A. and Jennings, J. R. (2005), 'Anterior cingulate activity correlates with blood pressure during stress', Psychophysiology **42**(6), 627–635.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M. et al. (2016), 'A multi-modal parcellation of human cerebral cortex', Nature **536**(7615), 171.

Greicius, M. D., Flores, B. H., Menon, V., Glover, G. H., Solvason, H. B., Kenna, H., Reiss, A. L. and Schatzberg, A. F. (2007), 'Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus', Biological psychiatry **62**(5), 429–437.

Greicius, M. D., Srivastava, G., Reiss, A. L. and Menon, V. (2004), 'Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional mri', Proceedings of the National Academy of Sciences **101**(13), 4637–4642.

Guo, J., Levina, E., Michailidis, G. and Zhu, J. (2011), 'Joint estimation of multiple graphical models', Biometrika **98**(1), 1–15.

Guo, Y. (2011), 'A general probabilistic model for group independent component analysis and its estimation methods', Biometrics **67**(4), 1532–1542.

Guo, Y. and Pagnoni, G. (2008), 'A unified framework for group independent component analysis for multi-subject fmri data', NeuroImage **42**(3), 1078–1093.

Guo, Y. and Tang, L. (2013), 'A hierarchical model for probabilistic independent component analysis of multi-subject fMRI studies', Biometrics **69**(4), 970–981.

Hayasaka, S. and Laurienti, P. J. (2010), 'Comparison of characteristics between region-and voxel-based network analyses in resting-state fmri data', Neuroimage **50**(2), 499–508.

He, Y. and Evans, A. (2010), 'Graph theoretical modeling of brain connectivity', Current opinion in neurology **23**(4), 341–350.

Hennig, J., Speck, O., Koch, M. A. and Weiller, C. (2003), 'Functional magnetic resonance imaging: a review of methodological aspects and clinical applications', Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine **18**(1), 1–15.

Hoff, P. D. (2009), 'Simulation of the matrix Bingham–von Mises–Fisher distribution, with applications to multivariate and relational data', Journal of Computational and Graphical Statistics **18**(2), 438–456.

Hyvärinen, A. and Oja, E. (2000), 'Independent component analysis: algorithms and applications', Neural networks **13**(4-5), 411–430.

Janes, A. C., Pizzagalli, D. A., Richardt, S., Frederick, B. d., Chuzi, S., Pachas, G., Culhane, M. A., Holmes, A. J., Fava, M., Evins, A. E. et al. (2010), 'Brain reactivity to smoking

cues prior to smoking cessation predicts ability to maintain tobacco abstinence', Biological psychiatry **67**(8), 722–729.

Jernigan, T. L., Brown, S. A. and Dowling, G. J. (2018), 'The adolescent brain cognitive development study.', Journal of research on adolescence: the official journal of the Society for Research on Adolescence **28**(1), 154–156.

Johnson, W. E., Li, C. and Rabinovic, A. (2007), 'Adjusting batch effects in microarray expression data using empirical bayes methods', Biostatistics **8**(1), 118–127.

Jones, D. T., Machulda, M. M., Vemuri, P., McDade, E., Zeng, G., Senjem, M., Gunter, J., Przybelski, S., Avula, R., Knopman, D. S. et al. (2011), 'Age-related changes in the default mode network are more advanced in alzheimer disease', Neurology **77**(16), 1524–1531.

Kalli, M., Griffin, J. E. and Walker, S. G. (2011), 'Slice sampling mixture models', Statistics and computing **21**(1), 93–105.

Kemmer, P. B., Guo, Y., Wang, Y. and Pagnoni, G. (2015), 'Network-based characterization of brain functional connectivity in Zen practitioners', Frontiers in Psychology **6**, 603.

Khachouf, O. T., Chen, G., Duzzi, D., Porro, C. A. and Pagnoni, G. (2017), 'Voluntary modulation of mental effort investment: an fMRI study', Scientific Reports **7**(1), 17191.

