Generative Models of Brain Connectivity for Population Studies

by

Archana Venkataraman

S.B., Electrical Engineering, Massachusetts Institute of Technology (2006)
M.Eng., Electrical Engineering, Massachusetts Institute of Technology (2007)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology

September 2012

Signature of Author: .....
Department of Electrical Engineering and Computer Science

Certified by: .....
Polina Golland
Associate Professor of Electrical Engineering and Computer Science
Distinguished Alumnus (1964) Career Development Chair

Accepted by: .....
Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chairman, Department Committee for Graduate Students

Generative Models of Brain Connectivity for Population Studies

by

Archana Venkataraman

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

Connectivity analysis focuses on the interaction between brain regions. Such relationships inform us about patterns of neural communication and may enhance our understanding of neurological disorders. This thesis proposes a generative framework that uses anatomical and functional connectivity information to find impairments within a clinical population. Anatomical connectivity is measured via Diffusion Weighted Imaging (DWI), and functional connectivity is assessed using resting-state functional Magnetic Resonance Imaging (fMRI).

We first develop a probabilistic model to merge information from DWI tractography and resting-state fMRI correlations. Our formulation captures the interaction between hidden templates of anatomical and functional connectivity within the brain. We also present an intuitive extension to population studies and demonstrate that our model learns predictive differences between a control and a schizophrenia population. Furthermore, combining the two modalities yields better results than considering each one in isolation.

Although our joint model identifies widespread connectivity patterns influenced by a neurological disorder, the results are difficult to interpret and integrate with our region-centric knowledge of the brain. To alleviate this problem, we present a novel approach to identify regions associated with the disorder based on connectivity information. Specifically, we assume that impairments of the disorder localize to a small subset of brain regions, which we call *disease foci*, and affect neural communication to/from these regions. This allows us to aggregate pairwise connectivity changes into a region-based representation of the disease. Once again, we use a probabilistic formulation: latent variables specify a template organization of the brain, which we indirectly observe through resting-state fMRI correlations and DWI tractography. Our inference algorithm simultaneously identifies both the afflicted regions and the network of aberrant functional connectivity.

Finally, we extend the region-based model to include multiple collections of foci, which we call *disease clusters*. Preliminary results suggest that as the number of clusters increases, the refined model explains progressively more of the functional differences between the populations.

Thesis Supervisor: Polina Golland

Title: Associate Professor of Electrical Engineering and Computer Science

Acknowledgments

Any great endeavor relies on the guidance and support of the people around us. I have been incredibly fortunate to have Polina Golland as my thesis advisor. Polina introduced me to the field of medical image analysis, and she has helped broaden my network of associations within the community. More importantly, she has pushed me to define my own research directions and to explore challenging problems. At the same time, her insights have been crucial to this work, and her unbounded enthusiasm and optimism have kept me motivated through many research dry spells.

I would also like to thank my thesis committee members, Carl-Fredrik Westin, Bertrand Thirion and Devavrat Shah for their constructive feedback. C-F has been a long-time collaborator to this work, Bertrand has provided many useful suggestions, and my conversations with Devavrat helped inspire the network models in the second half of this thesis. Among my other mentors, I am especially grateful to Anantha Chandrakasan. He supported my professional goals when I was a naïve Freshman and has provided many opportunities during my undergraduate and graduate programs at MIT.

This thesis would not be possible without the intellectual contributions of Marek Kubicki, Yogesh Rathi, Danial Lashkari and Georg Langs. Marek provided the dataset used in this work along with clinical interpretations of the schizophrenia results. Similarly, Yogesh has been invaluable in helping me to process and understand the diffusion data. I found myself drawn to probabilistic graphical models after talking with Danial about his research. I became hooked when I, too, started running out of Greek and Roman symbols. Finally, Georg's visualization tool has been essential for displaying the connectivity results.

I am indebted to all the members of the MIT Medical Vision Group: Mert Sabuncu, Tammy Riklin-Raviv, Bjoern Menze, Georg Langs, Thomas Yeo, Wanmei Ou, Danial Lashkari, Ramesh Sridharan, Michal Depa, Adrian Dalca, George Chen, Andrew Sweet, Amelia Arbisser and Christian Wachinger. These individuals have donated countless hours to practice talks and proof-reading papers. I am especially grateful for the stimulating discussions during our group meetings, punctuated by lively banter about current events and, of course, the latest football stats.

Jason Chang and Randi Cabezas have made the office a fun place to spend (too) many hours every day. As evidenced by the periodic Nerf gun battles, they contribute to a friendly and open work environment that I'd be hard-pressed to find anywhere else.

I am also grateful to Ann and Terry Orlando, the housemasters of Ashdown House, for helping me transition into graduate life. They are wonderful and caring people, who are committed to maintaining a fun and vibrant community within Ashdown.

Austin Che has been my shoulder to lean on through the turmoil of life and graduate school. More than anyone else, he has supported me throughout my long tenure at MIT. I am thankful for his love and devotion.

I would not be where I am today without the love and support of my family. My parents have nurtured my natural curiosity since childhood. My mother encouraged me towards math and science from a very early age and shaped my lifelong goal to attend MIT. A professor herself, she is someone I have always looked up to and admired. My father was the fun-loving parent. But at the same time, he taught me to open my eyes and observe the world around me. My brother Vinayak was my early companion and guinea pig. It amazes me to see the confident young scientist he has become.

Finally, I would like to thank Duke. Thank you for melting away 25 years of unhealthy seriousness and for showing me that there really is a world outside the mismatched buildings of MIT. Wherever life takes us, you will always be my best friend.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The brain is a complex organ capable of storing and processing information from a myriad of sensory inputs. Anatomically, it is an intricate network in which cortical and subcortical processing centers are connected by neural axons. Functionally, the brain is partitioned into specialized regions that interact to perform complex tasks. Non-invasive imaging technologies allow us to probe this complex construction. Traditional analysis has focused on localized properties of the brain, which include segmenting a specific structure [6], quantifying changes in volume or tissue properties [72], and pinpointing functional activation [35].

Recently, the focus has shifted to brain connectivity, which measures the relationship *between* regions rather than characteristics of an individual locale. These patterns of interaction provide further insight into the organization of the brain and may deepen our understanding of neuropsychiatric disorders. This thesis is concerned with anatomical and functional connectivity between brain regions. Anatomical connectivity informs us about the neural pathways in the brain; it represents the brain's internal wiring. Diffusion Weighted Imaging (DWI) is often used to measure anatomical connectivity. DWI captures the anisotropic diffusion of water within the brain. We estimate the underlying white matter fiber bundles using tractography [4, 5]. Functional connectivity assesses neural synchrony, which relates to patterns of communication within the brain. Functional connectivity is typically measured via temporal correlations in resting-state functional Magnetic Resonance Imaging (fMRI) data [14, 32].

We explore two major ideas in this thesis: (1) multi-modal analysis of resting-state fMRI and DWI data, and (2) identifying the effects of a neurological disorder based on multi-modal information. Unlike conventional analysis, we formulate a unified, generative framework of anatomical and functional connectivity. In this way, we *model* the interactions in the brain. The goal is to enhance our understanding of the brain beyond collecting statistics. To date, little progress has been made in this area. At the time of initial publication, ours were the first stochastic models to combine anatomical and functional connectivity and to infer patterns of abnormal connectivity induced by a disease. [95, 97].

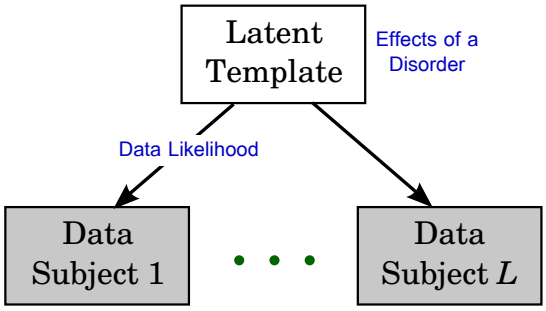


Figure 1-1: Motivation for our generative models of brain connectivity.

1.1 Departure from Traditional Analysis

Traditionally, connectivity analysis has focused on extracting and analyzing statistics of the data. Variations in these statistics across individuals or between different populations informs us about the brain. For example, functional systems are identified from resting-state fMRI data as voxels whose dynamics strongly correlate with that of a given seed region [8]. Similarly, multi-modal analysis searches for correlations between independently computed fMRI and DWI measures [47, 48, 61, 85]. The information is later pooled into a joint representation. Univariate tests and random effects analysis are commonly used in population studies of connectivity [40,60,66,105]. This approach relies on a statistical score, computed independently for each connection and modality, to determine connections that differ between a clinical population and normal controls. Although multi-pattern analyses of connectivity have been proposed [52, 53, 96], the majority of these methods identify but cannot explain the resulting connectivity structure.

Our approach, as illustrated in Fig. 1-1, is to treat anatomical connectivity and functional dynamics as components of an underlying generative process, which we observe via resting-state fMRI data and DWI tractography. Specifically, we formulate a set of *latent* variables to represent a (hidden) template for a given population. This is complemented with a simple but effective data likelihood model; observed data in individual subjects are generated stochastically based on the template. We further hypothesize that the effects of a disorder can be explained via changes to the latent structure.

A departure from conventional philosophy, this hierarchical organization affords several desirable properties. First, it is a natural representation for our real-world observations. Namely, fMRI and DWI measure correlates of neural activity and axonal pathways, respectively. They do not inform us directly about neural behavior and organization, which

are the primary quantities of interest in neuroscience. Second, the template-to-subject hierarchy accounts for variability within a population. This is because the subject data is generated probabilistically and tends to differ slightly between individuals. Finally, modeling the effects of a disorder within the latent structure is intuitive. In reality, the MRI scanner converts brain information to fMRI/DWI data. Although there will be deviations due to experimental noise, we do not expect the physics of this process to change when presented with a clinical subject. Rather, there is some abnormality within the (hidden) structural and/or functional workings of the brain.

1.2 Joint Model of Anatomical and Functional Connectivity

This thesis presents a generative model that combines information from resting-state fMRI and DWI data. Specifically, we define latent anatomical and functional connectivity variables, which specify a template organization of the brain. Anatomical connectivity indicates whether or not there is an underlying neurological connection between two regions. However, it does not quantify the number or trajectory of the corresponding white matter fibers. Functional connectivity describes how two regions co-activate (positive synchrony, negative synchrony or no relationship). Our multi-modal assumption is that latent anatomical connectivity influences the observed fMRI correlations. This relationship accounts for the neuro-scientific finding that a high degree of anatomical connectivity predicts higher functional correlations [41, 48]. We model the clinical population as a corrupted version of the healthy templates.

We efficiently estimate the templates of latent connectivity for each population using the EM algorithm [23]. The EM algorithm optimizes the model parameters by maximizing the data likelihood. In the process, it infers the posterior probability distribution of the latent variables, and consequently identifies differences between the groups.

We apply our model to a population study of schizophrenia. Implicitly, we require some amount of consistency across subjects to accurately infer the latent templates. Voxel-wise connectivity measures are too variable, especially for the DWI data. Instead, we rely on Brodmann areas to provide anatomically meaningful correspondences across subjects at the level of functional divisions within the brain. These regions are also large enough to ensure consistent tractography across subjects. We extract fMRI and DWI measures between pairs of regions. This process is illustrated in Fig. 1-2.

We demonstrate that our model learns stable and predictive connectivity differences attributed to the disease.

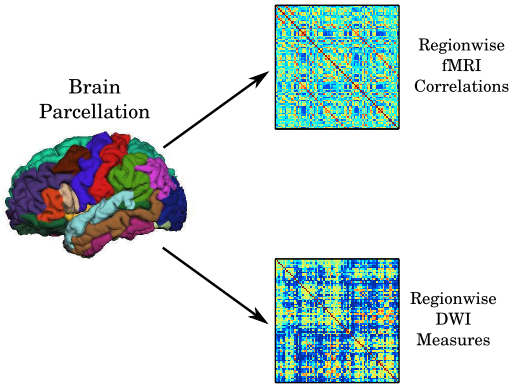


Figure 1-2: Graphical illustration of connectivity measures in a single subject. We partition the brain into N consistent regions, as shown on the left. We then extract symmetric $N \times N$ matrices of fMRI and DWI measures, as suggested on the right. The fMRI correlations are real-valued in the range $[-1, 1]$. The DWI measures are either zero, which represents a missing connection, or real-valued in the range $[0.2, 0.6]$.

1.3 Localizing the Effects of a Neurological Disorder

Our initial model demonstrates that neuropsychiatric disorders can be linked to aberrations in connectivity. However, the results are somewhat difficult to interpret and validate. In particular, the bulk of our knowledge about the brain is organized around regions (i.e., functional localization, tissue properties, morphometry) and not the connections between them. Moreover, it is nearly impossible to design non-invasive experiments that target a particular connection between two brain regions. These drawbacks are not unique to our specific framework; rather, they are present in nearly every clinical study of connectivity [52, 53, 96]. We address this problem by proposing a unified framework that pinpoints regions, which we call *foci*, whose connectivity patterns are most disrupted by a disorder.

Our second model assumes that the effects of a disorder are localized to a small collection of brain regions. The resulting impairments engender pathways of abnormal functional connectivity emanating from the disease foci. This, in turn, causes the observed fMRI correlations to differ between a control and a clinical population. Due to the global interaction among latent variables, we employ a variational EM algorithm [57] to fit the model to the data. Our method effectively aggregates population differences in connectivity to jointly infer the diseased regions and the induced connectivity differences.

We consider two versions of the region-based formulation. The first variant is based on the functional information. Here, we consider abnormalities on the complete graph of pairwise functional connections. The second variant uses neural anatomy as a substrate

for modeling functional connectivity. Specifically, we constrain the graph of aberrant functional connections to coincide with latent anatomical pathways in the brain. Once again, we use DWI data to infer these underlying tracts. Since neural communication between brain regions is constrained by white matter fibers, our driving assumption is that the strongest effects of a disorder will occur along direct anatomical connections. Although we model whole-brain connectivity, we rely on functional abnormalities between anatomically connected regions to identify the disease foci.

We demonstrate that our methods learn a stable set of afflicted regions in schizophrenia. Our results identify the posterior cingulate, the superior temporal gyri and the transverse temporal gyri as the most affected regions in schizophrenia. These regions have been confirmed in the schizophrenia literature [37, 59].

1.3.1 Multi-Class Region Labels

A natural extension is to consider multiple sets of foci, distributed throughout the brain, that are collectively responsible for abnormal neural communication.

Our revised model assumes that the diseased regions can be partitioned into H groups, which we denote *disease clusters*; each cluster is responsible for a separate graph of abnormal functional connectivity. For simplicity, we focus on the fMRI-based model. The conditional distributions differ slightly from the previous model due to the multi-class region assignments. However, we derive a corresponding variational EM algorithm to estimate the model parameters.

We report preliminary results on our schizophrenia dataset when varying the number of disease clusters H. Interestingly, we observe a nesting property, whereby the disease clusters for smaller values of H are subsets of those for larger H. The core disease cluster includes the significant regions identified by the initial functional model.

1.4 Schizophrenia Dataset

In this thesis we focus on schizophrenia as a representative neurological disease. Schizophrenia is a poorly-understood disorder marked by widespread cognitive difficulties affecting intelligence, memory, and executive attention. These impairments are not localized to a particular cortical region, but rather, they reflect abnormalities in widely-distributed functional and anatomical networks [16, 37]. The current hypothesis is that schizophrenia disturbs the interaction (i.e. connectivity) between regions of the brain [34, 63]. Hence, this condition is well-suited to our whole-brain connectivity framework.

1.5 Contributions of This Thesis

To summarize, this thesis introduces three main contributions that advance the field of connectivity analysis for clinical applications:

1. Hierarchical generative models that use observed fMRI/DWI data to make inferences about hidden processes in the brain. Differences induced by a neurological disease are represented in the hidden layer.

2. Multi-modal analysis of resting-state fMRI and DWI data to identify abnormal connections distributed throughout the brain.

3. Aggregating population differences in connectivity to localize foci of a disease. We further extend this framework to identify clusters of diseased regions that together influence whole-brain connectivity patterns.

1.6 Thesis Outline

The next chapter reviews background material, including the two imaging modalities used in this work and their corresponding analysis methods. We also summarize the clinical findings of schizophrenia and outline the graphical model notation used throughout this thesis. Chapter 3 introduces our joint generative model to identify population differences in connectivity based on fMRI and DWI data. We build upon our initial framework in Chapter 4 to pinpoint foci of a neurological disorder. Chapter 5 develops an extension to our region-based model that assumes multiple clusters of abnormal activity. Finally, we conclude by discussing the implications of this dissertation and proposing future directions of research in Chapter 6.

Chapter 2

Background

Functional and anatomical imaging modalities provide complementary viewpoints of the brain. In this chapter we outline the evolution of connectivity analysis from the traditional expert-driven approach to data-driven models that incorporate multiple sources of information. We then summarize the literature surrounding schizophrenia and describe our clinical dataset. We conclude with an overview of the probabilistic graphical model notation used in this thesis.

2.1 Functional Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging (fMRI) captures vascular effects in the brain that are associated with changes in blood oxygenation. Specifically, oxygenated hemoglobin in the blood is diamagnetic and has different magnetic properties than de-oxygenated hemoglobin, which is paramagnetic. Blood Oxygen Level Dependent (BOLD) fMRI measures local shifts in oxygenation over time using a T2*-weighted imaging protocol. T2* relaxes more slowly in oxygen-rich regions, resulting in higher signal intensity [77]. Similar to other MRI-based imaging modalities, fMRI achieves a reasonable spatial resolution (2-5mm^3); however, the temporal resolution of the signals is limited (1-5 seconds between volumes).

The vascular signal obtained from BOLD fMRI provides an indirect measure of neural activity. It has been shown that active regions of the brain exhibit locally increased blood flow and oxygen metabolism, which may be linked to heightened energy utilization during neurological processes [46]. However, the precise relationship between the hemodynamic and the underlying neural signals is not well understood.

2.1.1 Functional Localization

Traditional fMRI studies assess the response to a given experimental paradigm in order to localize brain functionality. This *task-based* data is usually analyzed with the General Linear Model (GLM) [35, 36], which assumes a linear contribution of each experimental condition to the fMRI time courses. Mathematically, let y_i represent the observed fMRI

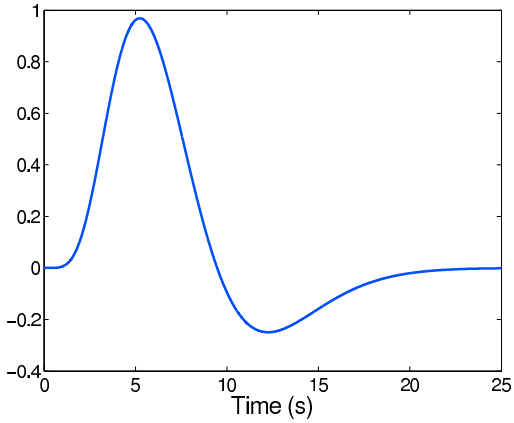


Figure 2-1: Parametric form of the Hemodynamic Response Function.

time course at spatial location i, and let X be the *design matrix* which contains temporal information about each experimental condition. The GLM uses the regression framework:

$$y_i = \mathrm{X}\beta_i + \epsilon_i, \tag{2.1}$$

where β_i is the unknown activation vector that contains the magnitudes of response to each stimulus, and ϵ_i is modeled as white Gaussian noise. The standard least-squares solution to Eq. (2.1) is given by

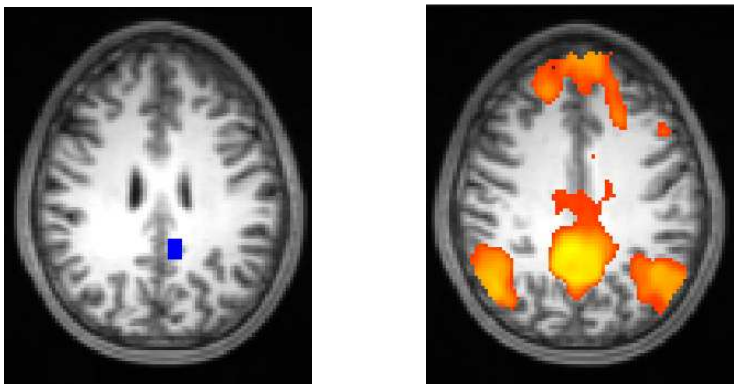
$$\hat{\beta}_i = (\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}\mathrm{X}^{\mathrm{T}}y_i. \tag{2.2}$$

A high-valued entry in $\hat{\beta}_i$ indicates that voxel i reacts strongly to the corresponding stimulus. Hence, the entries of $\hat{\beta}_i$ inform us about the role of voxel i in the brain.

Task-based fMRI assumes a parametric form for the vascular impulse response, known as the Hemodynamic Response Function (HRF) [36], which is depicted in Fig. 2-1. The HRF is convolved with the experimental protocol to obtain columns of the design matrix X. In contrast, *resting-state* fMRI measures spontaneous oscillations in the absence of any experimental paradigm [14, 32]. Although we can no longer use the GLM to solve for activation patterns, correlations within resting-state signals are believed to reflect the intrinsic functional connectivity of the brain [8]. The remainder of this thesis will focus exclusively on resting-state fMRI data.

2.1.2 Exploring Functional Coherence

Recent studies based on fMRI have revealed the presence of spontaneous, low-frequency (<0.08 Hz) fluctuations in the brain [8]. While independent of external cognitive stim-



(a) Posterior cingulate seed (b) Highly correlated voxels

Figure 2-2: Example of seed based correlation analysis. We specify (a) a seed in the posterior cingulate region of the default network and (b) identify voxels strongly correlated with the seed time course.

uli, these signals are strongly correlated across brain structures. Functional connectivity analysis aims to detect and characterize these coherent patterns of activity as a means of understanding the organization of the brain.