Kundu, S., Lukemire, J., Wang, Y. and Guo, Y. (2019), 'A novel joint brain network analysis using longitudinal alzheimer's disease data', Scientific reports **9**(1), 1–18.

Li, Y., Craig, B. A. and Bhadra, A. (2019), 'The graphical horseshoe estimator for inverse covariance matrices', Journal of Computational and Graphical Statistics **28**(3), 747–757.

Liu, Y., Liang, M., Zhou, Y., He, Y., Hao, Y., Song, M., Yu, C., Liu, H., Liu, Z. and Jiang, T. (2008), 'Disrupted small-world networks in schizophrenia', Brain **131**(4), 945–961.

Lukemire, J., Kundu, S., Pagnoni, G. and Guo, Y. (2020), 'Bayesian joint modeling of multiple brain functional networks', Journal of the American Statistical Association pp. 1–13.

Lukemire, J., Wang, Y., Verma, A. and Guo, Y. (2020), 'Hint: A hierarchical independent component analysis toolbox for investigating brain functional networks using neuroimaging data', Journal of Neuroscience Methods p. 108726.

Makalic, E. and Schmidt, D. F. (2015), 'A simple sampler for the horseshoe estimator', IEEE Signal Processing Letters **23**(1), 179–182.

McKeown, M. J., Jung, T.-P., Makeig, S., Brown, G., Kindermann, S. S., Lee, T.-W. and Sejnowski, T. J. (1998), 'Spatially independent activity patterns in functional mri data during the stroop color-naming task', Proceedings of the National Academy of Sciences **95**(3), 803–810.

McKeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., Bell, A. J. and Sejnowski, T. J. (1998), 'Analysis of fmri data by blind separation into independent spatial components', Human brain mapping **6**(3), 160–188.

McLachlan, G. and Peel, D. (2004), Finite mixture models, John Wiley & Sons.

Mejia, A. F., Nebel, M. B., Wang, Y., Caffo, B. S. and Guo, Y. (2020), 'Template independent component analysis: Targeted and reliable estimation of subject-level brain networks using big data population priors', Journal of the American Statistical Association **115**(531), 1151–1177.

Miller, D. R., Hayes, S. M., Hayes, J. P., Spielberg, J. M., Lafleche, G. and Verfaellie, M. (2017), 'Default mode network subsystems are differentially disrupted in posttraumatic stress disorder', Biological Psychiatry: Cognitive Neuroscience and Neuroimaging **2**(4), 363–371.

Minka, T. P. (2001), Automatic choice of dimensionality for pca, in 'Advances in neural information processing systems', pp. 598–604.

Mogensen, P. K. and Riseth, A. N. (2018), 'Optim: A mathematical optimization package for Julia', Journal of Open Source Software **3**(24), 615.

Mueller, S. G., Weiner, M. W., Thal, L. J., Petersen, R. C., Jack, C. R., Jagust, W., Trojanowski, J. Q., Toga, A. W. and Beckett, L. (2005), 'Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's disease neuroimaging initiative (ADNI)', Alzheimer's & Dementia **1**(1), 55–66.

Müller, P., Erkanli, A. and West, M. (1996), 'Bayesian curve fitting using multivariate normal mixtures', Biometrika **83**(1), 67–79.

O'Brien, S. M. and Dunson, D. B. (2004), 'Bayesian multivariate logistic regression', Biometrics **60**(3), 739–746.

Pagnoni, G. (2012), 'Dynamical properties of bold activity from the ventral posteromedial cortex associated with meditation and attentional skills', Journal of Neuroscience **32**(15), 5242–5249.

Peterson, C., Stingo, F. C. and Vannucci, M. (2015), 'Bayesian inference of multiple gaussian graphical models', Journal of the American Statistical Association **110**(509), 159–174.

Piironen, J., Vehtari, A. et al. (2017), 'Sparsity information and regularization in the horseshoe and other shrinkage priors', Electronic Journal of Statistics **11**(2), 5018–5051.

Polson, N. G. and Scott, J. G. (2010), 'Shrink globally, act locally: Sparse bayesian regularization and prediction', Bayesian statistics **9**, 501–538.

Polson, N. G., Scott, J. G. and Windle, J. (2013), 'Bayesian inference for logistic models using Pólya–Gamma latent variables', Journal of the American Statistical Association **108**(504), 1339–1349.

Pooley, R. A. (2005), 'Fundamental physics of mr imaging', Radiographics **25**(4), 1087–1099.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L. et al. (2011), 'Functional network organization of the human brain', Neuron **72**(4), 665–678.

Quiton, R. L. and Greenspan, J. D. (2007), 'Sex differences in endogenous pain modulation by distracting and painful conditioning stimulation', Pain **132**, S134–S149.

Reineberg, A. E., Andrews-Hanna, J. R., Depue, B. E., Friedman, N. P. and Banich, M. T. (2015), 'Resting-state networks predict individual differences in common and specific aspects of executive function', Neuroimage **104**, 69–78.

Rogers, B. P., Morgan, V. L., Newton, A. T. and Gore, J. C. (2007), 'Assessing functional connectivity in the human brain by fmri', Magnetic resonance imaging **25**(10), 1347–1357.

Rosa, M. J., Portugal, L., Hahn, T., Fallgatter, A. J., Garrido, M. I., Shawe-Taylor, J. and Mourao-Miranda, J. (2015), 'Sparse network-based models for patient classification using fmri', Neuroimage **105**, 493–506.

Rubinov, M. and Sporns, O. (2010), 'Complex network measures of brain connectivity: uses and interpretations', Neuroimage **52**(3), 1059–1069.

Rytty, R., Nikkinen, J., Paavola, L., Abou Elseoud, A., Moilanen, V., Visuri, A., Tervonen, O., Renton, A., Traynor, B., Kiviniemi, V. J. et al. (2013), 'Groupica dual regression analysis of resting state networks in a behavioral variant of frontotemporal dementia', Frontiers in human neuroscience **7**, 461.

Sahakian, B. J. and Owen, A. (1992), 'Computerized assessment in neuropsychiatry using cantab: discussion paper.', Journal of the Royal Society of Medicine **85**(7), 399.

Salustri, C., Tecchio, F., Zappasodi, F., Tomasevic, L., Ercolani, M., Moffa, F., Cassetta, E., Rossini, P. M. and Squitti, R. (2013), 'Sensorimotor cortex reorganization in alzheimer's

disease and metal dysfunction: a meg study', International Journal of Alzheimer's Disease **2013**.

Schmittmann, V. D., Jahfari, S., Borsboom, D., Savi, A. O. and Waldorp, L. J. (2015), 'Making large-scale networks from fmri data', PloS one **10**(9), e0129074.

Sethuraman, J. (1994), 'A constructive definition of dirichlet priors', Statistica sinica pp. 639–650.

Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., Mintun, M. A., Wang, S., Coalson, R. S. and Raichle, M. E. (2009), 'The default mode network and self-referential processes in depression', Proceedings of the National Academy of Sciences **106**(6), 1942–1947.

Shi, R. and Guo, Y. (2016), 'Investigating differences in brain functional networks using hierarchical covariate-adjusted independent component analysis', The annals of applied statistics **10**(4), 1930.

Smith, D. V., Utevsky, A. V., Bland, A. R., Clement, N., Clithero, J. A., Harsch, A. E., Carter, R. M. and Huettel, S. A. (2014), 'Characterizing individual differences in functional connectivity using dual-regression and seed-based approaches', Neuroimage **95**, 1–12.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R. et al. (2009), 'Correspondence of the brain's functional architecture during activation and rest', Proceedings of the National Academy of Sciences **106**(31), 13040–13045.