Expert-Driven Analysis

Seed based correlation analysis [8] is the most common approach for assessing functional connectivity. Here, an expert specifies a 'seed' region of interest (ROI) $\mathcal{R}$ within the desired functional system (e.g., motor cortex, visual cortex, default network). The seed is typically a small ball (3-5 voxels in diameter) embedded within the gray matter tissue. Let $\bar{y}_{\mathcal{R}} = \frac{1}{|\mathcal{R}|}\sum_{j\in\mathcal{R}} y_j$ denote the mean fMRI time course computed over all voxels j in $\mathcal{R}$. We can obtain the correlation coefficients

$$\rho_i = \frac{\langle y_i, \bar{y}_{\mathcal{R}} \rangle}{||y_i||\ ||\bar{y}_{\mathcal{R}}||} \tag{2.3}$$

between the mean signal $\bar{y}_{\mathcal{R}}$ and time course $\mathbf{y}_i$ for each voxel i in the brain. Functional systems are identified by thresholding the correlation coefficients at a user-specified value. Fig. 2-2 illustrates this procedure using a seed in the posterior cingulate cortex.

Seed based analysis has been extremely useful in identifying brain systems reliably across subjects. It has also shown promise in discovering new systems. One such example is the default network [12, 14], which is active when an individual is resting or performing an internally focused task, and it deactivates during sensory-motor experiments. The default network is believed to mediate internal cognition about the environment or past/future

events and may even play a role in memory formation [14]. The posterior cingulate cortex in Fig. 2-2 lies within the default network.

Data-Driven Methods

Despite being a valuable tool, seed based analysis is limited by our ability to specify the initial ROI. Furthermore, the choice of threshold can significantly affect the consistency of the results within a population. Currently, there is no universally-accepted method to select a threshold. In response to these limitations, data-driven approaches, such as clustering [19, 39, 94, 102] and Independent Component Analysis (ICA) [2, 17], are gaining prominence. Here, the aim is to partition the brain into different functional systems.

One simple but effective clustering method for resting-state fMRI is the K-Means algorithm [39, 94]. This method assumes that the time course y_i of voxel i is drawn from one of K multivariate Gaussian distributions, each with a unique mean m_k and a spherical covariance:

$$y_i = m_{k(i)} + \epsilon_i, \quad \forall i = 1, \ldots, N, \tag{2.4}$$

where N is the number of voxels in the volume, $k(i)$ is the cluster assignment for voxel i, and ϵ_i is $i.i.d.$ Gaussian noise. The K-Means algorithm alternates between assigning each voxel to the closest mean, as measured by the L^2 Euclidean distance

$$d^2(y_i, m_k) = ||y_i - m_k||^2, \quad \forall k = 1, \ldots, K, \tag{2.5}$$

and recomputing the mean signals of each cluster as the average of all time courses assigned to it. This naïve implementation is a natural data-driven extension to seed based analysis. Specifically, resting-state fMRI time courses are often normalized to have zero mean and unit variance. Therefore, $d^2(y_i, m_k) = 2 - 2\rho(y_i, m_k)$. Hence, minimizing L^2 distance is equivalent to maximizing the correlation coefficient.

Spectral Clustering is an alternative technique that relies on a pairwise affinity matrix constructed from the data points [76, 98]. The eigenvectors of this matrix induce a low-dimensional representation for the data, which in turn, encourages a natural grouping. This approach does not presume any parametric form for the data and can identify clusters with complex signal geometries [76]. Elements of the symmetric pairwise affinity matrix W are often modeled using a Gaussian kernel:

$$\mathrm{W}_{ij} = e^{-d^2(y_i, y_j)/2\sigma^2}, \tag{2.6}$$

where y_i and y_j represent two voxel time courses, $d^2(y_i, y_j)$ is given by Eq. (2.5), and σ^2 is the kernel width parameter.

Given the affinity matrix W, Normalized Cut Spectral Clustering [82] partitions the dataset to minimize the ratio of the sum of affinities W_{ij} between clusters to the sum of affinities within a cluster. We can formulate a continuous relaxation of this combinatorial problem via the eigenvalue equation:

$$\mathrm{D}^{-1/2}\mathrm{W}\mathrm{D}^{-1/2}z = \lambda z, \tag{2.7}$$

where D is a diagonal matrix such that $\mathrm{D}_{ii} = \sum_j \mathrm{W}_{ij}$. The left and right multiplications by $\mathrm{D}^{-1/2}$ in Eq. (2.7) correspond to a symmetric normalization of W where each entry W_{ij} is divided by $\sqrt{\mathrm{D}_{ii}\mathrm{D}_{jj}}$. The largest eigenvectors $\{z_1, \ldots, z_{K+1}\}$ of the matrix $\mathrm{D}^{-1/2}\mathrm{W}\mathrm{D}^{-1/2}$ tend to isolate voxel groups with small pairwise L^2 distances. The low-dimensional representation can be easily clustered using the simple K-Means algorithm.

In contrast to clustering methods, Independent Component Analysis (ICA) [51] assumes that resting-state fMRI data is a linear mixture of K spatially independent sources [2, 17]. Mathematically, let N be the number of voxels in the brain and T be the number of time points. The $T \times N$ data matrix $\mathrm{Y} = [y_1, \ldots, y_N]^{\mathrm{T}}$ is the product of a $T \times K$ mixing matrix M and a $K \times N$ component matrix C, i.e.,

$$\mathrm{Y} = \mathrm{MC}. \tag{2.8}$$

The rows of C contain independent spatial maps of the brain, and the matrix M specifies the time-varying contributions of each component. The goal of ICA is to determine a $K \times T$ *unmixing* matrix B such that

$$\mathrm{C} \approx \mathrm{BY}, \tag{2.9}$$

and such that the rows of C are maximally independent.

The InfoMax algorithm [65] selects the optimal unmixing matrix B^* by maximizing the joint entropy of the transformed component map estimates $\hat{\mathrm{C}} = g(\mathrm{Q})$, where

$$\mathrm{Q} = 2\mathrm{B}(\mathrm{Y}\mathrm{Y}^{\mathrm{T}})^{-1/2}\mathrm{Y} \tag{2.10}$$

is a scaled version of the data, and $g(\cdot)$ is the logistic function

$$g(z) = \frac{1}{1 + \exp^{-z}} \tag{2.11}$$

applied element-wise to each entry in the matrix Q. The highly nonlinear transformation between Q and $\hat{C}$ preserves higher order statistical information. The InfoMax algorithm is implemented iteratively; elements of B are updated in small batches.

Other methods identify functionally coherent regions by incorporating population-level information [92, 93]. For example, the method of [92] uses a dictionary learning framework to estimate a functional atlas. Similarly, the authors of [93] improve subject-wise functional connectivity estimates by imposing a common sparsity structure.

Limitations

The cluster assignments and spatial component maps estimated using data-driven methods are very similar to functional networks obtained from seed based analysis. However, unlike the expert approach, these results are often inconsistent across subjects [39]. Consequently, we can only identify large sensory-motor systems across individuals in a population. The specificity does not improve much with larger datasets. For example, a comprehensive study of functional connectivity was conducted using 1000 subjects [102]. The most refined parcellation consisted of 17 clusters, which is still quite coarse. The problem is exacerbated in a clinical setting, as the diseased population may no longer be homogeneous.

In this thesis we use a template-based framework in which all subjects share the same global organization. Subject variations are handled using a probabilistic setting. Our goal is to characterize patterns of connectivity within the brain and to understand how these interactions change in the presence of a neurological disorder. We do not attempt to delineate elements of the underlying network, i.e., regions. Instead, we specify them *a priori* based on a standard anatomical parcellation.

2.2 Diffusion Weighted Imaging

Diffusion Weighted Imaging (DWI) characterizes the anisotropic diffusion of water as it traverses soft tissue. Specifically, water diffuses more freely along rather than across white matter fiber bundles in the brain. This effect sheds light onto the structural organization of the brain. Since its introduction to the neuroscience community, DWI has been useful in a variety of clinical applications. It has been particularly successful in pinpointing abberations generated by white matter diseases such as multiple sclerosis and Alzheimer's disease [50]. It has also been used to localize the origins of acute-phase stroke [101].

A single DWI volume is obtained by applying a magnetic pulse sequence in a particular gradient direction v_k. The resulting intensity S_k in a particular voxel is modeled via the

modified Stejskal-Tanner Equations [5, 7]:

$$S_k = S_0 e^{-\mathrm{b} \cdot v_k^{\mathrm{T}} \mathrm{D}\, v_k}, \tag{2.12}$$

where S_0 is the intensity at the same voxel obtained with no gradient pulse, and the symmetric positive semi-definite *diffusion tensor* D reflects the directional mobility of water. The *b-value* b is pre-calculated based on the gradient pulse characteristics (timing, amplitude, shape). The values of S_0 and D vary for each voxel in the brain; however, b is often constant for the entire acquisition.

By collecting several images from unique gradient directions, one can estimate the entries $\mathrm{D}_{ij} = \mathrm{D}_{ji}$ of the symmetric diffusion tensor in Eq. (2.12).

2.2.1 Assessing Anatomical Connectivity

Scalar measures of connectivity can be derived from the diffusion tensor D estimated at each voxel. Let $\lambda_1 > \lambda_2 > \lambda_3$ denote the eigenvalues of D. The mean diffusivity, computed as $\bar{\lambda} = \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3)$, quantifies the aggregate amount of diffusion in a given voxel. In contrast, anisotropy measures the degree to which D deviates from a spherical (isotropic) shape. For example, Fractional Anisotropy (FA) is defined as follows:

$$FA = \frac{\sqrt{3[(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \hat{\lambda})^2]}}{\sqrt{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}.$$

In this thesis, we use FA as a scalar measure of connectivity for DWI data.

FA is a sensitive measure that is affected by a number of biological processes. For example, changes in myelination as well as inflammation of the underlying white matter fibers can drastically alter the local FA values. For clinical applications, statistical differences in anisotropy are often used to identify the effects of a neurological disease [50, 63].

The diffusion tensor can also be used to estimate white matter bundles via a procedure called tractography. Early work in this area focused on two broad classes of tractography algorithms. *Streamline tractography* creates a vector field using the principal diffusion direction (the eigenvector associated with the largest eigenvalue of D) coupled with smoothness constraints. A white matter tract corresponds to the path a particle would follow in this vector field from a given starting location [4]. In contrast, *stochastic tractography* generates a probability map of diffusion starting at a given location. The fiber tracts are modeled as sequences of unit vectors. Each vector is sampled from a posterior distribution, which combines a prior term based on the previous direction and a likelihood that is related to

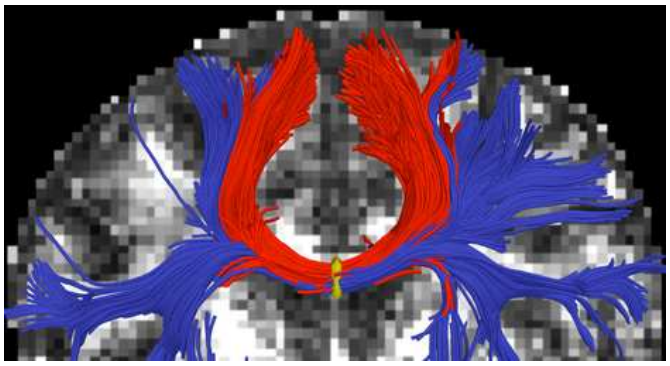


Figure 2-3: Results from two-tensor (blue) and single-tensor (red) tractography when seeded in the mid-saggital plane. The two-tensor method identifies more fiber trajectories.

the diffusion tensor D [33]. Stochastic tractography generates a large number of tracts from each starting location. The goal of subsequent processing is to provide a rough estimate of the underlying fibers and to quantify the uncertainty.

Given the tractography results, it is common to compute anatomical connectivity along the estimated white matter pathways. Such measures include the probability of diffusion between two brain regions, the number of fibers linking the regions, and the mean FA along the tracts connecting them.

There are drawbacks to both streamline and stochastic tractography. Traditional streamline algorithms suffer from diminished performance in regions with multiple fiber orientations. In particular, when two fibers cross or merge, the estimated diffusion tensor is nearly isotropic, and the algorithm may not follow the correct trajectory. Although stochastic tractography is more sensitive to multiple fibers, it has difficulty finding long-range connections. Since the algorithm samples each step, the probability of connecting two points tends to decay with the distance between them. In addition, stochastic tractography tends to find and follow large white matter bundles, which may not capture the full network of connections.

Recently, more sophisticated tractography methods have been proposed. For example, weighted mixtures and higher order tensors are used to handle complex diffusion patterns. Similarly, non-parametric methods estimate an Oriented Distribution Function (ODF) at each voxel to describe an arbitrary fiber configuration [1, 24]. In this thesis, we employ a two-tensor tractography algorithm [71], which relies on an unscented Kalman filter to fit a local neural fiber model. Fig. 2-3 illustrates that the two-tensor algorithm (blue) finds a denser set of connections than streamline tractography (red).

2.3 Traditional Multi-modal Analysis of Brain Connectivity

Early work in multimodal analysis focused on the relationship between task fMRI activations and the underlying anatomy. One popular technique is to use regions of fMRI activation as seed points for tractography [26, 42, 79]. Another approach is to quantify the relationship between anatomical connectivity and measures of functional co-activation in pre-defined regions of interest [55, 89]. Presently, the focus has shifted to resting-state fMRI for joint analysis [80]. Studies independently compute statistics of the fMRI and DWI signals (such as fMRI correlations, FA values, etc.) and search for correspondences between these metrics *a posteriori* [41, 48, 61]. For example, they identify connections with highly correlated anatomical and functional measures or connections along which the fMRI and DWI metrics are uniformly large.

These methods have yielded many insights into the nature of connectivity in the brain. For example, fMRI-guided tractography has improved the mapping of the motor, visual and language areas [26, 42, 79]. It has also been established that while a high degree of anatomical connectivity predicts higher functional correlations, the converse does not always hold [41, 48]. For example, strong functional correlations can be found between spatially distributed locations in the brain, whereas one is more likely to identify white matter tracts connecting nearby regions. We incorporate this latter finding into our joint connectivity model presented in the next chapter.

The main limitation of these simplistic approaches is that the analysis is performed separately on each modality, and information is later pooled into a joint representation. Furthermore, the methods perform independent tests for each connection. This ignores distributed *patterns* in functional and anatomical connectivity within the brain. We address these limitations by assuming that the structure and organization of the brain is captured by some underlying generative process, which gives rise to the observed fMRI/DWI measures. We use *both* modalities to infer population templates of connectivity and demonstrate that our method captures stable differences in a clinical population.

2.4 Exploring Network Structure of the Brain

Alternative methods have emerged to address the independence across connections, both in unimodal and multi-modal applications. These approaches treat the brain as a comprehensive network with interactions between nodes. Here, each node represents a voxel or region of interest, and edges are derived from fMRI correlations and DWI measures. Analysis of

these networks has revealed a consistent small-world (structural and functional) architecture of the brain. This organization implies clusters of tightly coupled nodes with a few long-range connections between them. One property of small-world networks is a short average path length, which indicates higher communication efficiency within the network [15, 84].

Standard graph metrics (centrality, clustering coefficient, degree distributions) have also been applied to the networks implied by imaging measures. This analysis revealed functional hubs in the default network [12, 13] and anatomical hubs in the medial cortex [43]. These hubs represent relay stations within the brain and facilitate communication between regions. Abnormalities within such hubs may be linked to disease.

Recent work also considered the relationship between the brain's structural organization and functional dynamics. It has been suggested that nodes with similar anatomical connectivity patterns tend to exhibit similar functionality. Additionally, structural connectivity should constrain the functional interactions, as all information is transmitted via neurons. The approach of [85] goes one step further and computationally models the interaction. The authors construct cortical connection graphs based on histological data of the macaw brain. They simulate the corresponding functional correlations using a dynamical system, which specifies the relationships within the graph. Although promising, this analysis has not been replicated using *in vivo* data from humans. The work of [48] explores how well the anatomical network structure explains large-scale properties of functional systems. Their results confirm previous findings that the presence of an anatomical connection is informative for prediction functional correlations, and that both anatomical and functional connectivity measures tend to decay with the distance between two regions. The results also suggest that multi-stage anatomical links explain some of the high functional correlations.

The method of [21] is the only alternative to our joint connectivity model that proposes a generative framework. The collection of fMRI time courses is modeled as a sample from a stationary autoregressive process. The observed covariance matrix is constrained by the anatomical connectivity information. Once the functional interactions are determined, the authors use a multiple linear regression framework to determine which anatomical connections contribute to a particular functional correlation.

Despite the advancements in multi-modal analysis, prior work focuses entirely on connections. In Chapter 4 we carry the analysis one step further and infer properties of individual brain regions. To the best of our knowledge, ours is the first generative framework to aggregate information from resting-state fMRI and DWI data to localize region effects.

2.5 Population Studies of Connectivity

Resting-state fMRI and DWI data are particularly attractive for clinical populations since patients are not required to perform challenging experimental paradigm. Univariate tests and random effects analysis are, to a great extent, the standard in population studies of connectivity [40, 66, 103]. These methods identify significantly different connections using a statistical score. Typically, the scores are computed independently for each functional correlation or DWI measure. Consequently, the analysis cannot capture abnormalities in distributed networks of connectivity within the brain.

Prior work has also explored multi-pattern analysis for functional connectivity [52,53,91, 96]. For example, [91] develops a random effects model for covariance matrices to identify functional connectivity differences in stroke patients. In contrast, [52, 53] employ group Independent Component Analysis (gICA) to represent the fMRI data as a set of spatially-independent regions with associated time courses. In [53], group functional connectivity is computed as the maximum lagged correlation between the component time courses; two-sample t-tests are used to identify significant population differences. In [52], a neural network is constructed for patient classification of first-episode schizophrenia. Similarly, we previously used a metric called Gini Importance to summarize multivariate patterns of interaction [96]. A patient classifier trained on these measures demonstrated superior classification accuracy than when trained on univariate statistics. Further details of this method are provided in the Appendix. In Chapter 3, we present a probabilistic framework for connectivity analysis. Differences between two populations are explained via changes in latent anatomical and functional connectivity variables. Connections are tied through global parameters, which modulate distributed effects of the disorder.

Although these studies identify connectivity patterns associated with a disease, the results are difficult to interpret given that much of our knowledge about the brain is organized around regions (i.e., functional localization, tissue properties) and not the connections between them. Furthermore, short of direct stimulation, we do not know how to design *in vivo* experiments that target a particular connection between two brain regions. Hence, validation of these findings is challenging, if not impossible. Our framework in Chapter 4 integrates population changes in functional connectivity to localize foci of a disorder.

2.6 Schizophrenia: Findings and Hypotheses

Schizophrenia is a neuropsychiatric disorder characterized by gross distortions in the perception of reality. Despite generating considerable interest in the neuroscience community,

the origins and expression of the disease are still poorly understood [87]. For example, structural findings only weakly and inconsistently correlate with the clinical and cognitive symptoms of schizophrenia [81]. Similarly, functional experiments report deficits in many cognitive domains, most notably memory and attention, but do not consistently report clinical correlates [74].

At present, the cognitive impairments of schizophrenia are thought to reflect underlying abnormalities in distributed brain networks. In particular, schizophrenia may compromise neural communication between cortical regions [34]. Recent studies have also focused on the degeneration of anatomical connectivity [63], fueled in part by post mortem and genetic evidence of myelination anomalies in patients with schizophrenia.

Findings from resting-state fMRI data include reduced connectivity in the brain's default network [9, 12], dorsolateral prefrontal cortex [103] and a widespread reduction in connectivity throughout the brain [66]. In contrast, although the majority of DWI studies report white matter abnormalities, there is no consensus on the location and nature of these changes [62, 63]. The most commonly reported anomalies are between frontal and temporal lobes and between the hemispheres [63]. They are believed to reflect the underlying neuro-developmental or neuro-degenerative processes affecting myelinated axons. Since DWI tractography and resting-state fMRI data provide different information about the underlying structure and dynamics of the brain, we believe that joint analysis of these modalities will improve our understanding of brain connectivity and of the effects that schizophrenia has on the connectivity pattern.

2.7 Schizophrenia Dataset

In this section we outline the acquisition protocol used to collect the data and our subsequent pre-processing steps. The fMRI/DWI connectivity measures are presented in later chapters.

2.7.1 Image Acquisition

The study collects data from 19 male patients with chronic schizophrenia and 19 healthy male controls. The control participants were group matched to the patients on age, handedness, parental socioeconomic status, and an estimated premorbid IQ. For each subject, an anatomical scan (SPGR, $TR = 7.4s$, $TE = 3ms$, $FOV = 26cm^2$, $res = 1mm^3$), a diffusion-weighted scan (EPI, $TR = 17s$, $TE = 78ms$, $FOV = 24cm$, $res = 1.66 \times 1.66 \times 1.7mm$, 51 gradient directions with $b = 900s/mm^2$, 8 baseline scans with $b = 0s/mm^2$) and a resting-state functional scan (EPI-BOLD, 196 volumes with $TR = 3s$, $TE = 30ms$, $FOV = 24cm$, $res = 1.875 \times 1.875 \times 3mm$) were acquired using a 3T GE Echospeed system.

2.7.2 Pre-Processing

SPGR We segment the structural images into 77 cortical and subcortical regions using Freesurfer [29]. The parcellation roughly corresponds to Brodmann areas in the brain. We then align each subject's structural image to the the corresponding DWI and fMRI volumes via non-rigid registration [3, 70]. This allows us to define a consistent set of regions across modalities and across subjects. Since our procedure never aligns images across subjects, we avoid the challenges of group-wise registration [30, 58, 106].

DWI The DWI data is corrected for eddy-current distortions using the FSL FLIRT algorithm [54, 83]. A two-tensor tractography algorithm is used to estimate the white matter fibers [71] for each subject.

The DWI connectivity is extracted along the identified fibers. Specifically, the fiber defines a binary categorization of values depending on whether or not two regions are connected. We prefer this scheme over continuous measures of connection strength (ex. probability of diffusion), which tend to decay as the distance between two regions increases. This decay is observed even if there are several white matter tracts bridging the regions.

fMRI We discarded the first five fMRI time points (burn-in) and performed motion correction by rigid body alignment and slice timing correction using FSL [83]. The fMRI data is then fed through a standard functional connectivity pre-processing pipeline [8]. Specifically, each volume is spatially smoothed using a Gaussian filter, and each voxel time course is temporally low-pass filtered with $0.08Hz$ cutoff and motion corrected via linear regression. Finally, we removed global contributions to the time courses from the white matter, ventricles and the whole brain. The fMRI signal in the ventricle correlates with respiration [90] and degrades the signal quality. Similarly, regressing the white matter and whole brain signals reduce the noise present in the data.

The fMRI connectivity is obtained by correlating the mean time courses of each brain region. This is a common measure to assess region-based functional connectivity.

2.8 Graphical Model Notation

This thesis uses a probabilistic framework to analyze brain connectivity. Graphical models allow us to efficiently describe the relationships in our formulation. Here, we introduce the notation used throughout this thesis. We refer the reader to [56, 57, 99] for a comprehensive overview of probabilistic graphical models and the corresponding inference algorithms.

The graphical model representation captures the conditional independence structure

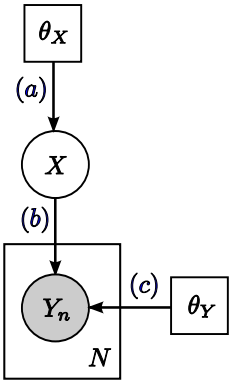


Figure 2-4: Example graphical model to illustrate our notation. Circles indicate random variables and squares denote non-random parameters. The shaded variables are observed.

between variables. Fig. 2-4 presents a simple example to highlight this property. Here, X and $\{Y_n\}$ are random variables, whereas θ_X and θ_Y are unknown non-random parameters. The shading indicates that X is a *latent* variable, but that we observe the values of $\{Y_n\}$. Note that $X, Y_n, \theta_X, \theta_Y$ need not be scalar. Instead, each symbol may represent a collection of random variables and parameters, respectively.

The arrows in Fig. 2-4 inform us about the conditional relationships. For example, arrow (a) signifies that the prior distribution of X depends on θ_X, i.e., $P(X;\theta_X)$. Similarly, arrows (b) and (c) specify the dependencies for Y_n. The plate at the bottom of Fig. 2-4 indicates that the random variables $Y_1, \ldots, Y_N$ are independent and identically distributed given X and θ_Y. Hence, the conditional distribution of $Y = \{Y_n\}$ factorizes as follows:

$$P(Y|X;\theta_Y) = \prod_{n=1}^{N} P(Y_n|X;\theta_Y). \tag{2.13}$$

We use a maximum likelihood framework to fit our models to the observed data. For the simple case presented in Fig. 2-4, we seek the optimal parameter values θ_X^*, θ_Y^* that maximize the data likelihood $P(Y;\theta_X,\theta_Y)$ after marginalizing the latent variables. Mathematically,

$$\{\theta_X^*, \theta_Y^*\} = \max_{\theta_X,\theta_Y} P(Y;\theta_X,\theta_Y) = \max_{\theta_X,\theta_Y} \int_X \left[P(X;\theta_X) \prod_{n=1}^{N} P(Y_n|X;\theta_Y) \right] dX \tag{2.14}$$

In this thesis we employ both the Expectation-Maximization (EM) algorithm and the Variational EM algorithm. In each case, we recover the posterior probability distribution of the latent variables $P(X|Y;\theta_X^*,\theta_Y^*)$ along with the model parameters.

2.9 Summary

To summarize, prior analysis of brain connectivity has relied on data statistics to identify patterns of interaction. In most cases, the correspondences between fMRI and DWI data, as well as the effects of a neurological disorder, are determined by univariate tests. While insightful, this type of analysis has produced widely varied results in clinical studies. Moreover, connectivity results are often difficult to interpret and validate. Nonetheless, few methods investigate the relationship between connections and individual brain regions.

In contrast to prior work, we take a generative approach to connectivity. In the next chapter we introduce a novel multi-modal framework to infer distributed patterns of abnormal connectivity induced by a neuropsychiatric disease. Chapter 4 builds on our original framework to aggregate population differences in connectivity to pinpoint regions that are most affected by a disorder. Finally, in Chapter 5 we extend the region-based model to multiple clusters of disease hubs.

Chapter 3

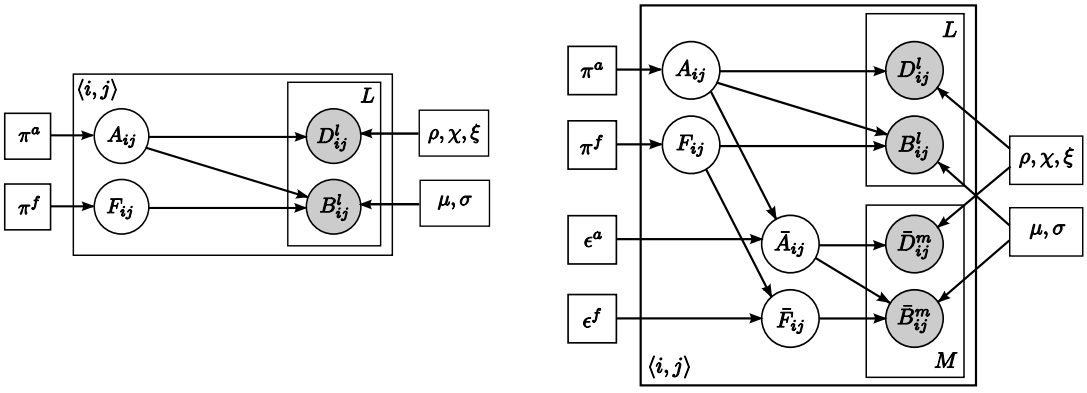
Modeling Anatomical and Functional Connectivity

As discussed in Chapter 2, the interaction between anatomical and functional imaging modalities offers a rich framework for understanding the effects of neuropsychiatric disorders. However, much of the prior research has focused on ad-hoc correspondences between independently computed statistics of the fMRI and DWI data. Additionally, clinical studies typically identify significant population differences separately within each modality and compare them *a posteriori*. This approach treats structural and functional connections as independent and ignores distributed patterns of connectivity.

In this chapter we propose and demonstrate a novel probabilistic framework to infer the relationship between resting-state fMRI correlations and DWI tractography. Specifically, we introduce the notion of *latent* anatomical and functional connectivity between brain regions. These variables represent an underlying process in the brain which cannot be observed directly from the data. The resulting model describes how the latent connectivity differs between two populations and makes intuitive assumptions about the fMRI and DWI generation process to construct the data likelihood. Our fMRI/DWI observation model is shared across subjects. Hence, we assume that the effects of a disorder can be explained via changes in latent anatomical and functional connectivity. To the best of our knowledge, ours is the first stochastic model to combine resting-state fMRI and DWI data to infer changes induced by a neurological disease.

We employ the EM algorithm to efficiently estimate templates of latent connectivity for each population and to identify group differences. The EM algorithm optimizes the model parameters by maximizing the data likelihood. We employ permutation tests and cross validation to verify the robustness of our method. We perform an extensive evaluation of the model on synthetic data. In addition, we learn stable patterns of interaction in a population study of schizophrenia. The work presented in this chapter was published in [97].

The remainder of this chapter is organized as follows. Section 3.1 introduces our gen-



(a) Joint fMRI/DWI Model (b) Model of Population Differences

Figure 3-1: (a) Joint connectivity model for a single population. A_{ij} represents the latent anatomical connectivity between regions i and j, and F_{ij} denotes the corresponding latent functional connectivity. D^l_{ij} and B^l_{ij} are the observed DWI and fMRI measurements, respectively, between regions i and j in the l^{th} subject. (b) Joint model for the effects of schizophrenia. The control population is generated according to the model in (a). The schizophrenia templates are identified by an overbar.

erative models; we develop the corresponding inference algorithms in Section 3.2. Section 3.3 presents the framework used for the empirical validation of our approach. Sections 3.4 and 3.5 report experimental results based on synthetic and clinical data, respectively. Section 3.6 discusses the behavior of our model, its advantages and its drawbacks.

3.1 Generative Model

We combine the DWI tractography and fMRI correlations in a unified generative model of the brain. Specifically, latent anatomical and functional connectivity specify a template organization of the brain for a given population. Anatomical connectivity indicates whether or not there is an underlying neurological connection between two regions. It does not quantify the number or trajectory of the corresponding white matter fibers. Functional connectivity describes how two regions co-activate (positive relationship, negative relationship or no relationship). We do not have access to these underlying variables. Rather, we observe noisy measurements via DWI tractography and resting-state fMRI correlations. Although DWI and fMRI signals vary across subjects, we assume they are generated probabilistically from a common latent template.

We first develop the probabilistic framework within a control population. This formulation serves as a foundation for modeling group differences, presented later in the chapter.

Table 3.1: Random variables (top) and non-random parameters (bottom) in the graphical models shown in Fig. 3-1. The latent variables are discrete; the observed variables and non-random parameters are continuous.

A_{ij}	Latent anatomical connectivity between regions i and j (control group)
$\bar{A}_{ij}$	Latent anatomical connectivity between regions i and j (clinical group)
F_{ij}	Latent functional connectivity between regions i and j (control group)
$\bar{F}_{ij}$	Latent functional connectivity between regions i and j (clinical group)
D_{ij}^{l}	Observed DWI measure between regions i and j in subject l (control group)
$\bar{D}_{ij}^{m}$	Observed DWI measure between regions i and j in subject m (clinical group)
B_{ij}^{l}	Observed fMRI correlation between regions i and j in subject l (control group)
$\bar{B}_{ij}^{m}$	Observed fMRI correlation between regions i and j in subject m (clinical group)
π^a	Prior for binary anatomical connectivity A_{ij}
π^f	Prior for multinomial functional connectivity F_{ij}
ϵ^a	Probability of change in anatomical connectivity
ϵ^f	Probability of change in functional connectivity
ρ_t	Probability of failing to find a white matter tract given $A_{ij} = t$ $(t = 0, 1)$
χ_t	Mean DWI value if there is a white matter tract given $A_{ij} = t$ $(t = 0, 1)$
ξ_t^2	Variance of DWI values if there is a white matter tract given $A_{ij} = t$ $(t = 0, 1)$
μ_{tk}	Mean fMRI value given $A_{ij} = t$ and $F_{ij} = k$ $(t = 0, 1,\ k = -1, 0, 1)$
σ_{tk}^2	Variance of fMRI values given $A_{ij} = t$ and $F_{ij} = k$ $(t = 0, 1,\ k = -1, 0, 1)$

3.1.1 Single Population Model

Fig. 3-1(a) depicts our model for a single population, and Table 3.1 summarizes our notation. The individual subject data arises from the latent population templates of connectivity. All latent and observed variables are generated independently for each pairwise connection; the data likelihood parameters are shared across connections.

Prior Let N be the total number of regions in the brain. We use A_{ij} and F_{ij} to denote the latent anatomical and functional connectivity indicators between region i and region j $(1 \leq i, j \leq N)$. The anatomical connectivity A_{ij} indicates the presence or absence of a *direct* anatomical pathway between two regions. We model A_{ij} as a binary random variable; the scalar parameter π^a specifies the *a priori* probability that a connection is present:

$$P(A_{ij}; \pi^a) = (\pi^a)^{A_{ij}} (1 - \pi^a)^{1 - A_{ij}}. \tag{3.1}$$

In contrast, the functional connectivity F_{ij} is a tri-state random variable. These states represent (1) little or no functional co-activation ($F_{ij} = 0$), (2) positive functional synchrony ($F_{ij} = 1$), and (3) negative functional synchrony ($F_{ij} = -1$) between two regions. Strong

negative correlations are often found in resting-state fMRI data. Since there is no consensus about their origin and significance [90], we isolate negative connectivity as a separate category. For notational convenience, we represent F_{ij} as a length-three indicator vector $F_{ij} = [F_{ij,-1} \quad F_{ij0} \quad F_{ij1}]^T$ with exactly one of its elements equal to one:

$$P(F_{ij}; \pi^f) = \prod_{k=-1}^{1} (\pi_k^f)^{F_{ijk}}, \tag{3.2}$$

where π_k^f is the probability that the functional connection between region i and region j is assigned to state k.

Although we model latent connectivity via discrete random variables, the posterior probability distributions of the variables $\{A, F\}$ provide a natural measure of connection strength. These distributions form the basis for subsequent analysis in population studies.

Below, we describe how the latent connectivity templates affect the observed measures in individual subjects. Empirically, we observe that the variability of the DWI and fMRI measures of connectivity across connections and across subjects can be reasonably approximated using Gaussian distributions (Section 3.5.1 provides more details). It is not surprising since both measures are computed as averages of the observed image data and should therefore approach Gaussian distributions as the number of elements increases. Moreover, using Gaussian likelihoods for the observed data greatly simplifies the learning/inference algorithm and allows for efficient fitting of the model parameters.

DWI Likelihood Let L be the number of subjects. The DWI measurement D_{ij}^l between regions i and j in the l^{th} subject depends on the anatomical connectivity A_{ij}. In this work we use the average FA along white matter fibers to assess DWI connectivity. The model can be readily extended to accommodate other measures of connectivity by redefining the data likelihood term below.

Our observation model for D_{ij}^l explicitly accounts for errors in tractography. These include missing tracts between anatomically connected regions and spurious tracts between isolated ones. In particular, if tractography identifies one or more white matter fibers between two regions, the value of D_{ij}^l is modeled as a Gaussian random variable whose mean and variance depend on the anatomical connectivity indicator A_{ij}. Otherwise, D_{ij}^l is set to zero. Mathematically,

$$P(D_{ij}^l | A_{ij} = t; \{\rho, \chi, \xi\}) = \rho_t \delta(D_{ij}^l) + (1 - \rho_t) \mathcal{N}(D_{ij}^l; \chi_t, \xi_t^2), \tag{3.3}$$

where $\delta(\cdot)$ is the Dirac delta function and $\mathcal{N}(\cdot\,;\chi,\xi^2)$ denotes a Gaussian distribution with mean χ and variance ξ^2. ρ_0, ρ_1 are the probability of failing to find a white matter tract between the regions in the absence ($A_{ij}=0$) or presence ($A_{ij}=1$) of a latent anatomical connection, respectively. Ideally, $\rho_0 = 1$ and $\rho_1 = 0$, i.e., a white matter tract should be found if and only if there is an underlying anatomical connection. However, detection via tractography is imperfect. In practice D^l_{ij} is strictly positive if a tract is found between the regions. The Gaussian distribution in Eq. (3.3) adequately captures the data variation as our results in Section 3.5.1 suggest.

fMRI Likelihood We model the BOLD fMRI correlation B^l_{ij} between regions i and j in the l^{th} subject as a Gaussian random variable whose mean and variance depend on *both* the latent functional connectivity F_{ij} and anatomical connectivity A_{ij}. This reflects the finding that direct anatomical connections predict higher functional correlations [48, 61]:

$$P(B^l_{ij}|A_{ij}=t, F_{ij}=k; \{\mu,\sigma\}) = \mathcal{N}(B^l_{ij}; \mu_{tk}, \sigma^2_{tk}). \tag{3.4}$$

In this work we compute B^l_{ij} using Pearson correlation coefficients. Once again, our empirical analysis in Section 3.5.1 suggests that the Gaussian likelihood in Eq. (3.4) provides a reasonable approximation for the data distribution [1].

Combining all the elements of the model in Eqs. (3.1)-(3.4), we obtain the joint log-likelihood of all hidden and observed variables for a single connection $\langle i,j\rangle$:

$$\begin{aligned}
&\log P(A_{ij}, F_{ij}, \mathbf{D}_{ij}, \mathbf{B}_{ij}; \{\pi, \mu, \sigma^2, \rho, \chi, \xi^2\}) \\
&\quad = A_{ij}\log(\pi^a) + (1-A_{ij})\log(1-\pi^a) + \sum_{k=-1}^{1} F_{ijk}\log(\pi^f_k) \\
&\quad + (1-A_{ij})\sum_{l=1}^{L}\log\Big(\rho_0\delta(D^l_{ij}) + (1-\rho_0)\mathcal{N}(D^l_{ij};\chi_0,\xi^2_0)\Big) \\
&\quad + A_{ij}\sum_{l=1}^{L}\log\Big(\rho_1\delta(D^l_{ij}) + (1-\rho_1)\mathcal{N}(D^l_{ij};\chi_1,\xi^2_1)\Big) \\
&\quad + \sum_{l=1}^{L}\sum_{k=-1}^{1}\Big[(1-A_{ij})F_{ijk}\,\log\mathcal{N}(B^l_{ij};\mu_{0k},\sigma^2_{0k}) + A_{ij}F_{ijk}\,\log\mathcal{N}(B^l_{ij};\mu_{1k},\sigma^2_{1k})\Big].
\end{aligned} \tag{3.5}$$

[1] Although the Beta distribution is a natural model for correlation values, which are limited to the range $[-1,1]$, inferring the Beta parameters is more difficult and potentially more sensitive to noise

3.1.2 Population Differences

Fig. 3-1(b) presents an extension of our model to a population study involving normal controls and schizophrenia patients. We assume that the differences between the groups are explained entirely by changes in latent connectivity and that the two populations share the same data likelihood model.

In particular, we model the control population according to Fig. 3-1(a) and treat the latent connectivity templates $\{\bar{A}, \bar{F}\}$ of the schizophrenia population as a "corrupted" version of the healthy template. Specifically, with (small) probability, each connection can switch its state:

$$P(\bar{A}_{ij}|A_{ij};\epsilon^a) = (\epsilon^a)^{A_{ij}(1-\bar{A}_{ij})+(1-A_{ij})\bar{A}_{ij}} \cdot (1-\epsilon^a)^{A_{ij}\bar{A}_{ij}+(1-A_{ij})(1-\bar{A}_{ij})}, \tag{3.6}$$

$$P(\bar{F}_{ij}|F_{ij};\epsilon^f) = \left(\frac{\epsilon^f}{2}\right)^{\left(1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}\right)} \left(1-\epsilon^f\right)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}. \tag{3.7}$$

Rather than parameterizing all possible connectivity differences, we rely on scalars ϵ^a and ϵ^f to govern the probability of change within each modality. For binary random variables A_{ij} and $\bar{A}_{ij}$, this implies that the probability of change in anatomical connectivity does not depend on the value of A_{ij}. A similar property holds for the tri-state random variables F_{ij} and $\bar{F}_{ij}$. Moreover, Eq. (3.7) assumes that functional connectivity switches to its other two states with equal probability. Empirically, our results are more robust using Eqs. (3.6-3.7) than if we infer all transition probabilities for each modality.

3.2 Inference

We employ the maximum likelihood (ML) framework to estimate the model parameters:

$$\hat{\Theta}^* = \underset{\Theta}{\operatorname{argmax}} \sum_{\langle i,j\rangle} \log P(\mathbf{D}_{ij}, \mathbf{B}_{ij}; \Theta) = \underset{\Theta}{\operatorname{argmax}} \sum_{\langle i,j\rangle} \log \sum_{A_{ij},F_{ij}} P(A_{ij}, F_{ij}, \mathbf{D}_{ij}, \mathbf{B}_{ij}; \Theta), \tag{3.8}$$

where Θ is the set of model parameters. $\Theta = \{\pi, \mu, \sigma^2, \rho, \chi, \xi^2\}$ for the single-population model; $\Theta = \{\pi, \mu, \sigma^2, \rho, \chi, \xi^2, \epsilon\}$ for the model of population differences.

We derive the Expectation-Maximization (EM) algorithm [23] for fitting the models. The EM algorithm constructs the joint posterior of all hidden variables, which is then used to infer population differences. The posterior distribution can be computed directly from the observed data and the model parameters $\{\pi, \mu, \sigma^2, \rho, \chi, \xi^2, \epsilon\}$. Since these parameters are shared across connections/subjects, our model is based on a small set of unknown values.

3.2.1 Single Population Model

We use $\mathbf{X}_{ij} = \{A_{ij}, F_{ij}\}$ and $\mathbf{Y}_{ij} = \{\mathbf{D}_{ij}, \mathbf{B}_{ij}\}$ to denote the hidden and observed variables, respectively, associated with the connection between region i and region j.

Since A_{ij} is a binary random variable and F_{ij} is a tri-state random variable, the latent vector $\mathbf{X}_{ij}$ assumes one of six distinct values. The EM algorithm iterates between estimating the posterior probability of the hidden variables $\mathbf{X}_{ij}$ and estimating the model parameters Θ. Due to the independence of pairwise connections, this problem reduces to a standard mixture model with six components.