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D. and Woolrich, M. W. (2011), 'Network modelling methods for fmri', Neuroimage **54**(2), 875–891.

Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., Nichols, T. E., Robinson, E. C., Salimi-Khorshidi, G., Woolrich, M. W. et al. (2013), 'Functional connectomics from resting-state fmri', Trends in cognitive sciences **17**(12), 666–682.

Stam, C. J. (2014), 'Modern network science of neurological disorders', Nature Reviews Neuroscience **15**(10), 683.

Stroop, J. R. (1935), 'Studies of interference in serial verbal reactions.', Journal of Experimental Psychology **18**(6), 643.

Supekar, K., Menon, V., Rubin, D., Musen, M. and Greicius, M. D. (2008), 'Network analysis of intrinsic functional brain connectivity in alzheimer's disease', PLoS computational biology **4**(6), e1000100.

Taren, A. A., Gianaros, P. J., Greco, C. M., Lindsay, E. K., Fairgrieve, A., Brown, K. W., Rosen, R. K., Ferris, J. L., Julson, E., Marsland, A. L. et al. (2017), 'Mindfulness meditation training and executive control network resting state functional connectivity: a randomized controlled trial', Psychosomatic medicine **79**(6), 674.

Taylor, V. A., Daneault, V., Grant, J., Scavone, G., Breton, E., Roffe-Vidal, S., Courtemanche, J., Lavarenne, A. S., Marrelec, G., Benali, H. et al. (2013), 'Impact of meditation training on the default mode network during a restful state', Social cognitive and affective neuroscience **8**(1), 4–14.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B. and Joliot, M. (2002), 'Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain', Neuroimage **15**(1), 273–289.

Van Den Heuvel, M. P. and Pol, H. E. H. (2010), 'Exploring the brain network: a review on resting-state fmri functional connectivity', European neuropsychopharmacology **20**(8), 519–534.

van der Pas, S., Szabó, B., van der Vaart, A. et al. (2017), 'Uncertainty quantification for the horseshoe (with discussion)', Bayesian Analysis **12**(4), 1221–1274.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., Consortium, W.-M. H. et al. (2013), 'The wu-minn human connectome project: an overview', Neuroimage **80**, 62–79.

Walker, S. G. (2007), 'Sampling the Dirichlet mixture model with slices', Communications in Statistics−Simulation and Computation **36**(1), 45–54.

Wang, H. (2012), 'Bayesian graphical lasso models and efficient posterior computation', Bayesian Analysis **7**(4), 867–886.

Wang, Y. and Guo, Y. (2019), 'A hierarchical independent component analysis model for longitudinal neuroimaging studies', NeuroImage **189**, 380–400.

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

Weishaupt, D., Köchli, V. D. and Marincek, B. (2008), How does MRI work?: an introduction to the physics and function of magnetic resonance imaging, Springer Science & Business Media.

Wilcox, C. E., Calhoun, V. D., Rachakonda, S., Claus, E. D., Littlewood, R. A., Mickey, J., Arenella, P. B. and Hutchison, K. E. (2017), 'Functional network connectivity predicts treatment outcome during treatment of nicotine use disorder', Psychiatry Research: Neuroimaging **265**, 45–53.

Yu, H. and Dauwels, J. (2016), Variational bayes learning of time-varying graphical models, in '2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)', IEEE, pp. 1–6.

Yu, M., Linn, K. A., Cook, P. A., Phillips, M. L., McInnis, M., Fava, M., Trivedi, M. H., Weissman, M. M., Shinohara, R. T. and Sheline, Y. I. (2018), 'Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fmri data', Human brain mapping **39**(11), 4213–4227.

Zhong, Y., Huang, L., Cai, S., Zhang, Y., von Deneen, K. M., Ren, A., Ren, J., Initiative, A. D. N. et al. (2014), 'Altered effective connectivity patterns of the default mode network in alzheimer's disease: an fmri study', Neuroscience letters **578**, 171–175.