For notational simplicity, we index the 6 states of $\mathbf{X}$ using a set $\{1, \ldots, 6\}$. We construct the associated prior distribution $P(\mathbf{X}_{ij} = h; \Theta)$ and data likelihood $P(\mathbf{Y}_{ij}|\mathbf{X}_{ij} = h; \Theta)$ by evaluating Eq. (3.1-3.4). For example, if the index h denotes the latent assignment $A_{ij} = t, F_{ij} = k$,

$$P(\mathbf{X}_{ij} = h; \Theta) = (\pi^a)^t (1 - \pi^a)^{(1-t)} \cdot \pi^f_k, \tag{3.9}$$

$$P(\mathbf{Y}_{ij}|\mathbf{X}_{ij} = h; \Theta) = \prod_{l=1}^{L} \left[\rho_t \delta(D^l_{ij}) + (1 - \rho_t)\mathcal{N}(D^l_{ij}; \chi_t, \xi^2_t)\right] \cdot \mathcal{N}(B^l_{ij}; \mu_{tk}, \sigma^2_{tk}). \tag{3.10}$$

Independence across pairwise connections gives rise to a simple sum in the log-likelihood of the observed and hidden variables:

$$\mathcal{L}(\mathbf{X}, \mathbf{Y}; \Theta) = \sum_{\langle i,j \rangle} \log P(\mathbf{X}_{ij}; \Theta) P(\mathbf{Y}_{ij}|\mathbf{X}_{ij}; \Theta).$$

E-Step: We fix the model parameter estimates $\hat{\Theta}$ and update the posterior probability estimates $\hat{p}_{ijh}$ of the latent variables:

$$\hat{p}_{ijh} = P(\mathbf{X}_{ij} = h|\mathbf{Y}_{ij}; \hat{\Theta}) \propto P(\mathbf{X}_{ij} = h; \hat{\Theta}) P(\mathbf{Y}_{ij}|\mathbf{X}_{ij} = h; \hat{\Theta}) \quad \text{s.t.} \quad \sum_{h=1}^{6} \hat{p}_{ijh} = 1. \tag{3.11}$$

The M-Step parameter updates depend on the marginal probabilities of each latent variable. For convenience, we define them below:

$$\hat{a}_{ij} = P(A_{ij} = 1|Y_{ij}; \hat{\Theta}) = \sum_{h:A_{ij}=1} \hat{p}_{ijh} \tag{3.12}$$

$$\hat{s}^{(tk)}_{ij} = P(A_{ij} = t, F_{ij} = k|Y_{ij}; \hat{\Theta}) = \sum_{h:A_{ij}=t,F_{ij}=k} \hat{p}_{ijh}. \tag{3.13}$$

M-Step: We fix the posterior probability estimates $\hat{p}_{ijh}$ and update the model parameter estimates $\hat{\Theta}$. Given a guess of the parameters $\hat{\Theta}$ from the previous iteration, we construct a lower bound to the log-likelihood $\Psi(\Theta, \hat{\Theta}) = E_{\mathbf{X}|\mathbf{Y}}\left[\log P(\mathbf{X}, \mathbf{Y}; \Theta)|\mathbf{Y}, \hat{\Theta}\right]$. With some algebraic manipulation, we obtain

$$\begin{aligned}\Psi(\Theta, \hat{\Theta}) &= \sum_{h=1}^{6} \sum_{\langle i,j \rangle} P(\mathbf{X}_{ij} = h|\mathbf{Y}_{ij}; \hat{\Theta}) \log \left[P(\mathbf{X}_{ij} = h; \Theta) P(\mathbf{Y}_{ij}|\mathbf{X}_{ij} = h; \Theta)\right] \\ &= \sum_{h=1}^{6} \sum_{\langle i,j \rangle} \hat{p}_{ijh} \left[\log P(\mathbf{X}_{ij} = h; \Theta) + \log P(\mathbf{Y}_{ij}|\mathbf{X}_{ij} = h; \Theta)\right]. \end{aligned} \tag{3.14}$$

The parameter updates are obtained by differentiating Eq. (3.14) with respect to Θ and setting the gradient equal to zero. For notational convenience, we let C represent the number of (distinct) pairwise connections, and we let L_{ij}^0 be the number of subjects for which $D_{ij}^l = 0$ (i.e., no tract was detected). The binomial and multinomial priors reduce to intuitive sums of the latent posterior probability estimates:

$$\hat{\pi}^a = \frac{1}{C} \sum_{\langle i,j \rangle} \hat{a}_{ij}, \qquad\qquad \hat{\pi}_k^f = \frac{1}{C} \sum_{\langle i,j \rangle} \sum_{h:F_{ijk}=1} \hat{p}_{ijh}. \tag{3.15}$$

The probability ρ is the empirical likelihood of not finding a white matter tract between two regions:

$$\hat{\rho}_1 = \frac{\sum_{\langle i,j \rangle} \hat{a}_{ij} L_{ij}^0}{\sum_{\langle i,j \rangle} L \cdot \hat{a}_{ij}}. \tag{3.16}$$

The Gaussian likelihood parameters for the DWI measurements D are given by the weighted empirical mean and empirical variance over all *non-zero* values:

$$\hat{\chi}_1 = \frac{\sum_{\langle i,j \rangle} \hat{a}_{ij} \sum_{l:D_{ij}^l > 0} D_{ij}^l}{\sum_{\langle i,j \rangle} \hat{a}_{ij} (L - L_{ij}^0)}, \tag{3.17}$$

$$\hat{\xi}_1^2 = \frac{\sum_{\langle i,j \rangle} \hat{a}_{ij} \sum_{l:D_{ij}^l > 0} (D_{ij}^l - \chi_s)^2}{\sum_{\langle i,j \rangle} \hat{a}_{ij} (L - L_{ij}^0)}. \tag{3.18}$$

The updates for ρ_0, χ_0, ξ_0^2 (corresponding to latent connectivity $A_{ij} = 0$) are trivially obtained from the above expressions by substituting $(1 - \hat{a}_{ij})$ for $\hat{a}_{ij}$.

The likelihood parameters for the functional observations B are similarly constructed

as weighted statistics of the data:

$$\hat{\mu}_{tk} = \frac{\sum_{\langle i,j \rangle} \hat{s}_{ij}^{(tk)} \sum_{l=1}^{L} B_{ij}^{l}}{\sum_{\langle i,j \rangle} L \cdot \hat{s}_{ij}^{(tk)}}, \tag{3.19}$$

$$\hat{\sigma}_{tk}^{2} = \frac{\sum_{\langle i,j \rangle} \hat{s}_{ij}^{(tk)} \sum_{l=1}^{L} (B_{ij}^{l} - \hat{\mu}_{tk})^{2}}{\sum_{\langle i,j \rangle} L \cdot \hat{s}_{ij}^{(tk)}}. \tag{3.20}$$

3.2.2 Modeling Population Differences

The algorithms presented above can be easily extended to the two-population model in Figure 3-1(b). This complete model is the primary focus of our work in the following sections. Below, we let $\mathbf{X}_{ij} = \{A_{ij}, F_{ij}, \bar{A}_{ij}, \bar{F}_{ij}\}$ and $\mathbf{Y}_{ij} = \{D_{ij}^{l}, B_{ij}^{l}, \bar{D}_{ij}^{m}, \bar{B}_{ij}^{m}\}$ denote the hidden and observed variables, respectively, of the connection between regions i and j.

Both A_{ij} and $\bar{A}_{ij}$ are binary random variables and both F_{ij} and $\bar{F}_{ij}$ are tri-state random variables. Therefore, the latent vector $\mathbf{X}_{ij}$ assumes one of 36 distinct values. Once again, we index the latent states of $\mathbf{X}_{ij}$ using $h \in \{1, \ldots, 36\}$ and map the estimation problem to the standard mixture model with 36 components.

E-Step: We construct the full prior and likelihood distributions $P(\mathbf{X}_{ij} = h; \Theta)$ and data likelihood $P(\mathbf{Y}_{ij} | \mathbf{X}_{ij} = h; \Theta)$ using Eqs. (3.1-3.7). The posterior estimate $\hat{p}_{ijh}$ is computed analogously to Eq. (3.11) for each value of (i, j, h).

We define the following marginal posterior probabilities of the clinical templates:

$$\hat{r}_{ij} = P(\bar{A}_{ij} = 1 | Y_{ij}; \hat{\Theta}) = \sum_{h:\bar{A}_{ij}=1} \hat{p}_{ijh} \tag{3.21}$$

$$\hat{u}_{ij}^{(tk)} = P(\bar{A}_{ij} = t, \bar{F}_{ij} = k | Y_{ij}; \hat{\Theta}) = \sum_{h:\bar{A}_{ij}=t, \bar{F}_{ij}=k} \hat{p}_{ijh}. \tag{3.22}$$

M-Step: As in the preceding section, we let L_{ij}^{0} be the number of control subjects for whom $D_{ij}^{l} = 0$ and M_{ij}^{0} be the number of schizophrenia patients for whom $\bar{D}_{ij}^{m} = 0$.

Once again, the probability estimates are intuitive sums of the latent posteriors. In this case, we must also solve for the parameters $\epsilon^{a}, \epsilon^{f}$ in Eqs. (3.6-3.7):

$$\hat{\pi}^{a} = \frac{1}{C} \sum_{\langle i,j \rangle} \hat{a}_{ij}, \qquad \hat{\pi}_{k}^{f} = \frac{1}{C} \sum_{\langle i,j \rangle} \sum_{h:F_{ijk}=1} \hat{p}_{ijh}, \tag{3.23}$$

$$\hat{\epsilon}^{a} = \frac{1}{C} \sum_{\langle i,j \rangle} \sum_{h:A_{ij} \neq \bar{A}_{ij}} \hat{p}_{ijh}, \qquad \hat{\epsilon}^{f} = \frac{1}{C} \sum_{\langle i,j \rangle} \sum_{h:F_{ij} \neq \bar{F}_{ij}} \hat{p}_{ijh}. \tag{3.24}$$

Since both populations share the same data likelihood model, the updates for $\{\hat{\mu}, \hat{\sigma}^2, \hat{\rho}, \hat{\chi}, \hat{\xi}^2\}$ are derived from Eqs. (3.16-3.20) by incorporating one data term for each population:

$$\hat{\mu}_{tk} = \frac{\sum_{\langle i,j \rangle} \left[\hat{s}_{ij}^{(tk)} \sum_{l=1}^{L} B_{ij}^{l} + \hat{u}_{ij}^{(tk)} \sum_{m=1}^{M} \bar{B}_{ij}^{m} \right]}{\sum_{\langle i,j \rangle} \left[L \cdot \hat{s}_{ij}^{(tk)} + M \cdot \hat{u}_{ij}^{(tk)} \right]}, \tag{3.25}$$

$$\hat{\sigma}_{tk}^2 = \frac{\sum_{\langle i,j \rangle} \left[\hat{s}_{ij}^{(tk)} \sum_{l=1}^{L} (B_{ij}^{l} - \hat{\mu}_{tk})^2 + \hat{u}_{ij}^{(tk)} \sum_{m=1}^{M} (\bar{B}_{ij}^{m} - \hat{\mu}_{tk})^2 \right]}{\sum_{\langle i,j \rangle} \left[L \cdot \hat{s}_{ij}^{(tk)} + M \cdot \hat{u}_{ij}^{(tk)} \right]}, \tag{3.26}$$

$$\hat{\rho}_1 = \frac{\sum_{\langle i,j \rangle} \left[\hat{a}_{ij} L_{ij}^{0} + \hat{r}_{ij} M_{ij}^{0} \right]}{\sum_{\langle i,j \rangle} \left[L \cdot \hat{a}_{ij} + M \cdot \hat{r}_{ij} \right]}, \tag{3.27}$$

$$\hat{\chi}_1 = \frac{\sum_{\langle i,j \rangle} \left[\hat{a}_{ij} \sum_{l: D_{ij}^{l} > 0} D_{ij}^{l} + \hat{r}_{ij} \sum_{m: \bar{D}_{ij}^{m} > 0} \bar{D}_{ij}^{m} \right]}{\sum_{\langle i,j \rangle} \left[\hat{a}_{ij} (L - L_{ij}^{0}) + \hat{r}_{ij} (M - M_{ij}^{0}) \right]}, \tag{3.28}$$

$$\hat{\xi}_1^2 = \frac{\sum_{\langle i,j \rangle} \left[\hat{a}_{ij} \sum_{l: D_{ij}^{l} > 0} (D_{ij}^{l} - \hat{\chi}_1)^2 + \hat{r}_{ij} \sum_{m: \bar{D}_{ij}^{m} > 0} (\bar{D}_{ij}^{m} - \hat{\chi}_1)^2 \right]}{\sum_{\langle i,j \rangle} \left[\hat{a}_{ij} (L - L_{ij}^{0}) + \hat{r}_{ij} (M - M_{ij}^{0}) \right]}. \tag{3.29}$$

We have presented the EM algorithm for both models in Fig. 3-1. The posterior distributions over the latent variables play a crucial role in the clinical application of our model as follows.

3.2.3 Quantifying Group Differences

We assume that group differences are expressed in the latent templates $\{A_{ij}, F_{ij}, \bar{A}_{ij}, \bar{F}_{ij}\}$. Therefore, the main quantity of interest is the probability of change in the anatomical or functional connectivity templates for a given pairwise connection. We let $\hat{\epsilon}_{ij}^{a}$ denote the probability of a change in the anatomical connectivity between regions i and j, and we let $\hat{\epsilon}_{ij}^{f}$ denote the corresponding probability of change in functional connectivity. We estimate these values based on the inferred posterior probabilities $\{\hat{p}_{ijh}\}$:

$$\hat{\epsilon}_{ij}^{a} = \sum_{h: A_{ij} \neq \bar{A}_{ij}} \hat{p}_{ijh}\,, \qquad\qquad \hat{\epsilon}_{ij}^{f} = \sum_{h: F_{ij} \neq \bar{F}_{ij}} \hat{p}_{ijh}\,. \tag{3.30}$$

These values are the main output of our algorithm in the context of population studies.

3.3 Model Evaluation

This section details the methods used to validate our experimental results.

3.3.1 Model Significance

Although our model is based on standard Gaussian and multinomial probability distributions, the joint distribution is not Gaussian due to multiplicative interactions among latent variables and the effects of unknown non-random parameters. Therefore, we evaluate significance through non-parametric permutation tests. Specifically, we construct the distribution of the statistics in Eq. (3.30) under the null hypothesis by randomly permuting the subject labels (NC vs. SZ) $10,000$ times. For each permutation, we fit the model and compute the relevant statistics $\hat{\epsilon}^a_{ij}, \hat{\epsilon}^f_{ij}$. The significance (p-value) of each connection is equal to the proportion of permutations for which the computed statistic is greater than or equal to the value obtained under the true labeling.

3.3.2 Classification Accuracy

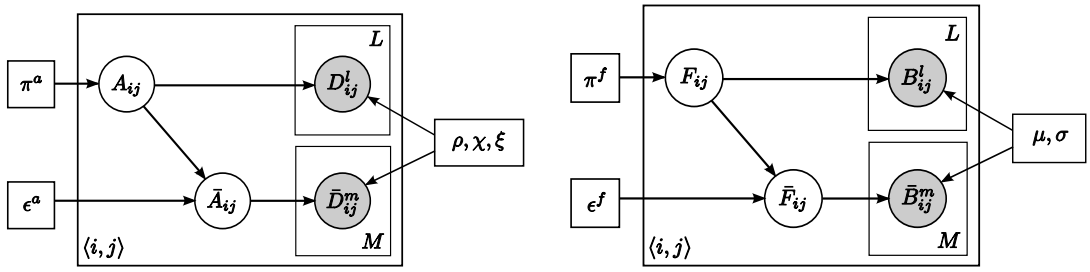
We also quantify the model's predictive power via ten-fold cross validation. We randomly divide the subjects into 10 groups, each with an equal number of controls and schizophrenia patients. [2] We fit the model using 9 training groups and employ a likelihood ratio test to predict the diagnoses (NC or SZ) of the held-out subjects. Our ratio test compares the likelihood of a new subject being generated from the control and the schizophrenic templates:

$$\frac{P(\{D_{ij}, B_{ij}\}|\{\hat{A}_{ij}, \hat{F}_{ij}\}, \hat{\Theta}^*)}{P(\{D_{ij}, B_{ij}\}|\{\hat{\bar{A}}_{ij}, \hat{\bar{F}}_{ij}\}, \hat{\Theta}^*)} \underset{'SZ'}{\overset{'NC'}{\gtrless}} 1, \tag{3.31}$$

where $\{\hat{A}_{ij}, \hat{F}_{ij}, \hat{\bar{A}}_{ij}, \hat{\bar{F}}_{ij}\}$ denotes the maximum *a posteriori* (MAP) estimate for the latent templates when fitted to the training data, and $\{D_{ij}, B_{ij}\}$ represents the observed DWI and fMRI data of a given test subject.

This process is repeated for each training-test set combination. In addition, we repeat the ten-fold cross validation 20 times using different groupings of subjects to evaluate the variability of the results. For comparison, we perform ten-fold cross validation using the support vector machine (SVM) classifiers trained on the fMRI correlations and DWI tractography measures individually, as well as on the combined dataset.

[2] Our clinical dataset consists of 19 patients and 19 controls, which we divide into nine groups of four subjects and one group of two subjects. These groups specify the testing sets in our classification experiments. Since we infer differences between the populations, it is important to maintain equal numbers of control and schizophrenia subjects in each group to avoid biasing the solution towards one population.



(a) DWI-only Model (b) fMRI-only Model

Figure 3-2: (a) DWI connectivity model for population differences. (b) fMRI connectivity model for population differences. A_{ij} represents the latent anatomical connectivity between regions i and j, and F_{ij} denotes the corresponding latent functional connectivity. D^l_{ij} and B^l_{ij} are the observed DWI and fMRI measurements, respectively, between regions i and j in the l^{th} subject. The schizophrenia templates are identified by an overbar. The variables, parameters and likelihood of (a) remains unchanged from the joint model. The likelihood in (b) is modified to reflect only the three functional connectivity states.

3.3.3 Baseline Methods

To evaluate the performance gain from combining fMRI and DWI data, we construct separate generative models for each modality, as depicted in Fig. 3-2. Here, we sever the connection between the anatomical connectivity templates $A, \bar{A}$ and the fMRI data.

Since the DWI data is independent of latent functional connectivity, all parameters, random variables and likelihoods remain unchanged for the DWI-only model (Fig. 3-2a).

The only modification in the fMRI-only model (Fig. 3-2b) involves the observed fMRI data. In particular, there are only three sets of likelihood parameters $\{\mu, \sigma^2\}$ corresponding to the three latent functional connectivity states. Formally, we replace the likelihood in Eq. (3.4) with

$$P(B^l_{ij}|F_{ij}; \{\mu, \sigma\}) = \prod_{k=-1}^{1} \mathcal{N}(B^l_{ij}; \mu_k, \sigma_k^2)^{F_{ijk}}, \quad k = -1, 0, 1 \tag{3.32}$$

for the control subjects and

$$P(\bar{B}^m_{ij}|\bar{F}_{ij}; \{\mu, \sigma\}) = \prod_{k=-1}^{1} \mathcal{N}(\bar{B}^m_{ij}; \mu_k, \sigma_k^2)^{\bar{F}_{ijk}}, \quad k = -1, 0, 1 \tag{3.33}$$

for the schizophrenia population.

We employ EM solutions, similar to those in Section 3.2. The empirical probabilities of change $\hat{\epsilon}^a_{ij}, \hat{\epsilon}^f_{ij}$ are computed according to Eq. (3.30) by replacing $\hat{p}_{ijh}$ with the posterior estimates of the appropriate model.

We perform permutation tests and cross validation using the DWI- and fMRI-specific models. These results allow us to evaluate the benefits of incorporating both imaging modalities in our analysis.

3.3.4 Implementation Details

In this section we describe the optimization choices in our implementation of the EM algorithm. We concentrate on the model of population differences.

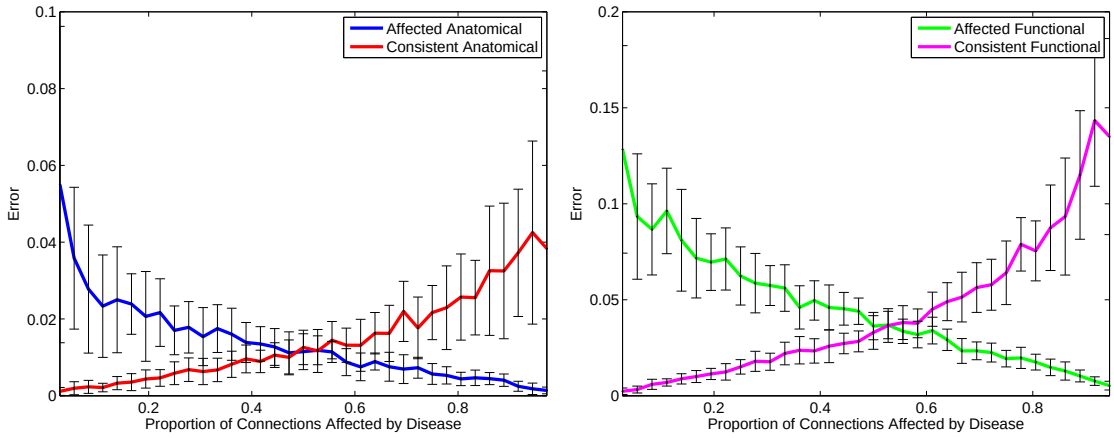
Initialization

Like many hill-climbing methods, the quality of our results depends on proper initialization. We initialize the model parameters $\Theta = \{\pi, \mu, \sigma^2, \rho, \chi, \xi^2, \epsilon\}$ based on empirical measures computed from the clinical data. In particular, we randomly sample the initial values of π^a, π^f_k, ϵ^a and ϵ^f from the interval $[0.3, 0.6]$. This scheme produces values near the center of the parameter space. We set σ^2_{tk} and ξ^2_t to the variances (across all connections and subjects) of the fMRI correlations and the non-zero DWI data, respectively. We set $\mu_{t0} = 0$, $\mu_{t1} = -\mu_{t,-1} = \sigma$, where σ^2 is the empirical variance. This captures our assumptions about the effect of latent functional connectivity on fMRI correlations and reflects the fact that the relationship between latent anatomical connectivity and fMRI data is less clear. Finally, we randomly sample χ_0 and χ_1 from the range of DWI FA values and generate $\rho_0 > \rho_1$.

It is sufficient to initialize the model parameters. The algorithm starts with computing the joint posterior probability distribution (E-Step). Empirically, we find that the results are stable with respect to different initializations of the model parameters. Therefore, we reinitialize the algorithm five times to sample the probability space, and subsequently select the maximum likelihood solution.

Convergence and Run Time

We ran the EM algorithm ten times using different subsets of subjects in each cross validation iteration. Convergence was based on the relative change in log-likelihood between consecutive iterations. On average, the algorithm converges in 87 iterations (E-step/M-step updates), and the average runtimes is 1.2 seconds per iteration. Thus, it requires on average 1.7 minutes to solve the model using EM. The iteration runtime scales linearly with the number of subjects. All simulations were performed using MATLAB on a single processor modern workstation.



(a) Anatomical (b) Functional

Figure 3-3: Proportion of mis-labeled connectivity relationships as a function of the proportion of latent connections affected by the disease. The bold lines represent the average error over 10 resamplings of the observed data, and the error bars represent one standard deviation. The likelihood parameterization is fixed according to the clinical dataset.

3.4 Experimental Results - Synthetic Data

We first evaluate the robustness and sensitivity of our algorithm using synthetic data. Throughout this section we fix the latent template and sample the observed data $\{B, \bar{B}, D, \bar{D}\}$, assuming 20 subjects in each population. We then infer the original latent templates from these noisy measurements. The error is computed as the proportion of connections for which the MAP connectivity estimates do not match the ground truth templates. We repeat the experiment ten times to collect error statistics.

In the first experiment we assume that the latent connectivity templates are similar for both populations. Specifically, the control templates have 180 pairwise connections for each of the 6 distinct values of latent connectivity templates $\{A_{ij}, F_{ij}\}$ ($N = 1080$, comparable to the clinical dataset), and we randomly alter a small percentage of connections to obtain the schizophrenia templates. This reflects our clinical hypothesis that the changes induced by schizophrenia occur in a small yet spatially-distributed subset of connections.

We fix the fMRI likelihood parameters according to the values estimated in our clinical dataset. This corresponds to $\sigma_{tk}^2 = 0.05$, $\mu_{0,-1} = -0.15$, $\mu_{1,-1} = -0.1$, $\mu_{00} = 0$, $\mu_{10} = 0.2$, $\mu_{01} = 0.3$, and $\mu_{11} = 0.5$. Likewise, we fix the DWI likelihood parameters at $\xi_t^2 = 0.005$, $\chi_0 = 0.45$, $\chi_1 = 0.35$, $\rho_0 = 0.6$ and $\rho_1 = 0.4$. The quantity $|\chi_1 - \chi_0| = 0.1$, which influences the separation between DWI distributions for present and absent latent anatomical connectivity, is equivalent to that of the clinical dataset. The values for ρ_0

and ρ_1 are much closer than what we estimate from real data. Otherwise, we find that anatomical connectivity is perfectly recovered, and we cannot probe the model's behavior.

Fig. 3-3 shows the errors in determining the latent templates both for the consistent connections and for the ones affected by the disorder. The bold lines in Fig. 3-3 represent the average error over ten independent samples of the entire data set. The error bars represent one standard deviation. Clearly, when the proportion of affected connections is small, the algorithm has slightly more difficulty identifying them. Similarly, if the proportion of affected connections is large, the algorithm has difficulty recovering the consistent connections. For example, when 10% of connections are affected, the model correctly identifies 90% of them. But if 90% of the connections are affected, the model recovers 99% of them. The maximum error is less than 15%. This suggests that our algorithm can accurately fit the model, which is promising for the application to clinical data.

In the second experiment, we explore the breakdown points of our model. We consider the case when the DWI likelihood distributions provide little information about latent anatomical connectivity as well as the case when the fMRI likelihood distributions are nearly uninformative about latent anatomical and/or functional connectivity. The Gaussian variances are fixed at $\sigma_{tk}^2 = \sigma^2 = 0.01$ and $\xi_t^2 = \xi^2 = 0.005$.

We parameterize the DWI model as follows:

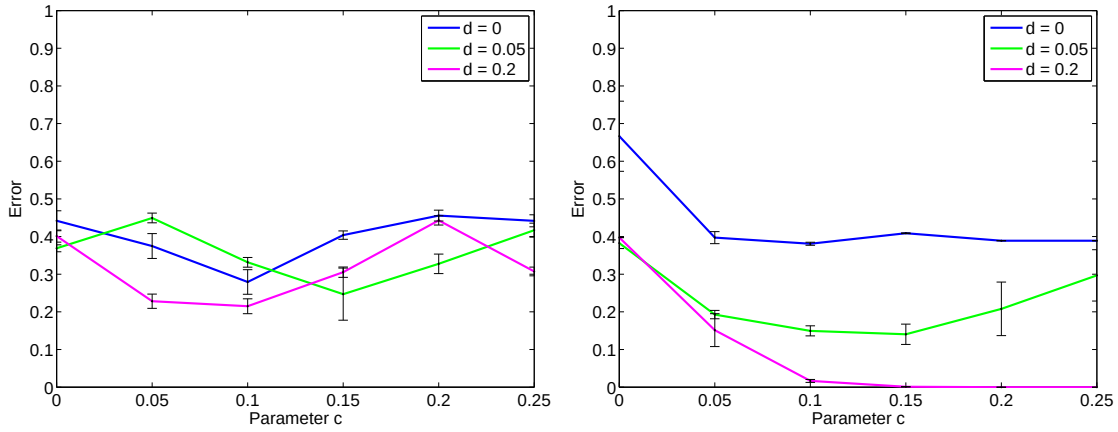
$$\begin{aligned} \chi_0 &= 0.5 - a & \chi_1 &= 0.5 + a \\ \rho_0 &= 0.5 + b & \rho_1 &= 0.5 - b \end{aligned} \tag{3.34}$$

where $a, b \geq 0$. The parameter a controls the difference in DWI distributions for the two underlying anatomical connectivity values. The parameter b determines to how much more likely one is to find a DWI tract between two regions given a direct anatomical connection than if no connection is present.

We parameterize the Gaussian means for the fMRI model as follows:

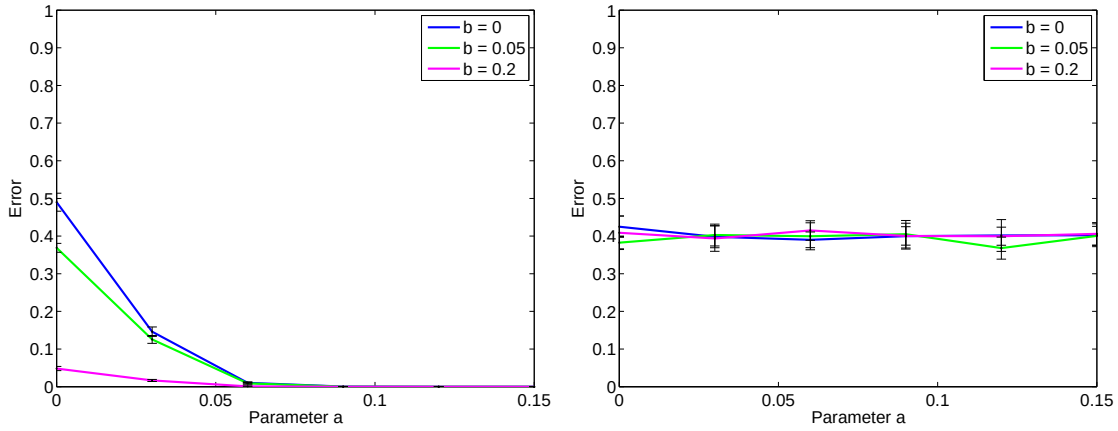
$$\begin{aligned} \mu_{00} &= \mu_{10} = 0 \\ \mu_{01} &= c = -\mu_{0,-1} \\ \mu_{10} &= d = -\mu_{1,-1} \end{aligned} \tag{3.35}$$

where $c, d \geq 0$. The parameters c and d control the functional separation in the absence and presence of a latent anatomical connection, respectively. The quantity $(d-c)$ relates to the effect anatomical connectivity has on the magnitude of fMRI correlations. This setup



(a) Anatomical (b) Functional

Figure 3-4: Proportion of mis-labeled connectivity relationships between the latent templates. The bold lines represent the average error over 10 resamplings of the observed data $\{B, \bar{B}, D, \bar{D}\}$. The error bars represent one standard deviation from the mean. The DWI likelihood parameterization is fixed at $a = 0$ and $b = 0.05$.



(a) Anatomical (b) Functional

Figure 3-5: Proportion of mis-labeled connectivity relationships between the latent templates. The error is averaged over 10 resamplings of the data. The fMRI likelihood parameterization is fixed at $c = 0$ and $d = 0.05$.

allows for adequate flexibility in manipulating the generative process while simultaneously reducing the number of free parameters to explore.

We assume a uniform distribution of latent connectivity values; the templates contain 30 pairwise connections for each of the 36 values of $\{A_{ij}, F_{ij}, \bar{A}_{ij}, \bar{F}_{ij}\}$ ($N = 1080$). We generate 20 subjects from each population and sweep the parameters in Eqs. (3.34-3.35). For each parameter set, we generate data from the two-population model in Fig. 3-1(b) and

solve for the latent connectivity. We repeat this procedure several times to ensure stability of the reported behavior.

We first specify the DWI parameterization such that the probability of finding a tract is slightly greater than 0.5 given a latent anatomical connection ($b = 0.05$), and such that there is no difference in DWI likelihood when a tract is observed ($a = 0$). Fig. 3-4 reports the errors in predicting latent connectivity changes between the populations. We observe that as c, d increase from zero, the algorithm uses the fMRI data and the slight difference in DWI likelihood to estimate latent functional connectivity. Another interesting observation is the predictable dip in error in Fig. 3-4(a) when $c, d \geq \sigma = 0.1$ and $|d - c| \geq \sigma$. In this case the fMRI likelihoods based on positive/negative latent functional connectivity are simultaneously far from zero and distinct given the presence or absence of a latent anatomical pathway. The algorithm uses the first separation (far from zero) to identify latent functional connectivity and the second (distinct based on anatomy) to infer latent anatomical connectivity.

Second, we fix the fMRI parameterization such that there is a slight separation between the mixture distributions given the presence of a latent anatomical connection ($c = 0, d = 0.05$). Fig. 3-5 reports the errors in predicting the connectivity changes. An informative DWI likelihood (higher values of a, b) allows us to correctly estimate the anatomical templates. However, it does not improve the estimates of latent functional connectivity. This is because our model does not include a direct link between the functional templates and the DWI data.

In summary, highly separable fMRI data allows us to estimate the functional templates and improves slightly our inference of latent anatomical connectivity. In contrast, highly separable DWI data produces accurate anatomical templates but does not improve the functional connectivity estimates. When both datasets are informative, the algorithm recovers all the latent templates and model parameters.

3.5 Experimental Results - Clinical Data

We demonstrate our model on the clinical study of schizophrenia (see Section 2.7). We compute the DWI connectivity D^l_{ij} between regions i and j in subject l by averaging FA along all fibers connecting the two regions. If no tracts are found, D^l_{ij} is set to zero. We extract the fMRI connectivity B^l_{ij} as the Pearson correlation coefficient between the mean time courses of regions i and j in subject l.

To inject prior clinical knowledge, we pre-select 8 brain structures (corresponding to 16 regions) that are believed to play a role in schizophrenia: the superior temporal gyrus,

rostral middle frontal gyrus, hippocampus, amygdala, posterior cingulate, rostral anterior cingulate, parahippocampal gyrus, and transverse temporal gyrus. We model only the 1096 $\left(16 \times 76 - \binom{16}{2}\right)$ pairwise connections between these ROIs and all other regions in the brain.

3.5.1 Empirical Study of Data Distributions

In this section we present aggregate properties of our data, which motivate our choice of likelihood parameterization in Section III. We group both populations together, as the differences induced by schizophrenia are subtle and do not affect the global distributions.

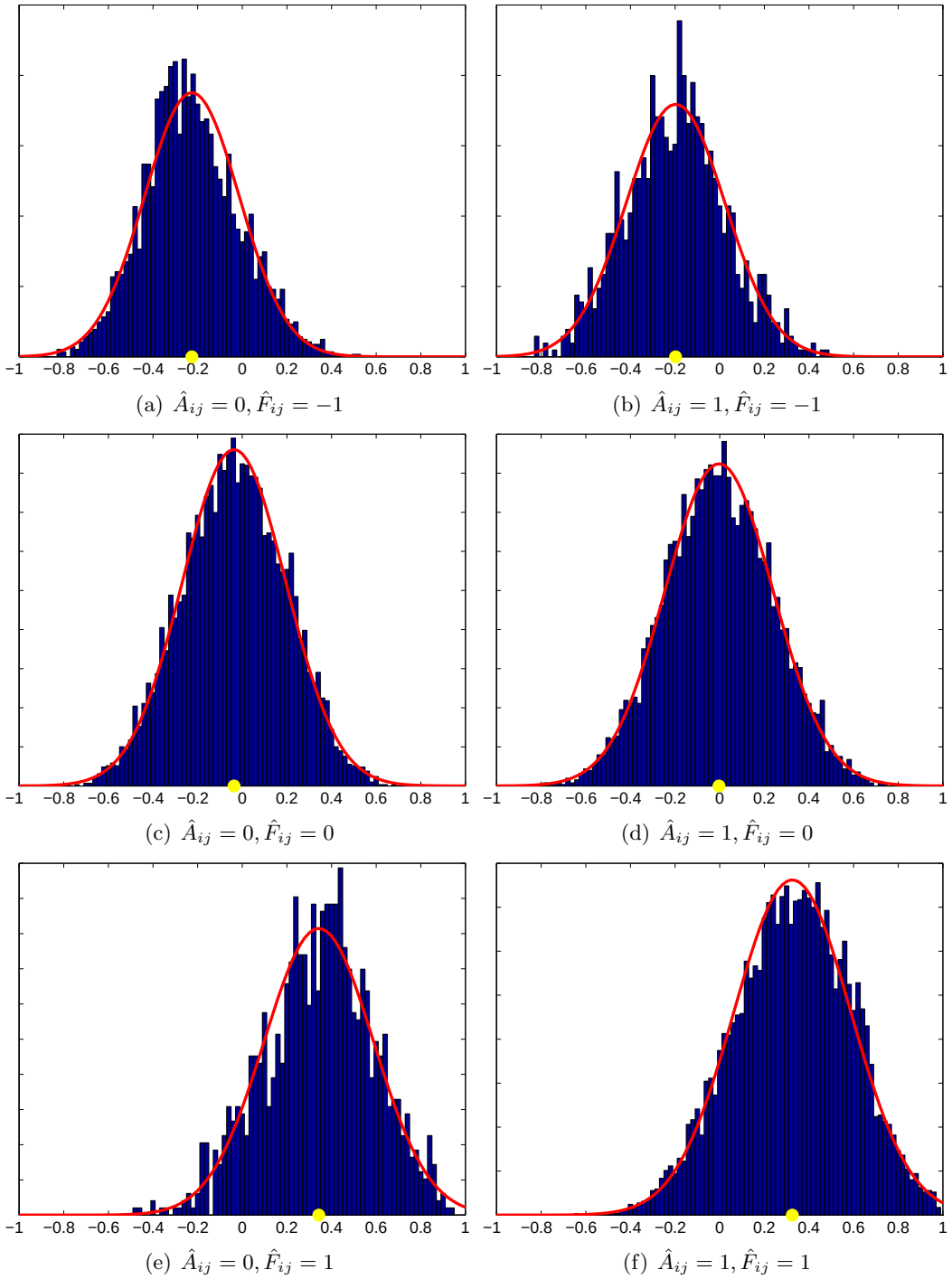
We first fit the distributions of fMRI correlations and DWI FA values to our likelihood model in Eqs. (3.3-3.4). Since we cannot access the latent connectivity A_{ij} and F_{ij}, we approximate these variables by working with average measures of the data across subjects. In particular, we threshold the proportion of subjects that exhibit white matter tracts between regions to estimate $\hat{A}_{ij}$. Similarly, we threshold the average fMRI correlations to estimate $\hat{F}_{ij}$. We then analyze the distribution of DWI FA values and fMRI correlations across all connections in all subjects.

Fig. 3-6 depicts the histograms of fMRI correlations for all combinations of roughly estimated latent connectivity. Fig. 3-7 illustrates histograms of the non-zero DWI values for the two types of anatomical connectivity. We have overlaid the fitted Gaussian distributions in each plot. The yellow dots correspond to empirical means. We observe that the variability in DWI and fMRI data across connections is across subjects are reasonably approximated using Gaussian distributions.

Fig. 3-7 also suggests that the average DWI measure is slightly higher for connections in which tractography identifies white matter tracts in only a few subjects ($\hat{A}_{ij} = 0$). We explore this phenomenon by considering the distribution of FA values along all fibers when (1) white matter tracts are detected in all subjects, and (2) white matter tracts are detected in only one subject. Our analysis considers the first scenario to represent a "true" anatomical connection and the second to be spurious fibers. Fig. 3-8 illustrates the histograms of two representative connections for each of the above cases.

Empirically, we observe that the distributions of FA values along spurious fibers is more uniformly distributed across a broad range of values ($FA \in [0.2, 0.8]$), whereas the distribution along true fibers is concentrated towards the lower end of this range ($FA \in [0.3, 0.5]$). The average FA for a false-positive connection is higher than the FA for a correctly-identified connection.

There are several factors which may contribute to this phenomenon. For example, since tractography is guided by the estimated tensors, perhaps the algorithm latches onto

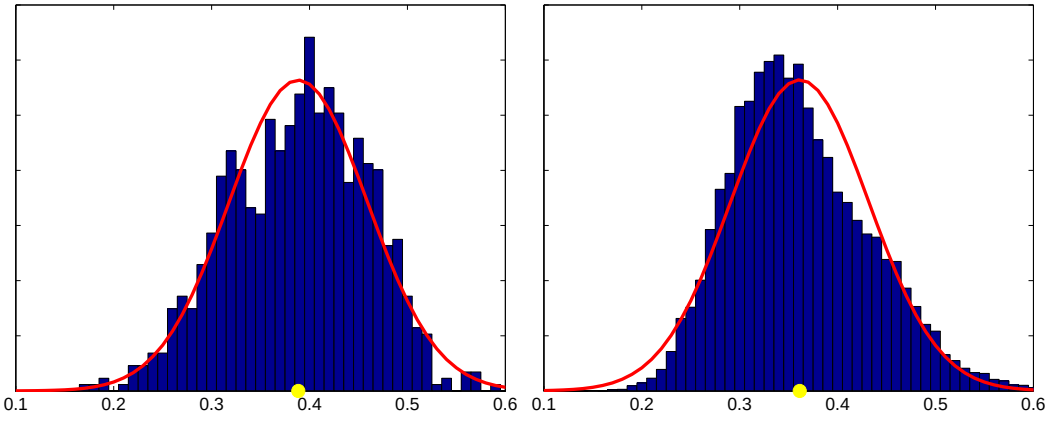


(a) $\hat{A}_{ij} = 0, \hat{F}_{ij} = -1$ (b) $\hat{A}_{ij} = 1, \hat{F}_{ij} = -1$

(c) $\hat{A}_{ij} = 0, \hat{F}_{ij} = 0$ (d) $\hat{A}_{ij} = 1, \hat{F}_{ij} = 0$

(e) $\hat{A}_{ij} = 0, \hat{F}_{ij} = 1$ (f) $\hat{A}_{ij} = 1, \hat{F}_{ij} = 1$

Figure 3-6: Histograms of fMRI correlations based on estimated connectivity. Gaussian distributions that have been fitted to the data are overlaid in red. The yellow dots correspond to empirical means.



(a) $\hat{A}_{ij} = 0$ (b) $\hat{A}_{ij} = 1$

Figure 3-7: Histograms of non-zero DWI data based on estimated anatomical connectivity. Gaussian distributions that have been fitted to the data are overlaid in red. The yellow dots correspond to empirical means.

artificially high anisotropy in the DWI images to produce these erroneous tracts. Our two-tensor tractography algorithm [71] may also play a role. We fit one tensor along the main fiber bundle and use a second tensor to account for residual anisotropy. Our tractography algorithm computes only FA along the main fiber, which may impact the overall distribution.

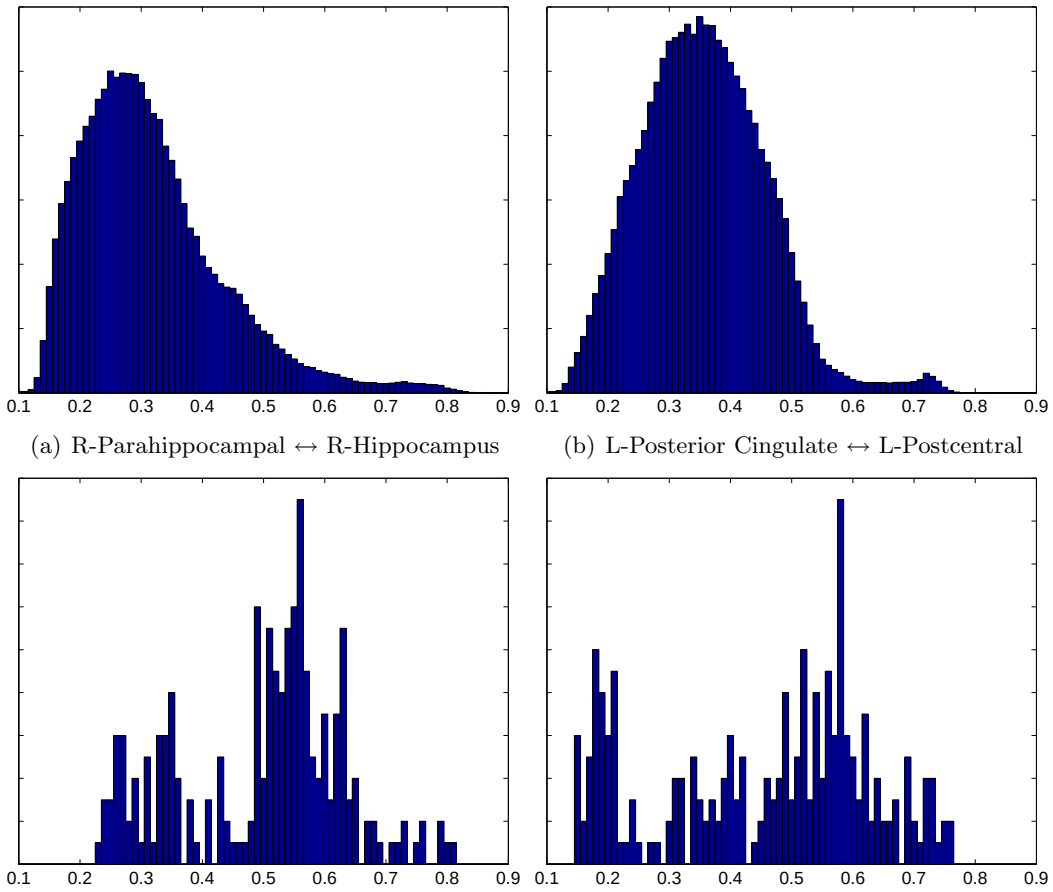
3.5.2 Joint Connectivity Model for the Clinical Data

We first fit the joint model in Fig. 3-1(a) to each population separately, as well as to the entire dataset. Table 3.2 reports the parameters of the three cases. We observe that the ML solutions $\{\pi.\mu, \sigma, \rho, \chi, \xi\}$ are largely consistent between the groups and for the combined case. This suggests that population differences have a larger influence on the latent connectivities rather than on the data likelihood parameters.

Table 3.2 highlights some interesting properties of the data. For example, $\mu_{0k} < \mu_{1k}$ for all k, which indicates that the presence of an anatomical connection between two regions increases the mean functional correlation. This result is consistent with prior work [48, 61]. Additionally, $\chi_0 > \chi_1$ implies that false-positive white matter tracts have *higher* mean FA values than correctly-identified white matter tracts. This is consistent with our empirical evaluation of the data in Section 3.5.1.

3.5.3 Population Study

Fig. 3-9 depicts the significantly different ($\hat{\epsilon}_{ij} > 0.75$) anatomical and functional connections identified by the algorithm. In this case, we identify connections with an uncorrected p-value



(a) R-Parahippocampal ↔ R-Hippocampus (b) L-Posterior Cingulate ↔ L-Postcentral

(c) R-Cuneus ↔ L-Transverse Temporal Gyrus (d) R-Rostral Ant Cingulate ↔ R-Inf Parietal

Figure 3-8: Histograms of FA values along fibers for representative connections detected in all subjects (a-b) and representative connections detected in a single subject (c-d).

of 0.05. Our results exhibit patterns that have previously been reported in the schizophrenia literature and are linked to clinical hypotheses regarding the disorder. For example, we observe that schizophrenia patients exhibit increased functional connectivity between the parietal/posterior cingulate region and the frontal lobe and reduced functional connectivity between the parietal/posterior cingulate region and the temporal lobe in Fig. 3-9. These results confirm the findings of functional abnormalities involving the default network and of widespread functional connectivity changes in schizophrenia [37, 66]. Likewise, the differences in anatomical connectivity are distributed across the brain.

Table 3.3 and Table 3.4 report the corresponding region pairs and significance values, where we have highlighted the connections that overlap with our discriminative analysis. Specifically, we perform robust feature selection using the Gini Importance (GI) score

Table 3.2: Parameters of the joint model in Fig. 3-1(a). The analysis is performed separately for the control (NC) and the schizophrenic (SZ) populations, as well as for the entire dataset (NC+SZ).

	π_A	$\pi_{F,-1}$	π_{F0}	π_{F1}	ρ_0	ρ_1	χ_0	χ_1	ξ_0^2	ξ_1^2
NC	0.37	0.49	0.40	0.11	0.65	0.10	0.42	0.34	0.005	0.003
SZ	0.37	0.42	0.44	0.14	0.67	0.11	0.41	0.34	0.005	0.003
NC+SZ	0.37	0.43	0.43	0.14	0.66	0.11	0.41	0.34	0.005	0.003

	$\mu_{0,-1}$	$\mu_{1,-1}$	μ_{00}	μ_{10}	μ_{01}	μ_{11}
NC	-0.13	-0.054	0.059	0.23	0.35	0.55
SZ	-0.19	-0.087	0.007	0.21	0.30	0.55
NC+SZ	-0.17	-0.071	0.015	0.21	0.29	0.55

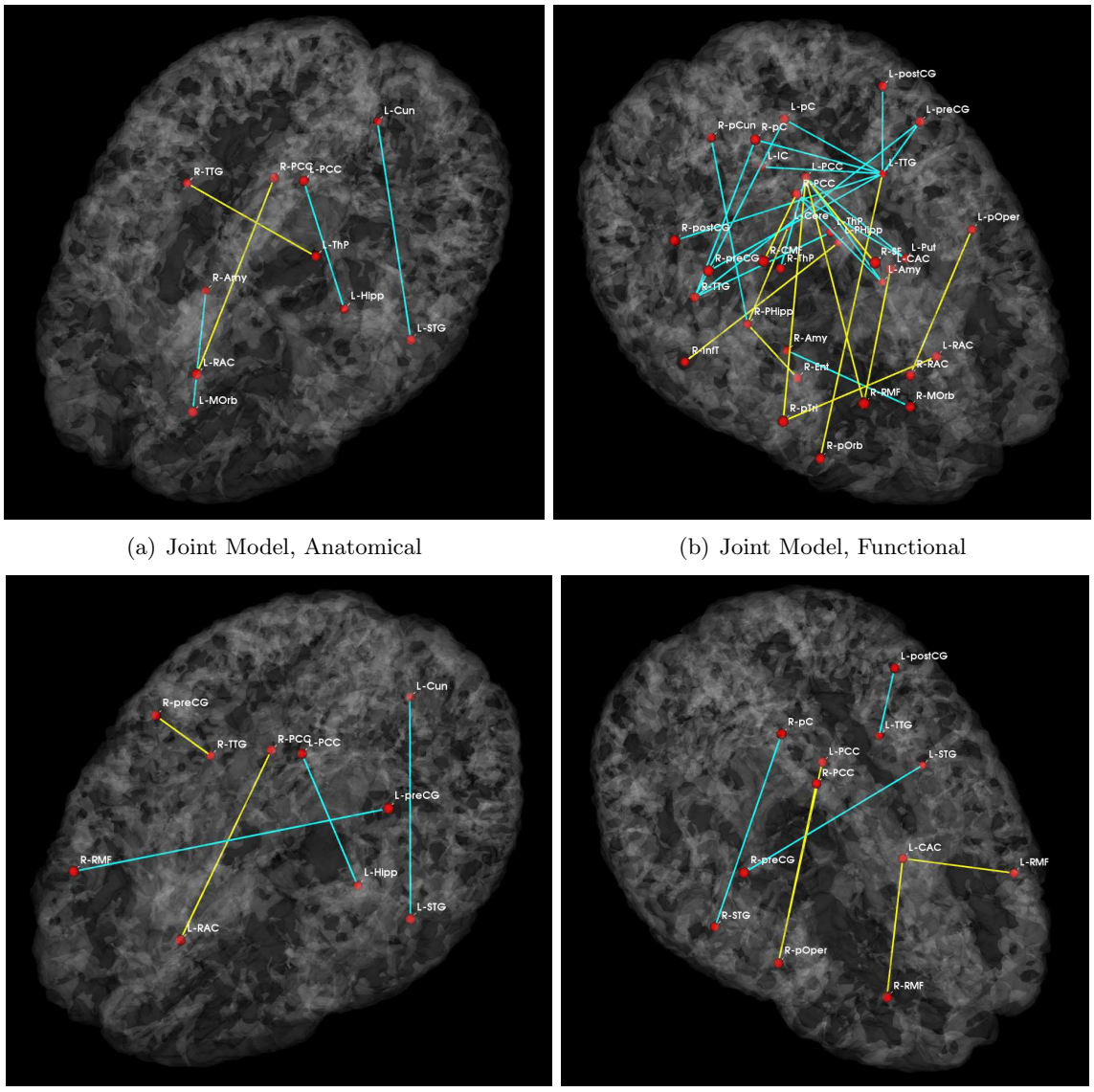
	$\sigma_{0,-1}^2$	$\sigma_{1,-1}^2$	σ_{00}^2	σ_{10}^2	σ_{01}^2	σ_{11}^2
NC	0.053	0.055	0.049	0.055	0.055	0.044
SZ	0.048	0.058	0.053	0.056	0.061	0.043
NC+SZ	0.050	0.057	0.053	0.056	0.061	0.044

derived from the Random Forest algorithm [11]. Details of this work are provided in Appendix A. We discuss the relevance of this comparison in the following section.

Fig. 3-10 shows representative DWI fibers for the significant anatomical connections identified by the joint model. In each case, we display the corresponding tracts within a single subject from the population with *higher* connectivity. We note that the results of the joint model do not completely agree with those of the single-modality models.

Finally, we observe consistency in parameter estimates across random subject re-labelings in the permutation procedure (not shown). This suggests that the main effects of the label permutations are reflected in the latent connectivity rather than in the data likelihood.

Fig. 3-11 reports classification accuracy for the joint generative model, for the individual generative models, and for the SVM classifiers. Training accuracy is presented as validation that the model does learn discriminative features. We acknowledge the low classification accuracy in Fig. 3-11 but emphasize that our model is not formulated for classification. In contrast, we aim to understand the interaction between fMRI and DWI data using a set of assumptions about connectivity and schizophrenia. Differences between the two populations are modeled through shifts in the likelihood parameters rather than by changes in specific fMRI and DWI values. Therefore, we do not expect our approach to achieve the classification performance of algorithms specifically tailored for classification. Rather, Fig. 3-11 illustrates three main points. First, the joint model achieves above-chance generalization accuracy. This suggests that the underlying connectivity might play a role in schizophrenia. Second,



(a) Joint Model, Anatomical

(b) Joint Model, Functional

(c) DWI-only Model

(d) fMRI-only Model

Figure 3-9: Significant anatomical and functional connectivity differences ($p < 0.05$ and $\hat{\epsilon}^a_{ij}, \hat{\epsilon}^f_{ij} > 0.75$). Blue lines indicate higher connectivity in the control group; yellow lines indicate higher connectivity in the schizophrenia population. (a-b) are derived from the joint DWI/fMRI model. (c) depicts significant anatomical connections from the DWI-only model, and (d) illustrates significant functional connections from the fMRI-only model.

Table 3.3: Significant anatomical and functional connections based on the joint generative model in Fig. 3-1(b). The blue connections are identified by our discriminative feature selection method presented in Appendix A.

Anatomical Connections			
Region 1	Region 2	p	$\hat{\epsilon}_A^n$
L Posterior Cingulate (L-PCC)	L Hippocampus (L-Hipp)	0.0001	0.93
R Transverse Temporal (R-TTG)	L Thalamus-Proper (L-ThP)	0.012	0.99
L Superior Temporal (L-STG)	L Cuneus (L-Cun)	0.016	0.88
L Medial Orbitofrontal (L-MOrb)	R Amygdala (R-Amy)	0.032	0.99
R Posterior Cingulate (R-PCC)	L Rostral Ant. Cingulate (L-RAC)	0.039	0.96

Functional Connections			
Region 1	Region 2	p	$\hat{\epsilon}_F^n$
R Pars Triangularis (R-pTri)	L Posterior Cingulate (L-PCC)	0.0003	0.93
R Superior Frontal (R-SF)	L Posterior Cingulate (L-PCC)	0.0005	0.94
R Parahippocampal (R-PHipp)	R Enthorinal Cortex (R-Ent)	0.001	0.98
R Postcentral (R-postCG)	L Transverse Temporal (L-TTG)	0.0011	0.91
L Transverse Temporal (L-TTG)	L Precentral (L-preCG)	0.0013	0.95
R Posterior Cingulate (R-PCC)	L Amygdala (L-Amy)	0.0015	0.97
R Inferior Temporal (R-InfT)	L Parahippocampal (L-PHipp)	0.0019	0.95
R Rostral Mid. Frontal (L-RMF)	L Caudal Ant. Cingulate (L-CAC)	0.0025	0.90
R Parahippocampal (R-PHipp)	L Cerebellum (L-Cere)	0.0025	0.95
L Posterior Cingulate (L-PCC)	L Amygdala (L-Amy)	0.0027	0.93
L Transverse Temporal (L-TTG)	L Paracentral (L-pC)	0.0028	0.99
R Pars Triangularis (R-pTri)	L Rostral Ant. Cingulate (L-RAC)	0.003	0.92
R Precentral (R-preCG)	L Transverse Temporal (L-TTG)	0.0032	0.99
R Paracentral (R-pC)	L Transverse Temporal (L-TTG)	0.0042	0.99
R Caudal Mid. Frontal (R-CMF)	L Posterior Cingulate (L-PCC)	0.0056	0.96
R Rostral Ant. Cingulate (R-RAC)	L Pars Opercularis (L-pOper)	0.013	0.92
R Transverse Temporal (R-TTG)	L Precentral (L-preCG)	0.017	0.98
R Transverse Temporal (R-TTG)	L Paracentral (L-pC)	0.019	0.98
R Pars Orbitalis (R-pOrb)	L Transverse Temporal (L-TTG)	0.021	0.75
R Transverse Temporal (R-TTG)	R Paracentral (R-pC)	0.024	0.90
R Medial Orbitalis (R-MOrb)	R Amygdala (R-Amy)	0.029	0.77
R Posterior Cingulate (R-PCC)	R Caudal Mid Frontal (R-CMF)	0.032	0.98
L Posterior Cingulate (L-PCC)	L Putamen (L-Put)	0.033	0.96
L Transverse Temporal (L-TTG)	L Postcentral (L-postCG)	0.037	0.79
R Transverse Temporal (R-TTG)	L Thalamus-Proper (L-ThP)	0.038	0.80
L Posterior Cingulate (L-PCC)	R Thalamus-Proper (R-ThP)	0.038	0.78
R Precuneus (R-pCun)	R Parahippocampal (R-PHipp)	0.042	0.95
R Posterior Cingulate (L-PCC)	L Putamen (L-Put)	0.042	0.83
L Transverse Temporal (L-TTG)	L Isthmus Cingulate (L-IC)	0.047	0.83

Table 3.4: Significant anatomical (top) and functional (bottom) connections based on the single-modality generative models in Fig. 3-2. The green connection is identified by our discriminative feature selection method presented in Appendix A.

Anatomical Connections			
Region 1	Region 2	p-value	$\hat{\epsilon}_n^a$
L Posterior Cingulate (L-PCC)	L Hippocampus (L-Hipp)	0.011	0.75
R Posterior Cingulate (R-PCC)	L Rostral Ant. Cingulate (L-RAC)	0.029	0.99
L Superior Temporal (L-STG)	L Cuneus (L-Cun)	0.029	0.79
R Rostral Mid Frontal (R-RMF)	L Precentral (L-preCG)	0.033	0.93

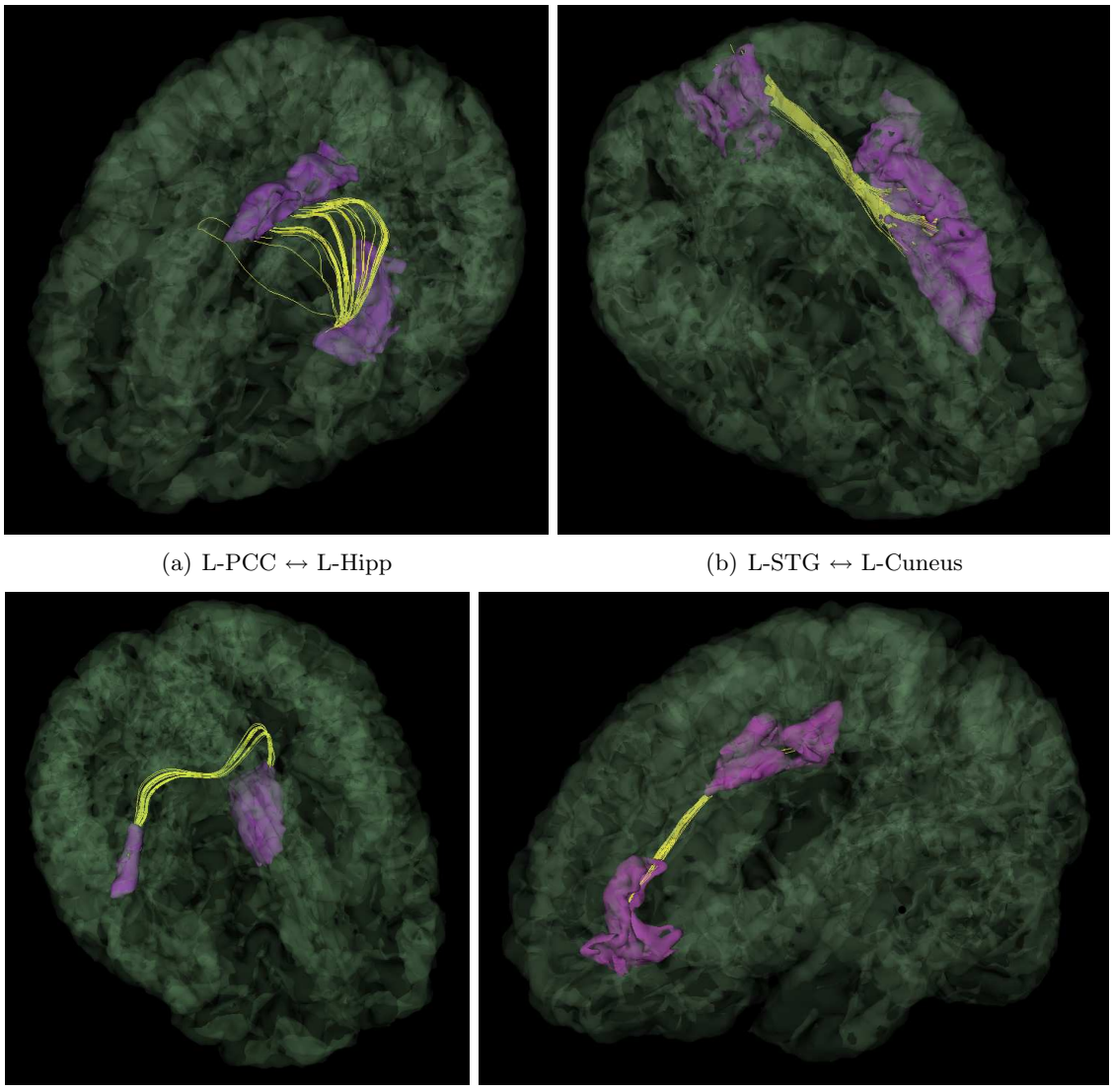
Functional Connections			
Region 1	Region 2	p-value	$\hat{\epsilon}_n^f$
R Posterior Cingulate (R-PCC)	R Pars Opercularis (R-pOper)	0.0011	0.78
R Rostral Mid. Frontal (R-RMF)	L Caudal Ant. Cingulate (L-CAC)	0.0028	0.95
R Pars Opercularis (R-pOper)	L Posterior Cingulate (L-PCC)	0.0033	0.99
L Transverse Temporal (L-TTG)	L Postcentral (R-postCG)	0.0047	0.97
R Precentral (R-preCG)	L Superior Temporal (L-STG)	0.020	0.84
L Rostral Mid. Frontal (L-RMF)	L Caudal Ant. Cingulate (L-CAC)	0.043	0.98
R Superior Temporal (R-STG)	R Paracentral (R-pC)	0.048	0.94

modeling anatomical and functional connectivity jointly yields predictive advantages over treating the fMRI and DWI data separately. Finally, even the SVM accuracy is low. In addition, Appendix A presents an auxiliary discriminative analysis of our functional data. We use state-of-the-art feature selection and classification algorithms to learn the relevant connections and group subjects. Despite our sophisticated approach, the generalization accuracy is only 75%. This underscores the well-documented challenge of finding robust functional and anatomical changes induced by schizophrenia [22, 63].

We note that much of the prior work on classification in schizophrenia did not rely on the modalities used in this paper. In particular, most reported classifiers consider volumetric changes found in T1 MRI [28] or activation patterns from task-based fMRI [31]. A few studies have focused on resting-state fMRI [52] or DTI tractography [75]. In all cases specialized features and classifiers were fine-tuned in order to obtain high accuracy.

3.6 Discussion

We formulate a generative model to infer changes in functional and anatomical connectivity induced by schizophrenia using both resting-state fMRI correlations and DWI tractography, and we present an algorithm for maximum likelihood estimation of the model parameters. We simultaneously obtain the joint posterior probability distribution of all the hidden vari-



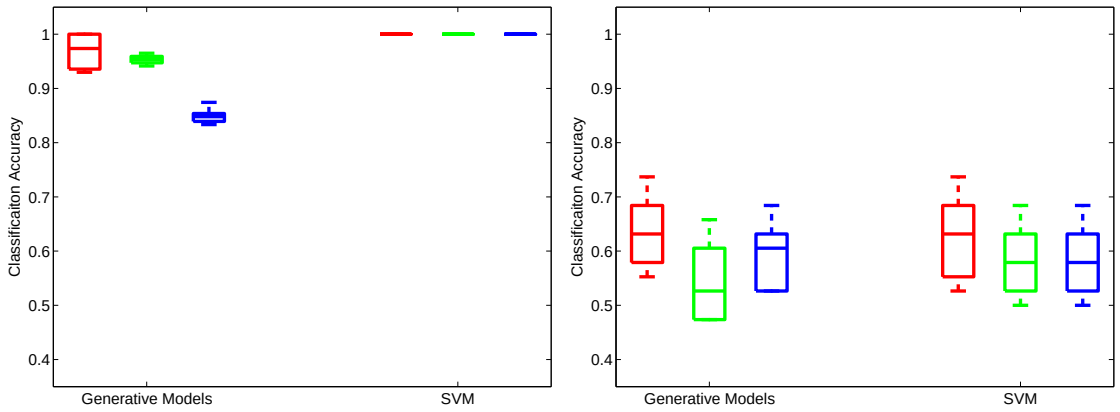
(a) L-PCC ↔ L-Hipp

(b) L-STG ↔ L-Cuneus

(c) R-TTG ↔ L-Thalamus Proper

(d) R-PCC ↔ R-RAC

Figure 3-10: Representative DWI fibers for each of the significant anatomical connections identified by the joint model in Table 3.3. The corresponding ROIs are displayed in pink; the fibers are depicted in yellow.



(a) Training Accuracy (b) Testing Accuracy

Figure 3-11: Training and testing accuracy of ten-fold cross validation using the joint generative model, the individual fMRI and DWI models and a linear SVM classifier. Red results are obtained using both modalities; green results are based only the DWI data; blue results are acquired from the fMRI data. The box denotes the upper and lower quartiles, the line indicates the median values, and the whiskers correspond to the 10^{th} and 90^{th} percentiles.

ables, which allows us to identify population differences.

One interesting observation is the symmetry of functional connectivity differences across the hemispheres in Fig. 3-9(b,d). In particular, if a given functional connection shows significant differences between the populations, then functional connections involving those same regions in the opposite hemisphere tend to also be significant. This may arise from the well-documented symmetry found in resting-state fMRI correlations [90].

In contrast to functional connectivity, the model identifies few significant anatomical connections, only two of which are consistent between the algorithms. Moreover, the inter-hemispheric connections in Fig. 3-9(a) do not correspond to direct neural pathways within the brain. Rather, these connections arise from artifacts in the DWI images as well as from the behavior of our two-tensor tractography algorithm [71]. In particular, our algorithm recovers a much richer set of white matter fibers relative to single-tensor methods. However, this set includes a larger number of false-positive tracts.

The results may also be influenced by our selection of regions. If the regions are too small, the variability in DWI tractography across subjects makes it difficult to infer the template anatomical connectivity and group-level parameters. However, larger regions smooth out important functional connectivity information. In this work, we rely on regions identified by Freesurfer. These estimates of Brodmann areas provide anatomically meaningful correspondences across subjects that roughly correspond to functional divisions within the

brain. Presently, we select the correlation between mean time courses as a measure of functional connectivity. However, other statistics can be incorporated as well (for example, the mode and/or variance of the distribution of voxel-wise correlations as well as fitting to parameterized distributions). Finally, we emphasize that our framework applies readily to any set of ROIs that are defined consistently across subjects.

Despite the limited differences in anatomical connectivity, one justification for including the DWI data is the improved classification. We observe that combining fMRI and DWI data achieves better generalization accuracy than that of similar models built from one of these modalities. Additionally, most significant functional connections obtained through the individual fMRI model are not consistent with those obtained via the joint models.

A second justification for including anatomy is the overlap between the significant functional connections identified by our generative models and the GI-based connections presented in Appendix A. We observe that seven of the significant joint model connections in Table 3.3 are among the most robust, predictive functional connections reported in Table A.3. In contrast, there is only one consistent connections between the functional model (Table 3.4) and our feature selection method.

Our experience with the algorithms suggests that that the joint model focuses on the presence or absence of a white matter tract between two regions (rather than differences in FA) to determine latent anatomical connectivity. In particular, if several subjects exhibit a connection, then A_n is likely to be one; otherwise, it is likely to be zero. This is supported by results in synthetic data. Given a large difference in the probabilities of not finding a tract (e.g., $\rho_0 \approx 0.65$ and $\rho_1 \approx 0.1$, as estimated from the data), our algorithm correctly distinguishes latent anatomical connectivity, regardless of FA values. Once the anatomical connectivity pattern has been determined, the algorithm partitions the functional correlations into two groups. The mean functional correlation increases when there is a latent anatomical connection, which is reflected in the parameter estimates. The algorithms can reassign "borderline" connections based on the parameter/posterior estimates. We believe that this partition of fMRI correlations based on anatomical connectivity stabilizes the estimates of latent functional connectivity. This, in turn, allows the joint model to better explain differences between two populations.

The significant connections in Fig. 3-9 may reveal underlying neurological changes induced by the disease. We observe increased functional connectivity between the parietal/posterior cingulate region and the frontal lobe and reduced functional connectivity between the parietal/posterior cingulate region and the temporal lobe in the schizophrenia population. Increased connectivity between the default network and the medial frontal lobe,

both at rest and during task, has been reported in schizophrenia [37, 104]. It is believed to interfere with perception of the external world through the misdirecting of attentional resources. Interestingly, decreased connectivity within the default network has been described as well [9, 103]. The later study reported decreased functional connectivity between the posterior cingulate gyrus and the hippocampus, which is consistent with our findings. The relationship between disruptions in functional connectivity and the integrity of the fornix has also been suggested. Similar to [103], our results reveal anatomical abnormalities within the two consistent anatomical connections (between the posterior cingulate and the hippocampus and between the superior temporal gyrus and the cuneus), which exhibits reduced anatomical connectivity in schizophrenia. We also observe a relationship between anatomical and functional connectivity disruptions within the posterior/temporal parts of the default network. Along with prior findings, our results suggest an inverse relationship between connectivity in the temporal and frontal parts of the default network. Such "anti-correlations" have been previously described between the default and task-related networks, but never within the default network itself.

We recognize the limitations of our joint generative model, especially those related to its simplicity. For example, we consider only direct anatomical connections between two regions while ignoring multistage pathways. In reality, there is some interaction between connections, which can be used to extract anatomical and functional networks within the brain. We model latent connectivity via discrete random variables, which may marginalize subtle variations between groups, and we assume that all subject data are drawn from the same distribution, whereas the strength of fMRI correlations and FA values can vary across subjects. Finally, the relationship between the modalities is captured through the link from anatomical connectivity to fMRI correlations.

These choices are deliberate on our part. Since the interaction between resting-state fMRI correlations and DWI tractography is neither well understood nor well characterized, we avoid placing strong prior assumptions on the structural-functional relationship. Our goal at this stage is to model what we observe from the data using a simple, robust framework. Furthermore, given the potentially large amounts of inter-subject variability and external noise, we intentionally simplify the model to reduce the number of parameters and avoid over-fitting. We address some of these limitations in the following chapter.

Chapter 4

Identifying Foci of a Neurological Disorder

Chapter 3 demonstrated that impairments of a neuropsychiatric disorder can be observed through aberrations in connectivity. Although our generative models can identify functional and anatomical connections influenced by the disease, connectivity results are difficult to interpret and validate. At present, the bulk of our knowledge about the brain is organized around regions (i.e., functional localization, tissue properties, morphometry) and not the connections between them. Moreover, it is nearly impossible to design non-invasive experiments that target a particular connection between two brain regions.

In this chapter we build upon our original framework to pinpoint regions, which we call *foci*, whose connectivity patterns are most disrupted by the disorder. Our method effectively translates differences in connectivity between a control and a clinical population into estimates of the regions associated with the disease. Using a probabilistic setting, we define a latent (hidden) graph that characterizes the network of abnormal functional connectivity emanating from the affected brain regions. This generates population differences in the observed fMRI correlations. We employ the variational EM algorithm to fit the model to the observed data. Our algorithm jointly infers the regions affected by the disease and the induced connectivity differences. To the best of our knowledge, ours is the first stochastic model to relate connectivity information to region labels.

We present two versions of the model. The first variant considers the complete graph of pairwise functional connections. The second model uses neural anatomy as a substrate for modeling functional differences. In particular, we rely on Diffusion Weighted Imaging (DWI) tractography to estimate the underlying white matter fibers in the brain. The latent anatomical connectivity inferred from these fibers constrains the graph of aberrant functional connections. Previous work in joint modeling of resting-state fMRI and DWI data suggests that a direct anatomical connection between two regions predicts a higher functional correlation [21, 41, 61, 97]; however, multi-stage pathways may explain some of

the functional effects. Since neural communication between brain regions is constrained by white matter fibers, we hypothesize that the strongest effects of a disorder will occur along direct anatomical connections. Hence, we model whole-brain functional connectivity but only use functional abnormalities between anatomically connected regions to identify the disease foci. The work in this chapter will appear in [95].

The remainder of this chapter is organized as follows. We present our generative model in Section 4.1 and develop the corresponding inference algorithm in Section 4.2. Section 4.3 describes the framework used for the empirical validation of our approach. Sections 4.4 and 4.5 report experimental results based on synthetic and clinical data, respectively. Finally, Section 4.6 discusses the behavior of our model, its advantages and drawbacks, and future directions of research.

4.1 Generative Model

We assume that the disorder is characterized by impairments in a small subset of brain regions, which we designate as *foci*. The impairments affect neural signaling along pathways associated with the diseased regions. We use a probabilistic framework to represent the interaction between regions and the effects of the disease. Once again, *latent* variables specify a template organization of the brain, which we cannot directly access. Instead, we observe noisy measurements of the hidden structure via resting-state fMRI correlations and DWI tractography. The fMRI and DWI signals are generated stochastically from a group-wise latent template shared by all subjects.

We first develop the model for functional data. This formulation serves as a foundation for incorporating anatomical connectivity, as presented later in the section. Table 4.1 summarizes our notation in this paper. Many of the variables and probability distributions presented in this section are identical to our previous formulation in Chapter 3. However, due to the novelty and complexity of our approach, we opt for a complete description of the models, which includes necessary repetition.

4.1.1 Functional Model

Fig. 4-1 depicts a network diagram of the brain and the corresponding graphical model for the functional connectivity data. The nodes in Fig. 4-1(a) denote regions in the brain, and edges correspond to pairwise functional connections between them. The green nodes/edges are healthy and the red nodes/edges are diseased.

Based on the region assignments, we define a binary graph T of aberrant functional connectivity using a simple set of rules: (1) a connection between two diseased regions is

Table 4.1: Random variables (top) and non-random parameters (bottom) in our graphical models shown in Figs. 4-1 and Fig. 4-2. The latent variables are discrete; the observed variables and non-random parameters are continuous.

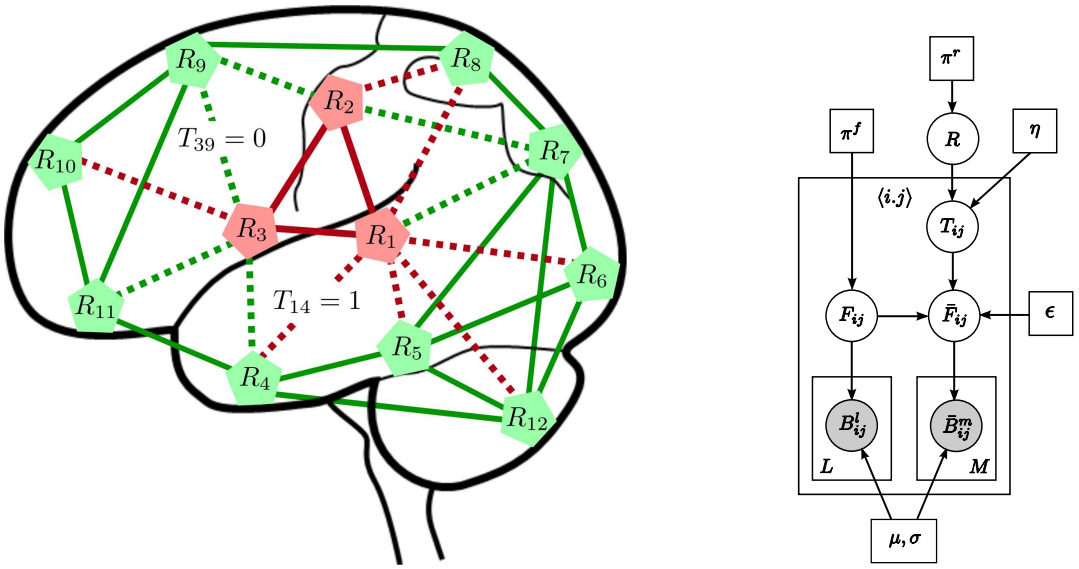
R	Binary vector that indicates the state (healthy/diseased) for each brain region i
A_{ij}	Latent anatomical connectivity between regions i and j
T_{ij}	Edge $\langle i, j \rangle$ in the latent graph of abnormal functional connectivity
F_{ij}	Latent functional connectivity between regions i and j (control group)
$\bar{F}_{ij}$	Latent functional connectivity between regions i and j (clinical group)
B^l_{ij}	Observed fMRI correlation between regions i and j in subject l (control group)
$\bar{B}^m_{ij}$	Observed fMRI correlation between regions i and j in subject m (clinical group)
D^l_{ij}	Observed DWI measure between regions i and j in subject l (control group)
$\bar{D}^m_{ij}$	Observed DWI measure between regions i and j in subject m (clinical group)
π^r	Prior for binary region indicator R_i
π^f	Prior for multinomial functional connectivity F_{ij}
π^a	Prior for binary anatomical connectivity A_{ij}
η	Probability of a diseased connection between a healthy and diseased node
ϵ	Probability of deviating from the latent graph of aberrant functional connectivity
μ_k	Mean fMRI correlations given $F_{ij} = k$ $(k = -1, 0, 1)$
σ^2_k	Variance of fMRI correlations given $F_{ij} = k$ $(k = -1, 0, 1)$
ρ_t	Probability of failing to find a white matter tract given $A_{ij} = t$ $(t = 0, 1)$
χ_t	Mean DWI value if there is a white matter tract given $A_{ij} = t$ $(t = 0, 1)$
ξ^2_t	Variance of DWI values if there is a white matter tract given $A_{ij} = t$ $(t = 0, 1)$

always abnormal ($T_{ij} = 1$, solid red lines in Fig. 4-1(a)), (2) a connection between two healthy regions is always healthy ($T_{ij} = 0$, solid green lines), and (3) a connection between a healthy and a diseased region is abnormal with probability η (dashed lines). We use the latent functional connectivity variables F_{ij} and $\bar{F}_{ij}$ to model the synchrony between two regions in the control and clinical populations, respectively. Ideally, $\bar{F}_{ij} \neq F_{ij}$ for abnormal connections and $\bar{F}_{ij} = F_{ij}$ for healthy connections. However, due to noise and intersubject variability, we assume that the latent templates can deviate from the graph T with (small) probability ϵ, which we estimate from the data.

The observed fMRI correlations B^l_{ij} provide noisy information about the latent network.

Disease Foci Let N be the total number of regions in the brain. The random variable $R = [R_1, \ldots, R_N]$ is a binary vector that indicates the state, healthy ($R_i = 0$) or diseased ($R_i = 1$), for each brain region ($i = 1, \ldots, N$). We assume an *i.i.d.* Bernoulli prior for the elements of R:

$$P(R_i; \pi^r) = (\pi^r)^{R_i}(1 - \pi^r)^{1-R_i}, \tag{4.1}$$



(a) Network Model of Brain Connectivity

(b) Graphical Model

Figure 4-1: (a) A network model of connectivity for the functional data. The nodes correspond to regions in the brain, and the lines denote pairwise functional connections between them. Only a subset of edges is shown; the model is defined on the full graph of pairwise connections. The green nodes and edges are normal. The red nodes are foci of the disease; red edges specify pathways of abnormal functional connectivity. The solid lines are deterministic given the region labels; the dashed lines are probabilistic. (b) The corresponding graphical model. Vector R specifies diseased regions. F_{ij} denotes the latent functional connectivity between regions i and j. B^l_{ij} is the observed fMRI measurements in the l^{th} subject. Variables associated with the diseased population are identified by an overbar.

where the scalar parameter π^r specifies the *a priori* probability that a region is diseased. The prior is shared by all nodes in the network.

Graph of Abnormal Connectivity The binary graph T represents the abnormal functional connectivity emanating from the disease foci. Each edge T_{ij} is generated independently given the labels of regions i and j:

$$P(T_{ij}|R_i, R_j; \eta) = \begin{cases} \delta_T(T_{ij}), & R_i = R_j = 0, \\ 1 - \delta_T(T_{ij}), & R_i = R_j = 1, \\ \eta^{T_{ij}}(1-\eta)^{1-T_{ij}}, & R_i \neq R_j, \end{cases} \tag{4.2}$$

where $\delta_T(\cdot)$ is an indicator function that equals to one if and only if its argument is zero,

and η is the scalar parameter that represents the probability of a connection between a healthy and a diseased region being altered.

Latent Functional Connectivity Using Eq. 3.2 we model the latent functional connectivity F_{ij} of the control population as a tri-state random variable drawn from a multinomial distribution with parameter π^f:

$$P(F_{ij};\pi^f) = \prod_{k=-1}^{1} \left(\pi_k^f\right)^{F_{ijk}}. \tag{4.3}$$

Once again, these states represent little or no functional co-activation ($F_{ij} = 0$), positive functional synchrony ($F_{ij} = 1$), and negative functional synchrony ($F_{ij} = -1$).

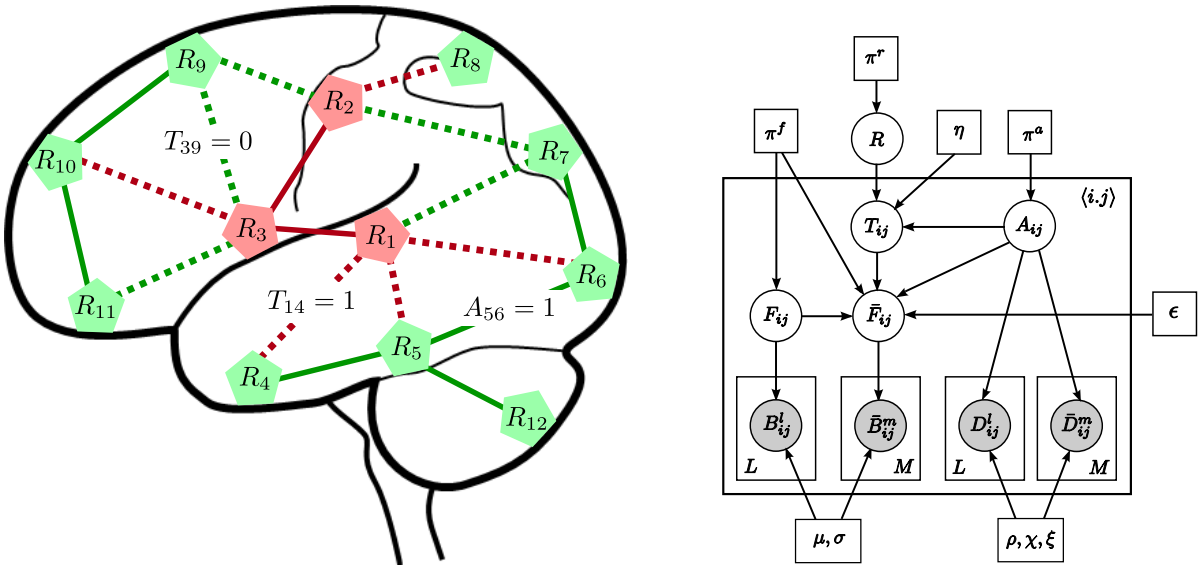
The latent functional connectivity $\bar{F}_{ij}$ of the clinical population is also tri-state and is based on F_{ij} and the graph T_{ij}. If the edge $\langle i,j\rangle$ is healthy ($T_{ij} = 0$), the functional connectivity of the clinical population is equal to that of the control population with probability $1-\epsilon$, and it differs with probability ϵ. Conversely, if the edge $\langle i,j\rangle$ is diseased ($T_{ij} = 1$), then the functional connectivity of the clinical population differs from the control population with probability $1-\epsilon$, and it is equal with probability ϵ. Formally,

$$P(\bar{F}_{ij}|F_{ij},T_{ij};\epsilon) = \left[(1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\right]^{1-T_{ij}}\left[\epsilon^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{1-\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\right]^{T_{ij}}. \tag{4.4}$$

fMRI Likelihood Let L be the number of subjects in the control population and M be the number of subjects in the clinical population. The BOLD fMRI correlation B_{ij}^l between regions i and j in the l^{th} subject of the control population is a noisy observation of the functional connectivity indicator F_{ij}. In particular, B_{ij}^l is a Gaussian random variable whose mean and variance depend on the value of F_{ij}:

$$P(B_{ij}^l|F_{ij};\{\mu,\sigma^2\}) = \prod_{k=-1}^{1} \mathcal{N}\left(B_{ij}^l;\mu_k,\sigma_k^2\right)^{F_{ijk}}, \tag{4.5}$$

where $\mathcal{N}(\cdot\,;\mu,\sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2. We fix $\mu_0 = 0$ to center the parameter estimates. This acts as a type of regularization for the model. Without this constraint, the ML data distributions overlap significantly, which makes it difficult to infer latent functional connectivity. The likelihood for the clinical population $\bar{B}_{ij}^m$ has the same functional form and parameter values as Eq. (4.5) but uses the clinical template $\bar{F}_{ij}$ instead of the control template F_{ij}.



(a) Network Model of Brain Connectivity (b) Graphical Model

Figure 4-2: (a) A network model of connectivity. The nodes correspond to regions in the brain, and the lines denote anatomical connections between them. The green nodes and edges are normal. The red nodes are foci of the disease; red edges specify pathways of abnormal functional connectivity. The solid lines are deterministic given the region labels; the dashed lines are probabilistic. (b) Corresponding graphical model. Vector R specifies diseased regions. A_{ij} represents the latent anatomical connectivity between regions i and j. F_{ij} denotes the corresponding latent functional connectivity. D^l_{ij} and B^l_{ij} are the observed DWI and fMRI measurements, respectively, in the l^{th} subject. Variables associated with the diseased population are identified by an overbar.

4.1.2 Multi-modal Analysis

Since functional communication in the brain is constrained by neural axons, our second model assumes that the salient effects of a disorder will occur along anatomical pathways. This extension is illustrated in Fig. 4-2. The edges in Fig. 4-2(a) correspond to neural connections, which are captured by latent anatomical connectivity A_{ij}. Specifically, the presence or absence of an edge $\langle i, j \rangle$ in the network is governed by the binary value of A_{ij}. The anatomical network structure is shared between the control and clinical populations. The regions in this work correspond to (large) Brodmann areas. Our results in Chapter 3 suggest that the anatomical differences between schizophrenia patients and normal controls are very small in this case. Once again, the observed DWI measurements D^l_{ij} and fMRI correlations B^l_{ij} provide noisy information about the latent network structure.

Latent Anatomical Connectivity Once again, the latent anatomical connectivity variable A_{ij} indicates the presence or absence of a *direct* anatomical pathway between regions i and j. We model A_{ij} as a binary random variable according to Eq. (3.1):

$$P(A_{ij};\pi^a) = (\pi^a)^{A_{ij}}(1-\pi^a)^{1-A_{ij}}. \tag{4.6}$$

where π^a is the *a priori* probability that a connection is present.

Graph of Abnormal Connectivity The binary graph T of aberrant functional connectivity is now defined along latent anatomical pathways. Therefore, we modify the rules from Section 4.1.1 and generate the edge T_{ij} between regions i and j as follows:

$$P(T_{ij}|R_i,R_j;\eta) = \begin{cases} \delta_T(T_{ij}), & A_{ij}=0, \\ \delta_T(T_{ij}), & A_{ij}=1, R_i=R_j=0, \\ 1-\delta_T(T_{ij}), & A_{ij}=1, R_i=R_j=1, \\ \eta^{T_{ij}}(1-\eta)^{1-T_{ij}}, & A_{ij}=1, R_i\neq R_j, \end{cases} \tag{4.7}$$

In particular, the first condition in Eq. (4.7) states that $T_{ij}=0$ when the corresponding anatomical connection is absent.

Functional Connectivity of the Clinical Population We adapt the distribution for the latent functional connectivity $\bar{F}_{ij}$ in Eq. (4.4) to reflect the anatomical constraint:

$$P(\bar{F}_{ij}|F_{ij},T_{ij},A_{ij};\epsilon) = \left(\left[(1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \right]^{1-T_{ij}} \left[\epsilon^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{1-\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \right]^{T_{ij}} \right)^{A_{ij}} \left(\prod_{k=-1}^{1} \left(\pi_k^f\right)^{\bar{F}_{ijk}} \right)^{1-A_{ij}}. \tag{4.8}$$

If there is a latent anatomical connection between regions i and j ($A_{ij}=1$), then $\bar{F}_{ij}$ is generated according to Eq. (4.4). If there is no anatomical connection ($A_{ij}=0$), then the final term of Eq. (4.8) implies that $\bar{F}_{ij}$ is drawn from the prior π^f, irrespective of F_{ij}.

DWI Likelihood Using Eq. (3.3), the DWI measurement D_{ij}^l for the l^{th} subject in the control population is a noisy observation of the anatomical connectivity A_{ij}:

$$P(D_{ij}^l|A_{ij};\{\rho,\chi,\xi^2\}) = \mathcal{P}_0(D_{ij}^l;\{\rho,\chi,\xi^2\})^{1-A_{ij}} \cdot \mathcal{P}_1(D_{ij}^l;\{\rho,\chi,\xi^2\})^{A_{ij}}, \tag{4.9}$$

where $\mathcal{P}_t(D_{ij}) = \rho_t \delta(D_{ij}) + (1-\rho_t)\mathcal{N}(D_{ij}; \chi_t, \xi_t^2)$ for $t = 0, 1$, and $\delta(\cdot)$ is the Dirac delta function. The data $\bar{D}_{ij}^m$ of the clinical population follows the same likelihood.

We do not modify the fMRI likelihood to reflect latent anatomical connectivity, as in Eq. (3.4). Rather, we assume that *both* the interaction between anatomy and function and the effects of a disorder occur within the latent structure.

4.2 Variational Inference

Since we are primarily interested in the region labels R, we opt to marginalize out the graph structure T. This simplifies the relationship between R and the observed data.

The only term which is affected by the marginalization is the conditional distribution of the clinical template $\bar{F}_{ij}$, which now depends on the values of R_i and R_j. Specifically, we have

$$P(\bar{F}_{ij}|F_{ij}, R_i, R_j; \eta, \epsilon) = \begin{cases} (1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & R_i = R_j = 0, \\ \epsilon^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{1-\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & R_i = R_j = 1, \\ \epsilon_1^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{1-\epsilon_1}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & R_i \neq R_j, \end{cases} \tag{4.10}$$

for the functional model and

$$P(\bar{F}_{ij}|F_{ij}, R_i, R_j, A_{ij}; \eta, \epsilon) = \begin{cases} (1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & A_{ij} = 1, R_i = R_j = 0, \\ \epsilon^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{1-\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & A_{ij} = 1, R_i = R_j = 1, \\ \epsilon_1^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}} \left(\frac{1-\epsilon_1}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & A_{ij} = 1, R_i \neq R_j, \\ \prod_{k=-1}^{1} \left(\pi^f\right)^{\bar{F}_{ijk}}, & A_{ij} = 0. \end{cases} \tag{4.11}$$

for the joint model, where $\epsilon_1 = \eta\epsilon + (1-\eta)(1-\epsilon)$. It is easy to see that ϵ_1 reflects the coupling between the graph prior η and latent noise variable ϵ when the region labels differ.

We employ a maximum likelihood (ML) framework to fit the model to the data. Unlike our previous formulation in Chapter 3, the region variable R induces a complex coupling between pairwise connections forcing us to adopt a variational approximation [57] for the posterior probability distribution when deriving the EM algorithm for parameter estimation.

4.2.1 Functional Model

Let $Y = \{B, \bar{B}\}$ and $\Theta = \{\pi, \eta, \epsilon, \mu, \sigma^2\}$ denote the observed fMRI measurements and the set of model parameters, respectively. Our variational posterior assumes the following form:

$$Q(R, F, \bar{F}) = Q^r(R) \cdot Q^c(F, \bar{F}) = Q^r(R) \prod_{<i,j>} Q^c_{ij}(F_{ij}, \bar{F}_{ij}), \tag{4.12}$$

where $Q^r(\cdot)$ is a distribution over the length-N binary vector R and $Q^c_{ij}(\cdot)$ is an 9-state multinomial distribution corresponding to all configurations of latent functional connectivity. This factorization yields a tractable inference algorithm while preserving the dependency between F_{ij}, and $\bar{F}_{ij}$ given the region indicator vector R.

We use a variational EM formulation [23] to obtain the posterior distribution $Q(\cdot)$ and model parameters Θ which minimize the variational free energy

$$\mathcal{FE} = -E_Q \left[\log P(R, F, \bar{F}, Y; \Theta)\right] - \mathcal{H}(Q), \tag{4.13}$$

where the joint log-likelihood of all hidden and observed variables is obtained by combining the prior and likelihood distributions from Section 4.1.1 with Eq. (4.10):

$$\begin{aligned}
\log P(R,F, \bar{F}, Y; \Theta) = {} & \log(\pi^r) \sum_{i=1}^{N} R_i + \log(1-\pi^r) \sum_{i=1}^{N} (1 - R_i) + \sum_{\langle i,j \rangle} \sum_{k=-1}^{1} F_{ijk} \log\left(\pi^f_k\right) \\
& + \sum_{\langle i,j \rangle} (1-R_i)(1-R_j) \left(F_{ij}^{\mathrm{T}} \bar{F}_{ij} \log(1-\epsilon) + (1 - F_{ij}^{\mathrm{T}} \bar{F}_{ij}) \log\left(\frac{\epsilon}{2}\right)\right) \\
& + \sum_{\langle i,j \rangle} (R_i R_j) \left(F_{ij}^{\mathrm{T}} \bar{F}_{ij} \log(\epsilon) + (1 - F_{ij}^{\mathrm{T}} \bar{F}_{ij}) \log\left(\frac{1-\epsilon}{2}\right)\right) \\
& + \sum_{\langle i,j \rangle} (R_i(1-R_j) + (1-R_i)R_j) \left(F_{ij}^{\mathrm{T}} \bar{F}_{ij} \log(\epsilon_1) + (1 - F_{ij}^{\mathrm{T}} \bar{F}_{ij}) \log\left(\frac{1-\epsilon_1}{2}\right)\right) \\
& + \sum_{\langle i,j \rangle} \sum_{k=-1}^{1} \left[F_{ijk} \sum_{l=1}^{L} \log \mathcal{N}\left(B^l_{ij}; \mu_k, \sigma^2_k\right) + \bar{F}_{ijk} \sum_{m=1}^{M} \log \mathcal{N}\left(\bar{B}^m_{ij}; \mu_k, \sigma^2_k\right)\right].
\end{aligned} \tag{4.14}$$

E-Step: For a fixed setting of model parameters $\hat{\Theta}$, the free energy in Eq. (4.13) can be expanded as follows:

$$\mathcal{FE} = -\sum_{R} Q^r(R) \sum_{F,\bar{F}} \log P(R, F, \bar{F}, Y; \hat{\Theta}) + \sum_{R} Q^r(R) \log Q^r(R) - \mathcal{H}(Q^c)$$

$$= \sum_R Q^r(R) \left[-E_{Q^c} \left[\log P(R, F, \bar{F}, Y; \hat{\Theta}) \right] + \log Q^r(R) \right] - \mathcal{H}(Q^c). \tag{4.15}$$

We define the (normalized) probability distribution $\tilde{P}(R; \Theta)$ as

$$\tilde{P}(R; \hat{\Theta}) \propto \exp \left\{ E_{\hat{Q}^c} \left[\log P(R, F, \bar{F}, Y; \hat{\Theta}) \right] \right\}. \tag{4.16}$$

By substituting Eq. (4.16) into Eq. (4.15), it is trivial to show that

$$\mathcal{FE} = KL \left(Q^r(R) || \tilde{P}(R; \hat{\Theta}) \right) + const., \tag{4.17}$$

where $KL(p||q)$ is the the Kullback-Leibler (KL) divergence from the distribution $p(\cdot)$ to the distribution $q(\cdot)$, and the additional constants do not depend on R.

Using a similar expansion, we can also show that

$$\mathcal{FE} = KL \left(Q^c(F, \bar{F}) || \tilde{P}(F, \bar{F}; \hat{\Theta}) \right) + const., \tag{4.18}$$

where $\tilde{P}(F, \bar{F}; \hat{\Theta}) \propto \exp \left\{ E_{\hat{Q}^r} \left[\log P(R, F, \bar{F}, Y; \hat{\Theta}) \right] \right\}$

Since the KL divergence is non-negative, Eqs. (4.17-4.18) give us the following fixed-point equations for the variational posterior $\hat{Q}(\cdot)$:

$$\hat{Q}^c(F, \bar{F}) = \tilde{P}(F, \bar{F}; \hat{\Theta}) \propto \exp \left\{ E_{\hat{Q}^r} \left[\log P(R, F, \bar{F}, Y; \hat{\Theta}) \right] \right\}, \tag{4.19}$$

$$\hat{Q}^r(R) = \tilde{P}(R; \hat{\Theta}) \propto \exp \left\{ E_{\hat{Q}^c} \left[\log P(R, F, \bar{F}, Y; \hat{\Theta}) \right] \right\}. \tag{4.20}$$

We alternatively update $\hat{Q}^r(R)$ and $\hat{Q}^c(F, \bar{F})$, according to the above expressions, until convergence. Specifically, we employ Gibbs sampling to obtain samples $\mathcal{S} = \{R^s\}$ from $\hat{Q}^r(R)$. Based on the joint log-likelihood in Eq. (4.14), the right-hand side of Eq. (4.19) can be expressed in terms of

$$\hat{q}_{ij}^{00} \triangleq E_{\hat{Q}^r}[(1 - R_i)(1 - R_j)], \tag{4.21}$$

$$\hat{q}_{ij}^{11} \triangleq E_{\hat{Q}^r}[R_i R_j], \tag{4.22}$$

$$\hat{q}_{ij}^{10} \triangleq E_{\hat{Q}^r}[R_i(1 - R_j) + (1 - R_i)R_j]. \tag{4.23}$$

We approximate these quantities using averages of R_i and $R_i R_j$ over the elements of $\mathcal{S}$.

To update $\hat{Q}^c(\cdot)$, we evaluate the right-hand side of Eq. (4.19) for each configuration $F_{ij} = k$, $\bar{F}_{ij} = k'$ $(k, k' \in \{-1, 0, 1\})$ and normalize over all nine combinations of k, k' to

obtain a valid probability distribution.

According to the joint log-likelihood in Eq. (4.14), the right-hand side of Eq. (4.20) is given in terms of $E_{\hat{Q}^c}[F_{ij}^{\mathrm{T}}\bar{F}_{ij}]$. Since F_{ij} and $\bar{F}_{ij}$ are indicator variables, this quantity can be evaluated as

$$\hat{p}_{ij} \triangleq E_{\hat{Q}^c}[F_{ij}^{\mathrm{T}}\bar{F}_{ij}] = \sum_{k=-1}^{1} \hat{Q}_{ij}^c(F_{ij}=k, \bar{F}_{ij}=k). \tag{4.24}$$

Similar to Section 3.2 the model parameter estimates $\hat{\Theta}$ rely on marginal probabilities of $\hat{Q}^c(F,\bar{F})$. We compute these quantities after convergence of the variational posterior distribution $\hat{Q}(\cdot)$:

$$\hat{s}_{ijk} = \hat{P}(F_{ij}=k|Y;\Theta) = \sum_{\bar{F}_{ij}} \hat{Q}_{ij}^c(F_{ijk}=1, \bar{F}_{ij}), \tag{4.25}$$

$$\hat{u}_{ijk} = \hat{P}(\bar{F}_{ij}=k|Y;\Theta) = \sum_{F_{ij}} \hat{Q}_{ij}^c(F_{ij}, \bar{F}_{ijk}=1). \tag{4.26}$$

M-Step: We fix the posterior probability estimates $\hat{Q}(R,F,\bar{F})$ and update the model parameter estimates $\hat{\Theta}$ by differentiating Eq. (4.13) with respect to each element of Θ and setting the gradient equal to zero.

The update for π^r involves averaging the proportion of diseased regions across Gibbs samples:

$$\hat{\pi}^r = \frac{1}{NS}\sum_{i=1}^{N}\sum_{s=1}^{S} R_i^s. \tag{4.27}$$

The multinomial prior reduces to an average over the marginal posterior distribution:

$$\hat{\pi}_k^f = \frac{1}{C}\sum_{\langle i,j\rangle} \hat{s}_{ijk}, \tag{4.28}$$

where C is the total number of pairwise connections. The fMRI likelihood parameter estimates are computed as weighted statistics of the data:

$$\mu_k = \begin{cases} \frac{\sum_{\langle i,j\rangle}\left[\hat{s}_{ijk}\sum_l B_{ij}^l + \hat{u}_{ijk}\sum_m \bar{B}_{ij}^m\right]}{\sum_{\langle i,j\rangle}\left[L\cdot\hat{s}_{ijk}+M\cdot\hat{u}_{ijk}\right]}, & k=\pm 1, \\ 0, & k=0, \end{cases} \tag{4.29}$$

$$\sigma_k^2 = \frac{\sum_{\langle i,j\rangle}\left[\hat{s}_{ijk}\sum_l (B_{ij}^l-\hat{\mu}_k)^2 + \hat{u}_{ijk}\sum_m(\bar{B}_{ij}^m-\hat{\mu}_k)^2\right]}{\sum_{\langle i,j\rangle}\left[L\cdot\hat{s}_{ijk}+M\cdot\hat{u}_{ijk}\right]}, \tag{4.30}$$

where we have fixed $\mu_0 = 0$ for the component that represents zero functional synchrony to center the parameter estimates and regularize the model.

The parameters η and ϵ are tied through Eq. (4.4). We use Newton's method to jointly update η and ϵ. The only term of the free energy objective that depends on either η or ϵ is

$$E_{\hat{Q}}\left[\log P(\bar{F}|F,R;\hat{\Theta})\right] = \sum_{\langle i,j\rangle}\Big[\hat{q}_{ij}^{00}\left(\hat{p}_{ij}\log(1-\epsilon)+(1-\hat{p}_{ij})\log\left(\frac{\epsilon}{2}\right)\right) + \hat{q}_{ij}^{11}\left(\hat{p}_{ij}\log(\epsilon)+(1-\hat{p}_{ij})\log\left(\frac{1-\epsilon}{2}\right)\right)+\hat{q}_{ij}^{10}\left(\hat{p}_{ij}\log(\epsilon_1)+(1-\hat{p}_{ij})\log\left(\frac{1-\epsilon_1}{2}\right)\right)\Big], \tag{4.31}$$

where we have substituted the definitions from Eqs. (4.21-4.24) into the expression. The Newton's method update for η, ϵ is based on the following iteration:

$$\begin{bmatrix}\hat{\epsilon}_{n+1}\\ \hat{\eta}_{n+1}\end{bmatrix} = \begin{bmatrix}\hat{\epsilon}_{n}\\ \hat{\eta}_{n}\end{bmatrix} - \begin{bmatrix}\frac{\partial^2\mathcal{FE}}{\partial\epsilon^2} & \frac{\partial^2\mathcal{FE}}{\partial\epsilon\partial\eta}\\ \frac{\partial^2\mathcal{FE}}{\partial\epsilon\partial\eta} & \frac{\partial^2\mathcal{FE}}{\partial\eta^2}\end{bmatrix}^{-1}\begin{bmatrix}\frac{\partial\mathcal{FE}}{\partial\epsilon}\\ \frac{\partial\mathcal{FE}}{\partial\eta}\end{bmatrix}\Bigg|_{(\epsilon,\eta)=(\hat{\epsilon}_n,\hat{\eta}_n)} \tag{4.32}$$

The first and second derivatives of Eq. (4.31) with respect to η, ϵ are

$$\begin{aligned}
\frac{\partial\mathcal{FE}}{\partial\epsilon} &= \frac{1}{\epsilon}\sum_{\langle i,j\rangle}\left[\hat{q}_{ij}^{11}\cdot\hat{p}_{ij}+\hat{q}_{ij}^{00}(1-\hat{p}_{ij})\right] - \frac{1}{1-\epsilon}\sum_{\langle i,j\rangle}\left[\hat{q}_{ij}^{11}\cdot\hat{p}_{ij}+\hat{q}_{ij}^{00}\cdot(1-\hat{p}_{ij})\right]\\
&\quad + \frac{2\eta-1}{\epsilon_1}\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}\cdot\hat{p}_{ij} - \frac{2\eta-1}{1-\epsilon_1}\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}(1-\hat{p}_{ij}),\\
\frac{\partial\mathcal{FE}}{\partial\eta} &= \frac{2\epsilon-1}{\epsilon_1}\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}\cdot\hat{p}_{ij} - \frac{2\epsilon-1}{1-\epsilon_1}\sum_{\langle i,j\langle}\hat{q}_{ij}^{10}(1-\hat{p}_{ij}),\\
\frac{\partial^2\mathcal{FE}}{\partial\epsilon^2} &= -\frac{1}{\epsilon^2}\sum_{\langle i,j\rangle}\left[\hat{q}_{ij}^{11}\cdot\hat{p}_{ij}+\hat{q}_{ij}^{00}(1-\hat{p}_{ij})\right] - \frac{1}{(1-\epsilon)^2}\sum_{\langle i,j\rangle}\left[\hat{q}_{ij}^{11}\cdot\hat{p}_{ij}+\hat{q}_{ij}^{00}(1-\hat{p}_{ij})\right]\\
&\quad - \left(\frac{2\eta-1}{\epsilon_1}\right)^2\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}\cdot\hat{p}_{ij} - \left(\frac{2\eta-1}{1-\epsilon_1}\right)^2\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}(1-\hat{p}_{ij}),\\
\frac{\partial^2\mathcal{FE}}{\partial\eta^2} &= -\left(\frac{2\epsilon-1}{\epsilon_1}\right)^2\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}\cdot\hat{p}_{ij} - \left(\frac{2\epsilon-1}{1-\epsilon_1}\right)^2\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}(1-\hat{p}_{ij}),\\
\frac{\partial^2\mathcal{FE}}{\partial\eta\partial\epsilon} &= \frac{1}{\epsilon_1^2}\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}\cdot\hat{p}_{ij} - \frac{1}{(1-\epsilon_1)^2}\sum_{\langle i,j\rangle}\hat{q}_{ij}^{10}(1-\hat{p}_{ij}).
\end{aligned}$$

4.2.2 Joint Model

The variational EM algorithm can be easily extended to incorporate anatomical connectivity. Below, we let $Y = \{B, \bar{B}, D, \bar{D}\}$ denote the observed fMRI and DWI measurements, respectively, and we let $\Theta = \{\pi, \eta, \epsilon, \mu, \sigma^2, \rho, \chi, \xi^2\}$ be the set of model parameters. Since A_{ij} is binary and F_{ij} and $\bar{F}_{ij}$ are tri-state, the variational posterior is

$$Q(R, A, F, \bar{F}) = Q^r(R) \cdot Q^c(A, F, \bar{F}) = Q^r(R) \prod_{<i,j>} Q^c_{ij}(A_{ij}, F_{ij}, \bar{F}_{ij}), \tag{4.33}$$

where $Q^r(\cdot)$ is a distribution over the length-N binary vector R and $Q^c_{ij}(\cdot)$ is an 18-state multinomial distribution corresponding to all configurations of latent anatomical and functional connectivity.

E-Step: For a fixed setting of model parameters $\hat{\Theta}$, we alternatively update $\hat{Q}^r(R)$ and $\hat{Q}^c(A, F, \bar{F})$ according to the following expressions:

$$\hat{Q}^r(R) \propto \exp\left\{E_{\hat{Q}^c}\left[\log P(R, A, F, \bar{F}, Y; \Theta)\right]\right\}, \tag{4.34}$$

$$\hat{Q}^c(A, F, \bar{F}) \propto \exp\left\{E_{\hat{Q}^r}\left[\log P(R, A, F, \bar{F}, Y; \Theta)\right]\right\}. \tag{4.35}$$

Once again, we use Gibbs sampling to obtain samples $\mathcal{S} = \{R^s\}$ from Eq. (4.34) and then evaluate $\hat{Q}^c(A, F, \bar{F})$ using averages of R_i and $R_i R_j$ over the elements of $\mathcal{S}$. We update $\hat{Q}^c(\cdot)$ by evaluating the right-hand side of Eq. (4.35) for all 18 configurations of $\{A_{ij}, F_{ij}, \bar{F}_{ij}\}$ and normalizing. $\hat{Q}^r(R)$ is given in terms of $E_{\hat{Q}^c}[A_{ij} \cdot F_{ij}^{\mathrm{T}} \bar{F}_{ij}]$ and $E_{\hat{Q}^c}[A_{ij}(1 - F_{ij}^{\mathrm{T}} \bar{F}_{ij})]$, which are evaluated similar to Eq. (4.24).

M-Step: As in the preceding section, we define the marginal posterior probability for latent anatomical connectivity

$$\hat{a}_{ij} = \sum_{F_{ij}, \bar{F}_{ij}} Q^{ij}_V(A_{ij} = 1, F_{ij}, \bar{F}_{ij}).$$

Additionally, we let L^0_{ij} be the number of control subjects for whom $D^l_{ij} = 0$ and M^0_{ij} be the number of schizophrenia patients for whom $\bar{D}^m_{ij} = 0$.

The updates for $\hat{\pi}^r$ and the fMRI likelihood parameters remain unchanged. The prior estimate for $\hat{\pi}^a$ is an intuitive average of marginal probabilities:

$$\hat{\pi}^a = \frac{1}{C} \sum_{\langle i,j \rangle} \hat{a}_{ij}, \tag{4.36}$$

where C is the total number of pairwise connections.

The prior π^f interacts with A, F and $\bar{F}$ due to Eqs. (4.3) and (4.8). Minimizing the free energy with respect to π^f_k results in the following update equation:

$$\hat{\pi}^f_k = \frac{\sum_{\langle i,j \rangle} \left[\hat{s}_{ijk} + \sum_{F_{ij}} \hat{Q}^c_{ij}(A_{ij} = 0, F_{ij}, \bar{F}_{ij} = k) \right]}{C + \sum_{\langle i,j \rangle} \sum_{F_{ij}} \hat{Q}^c_{ij}(A_{ij} = 0, F_{ij}, \bar{F}_{ij} = k)}. \tag{4.37}$$

The probability $\hat{\rho}_1$ is the empirical likelihood of not finding a white matter tract between two regions given an underlying anatomical connection

$$\hat{\rho}_1 = \frac{\sum_{\langle i,j \rangle} \hat{a}_{ij}(L^0_{ij} + M^0_{ij})}{\sum_{\langle i,j \rangle} \hat{a}_{ij}(L + M)}. \tag{4.38}$$

The Gaussian likelihood parameters for the DWI measurements are given by the weighted empirical mean and empirical variance over all *nonzero* values

$$\hat{\chi}_1 = \frac{\sum_{\langle i,j \rangle} \hat{a}_{ij} \left(\sum_{l: D^l_{ij} > 0} D^l_{ij} + \sum_{m: \bar{D}^m_{ij} > 0} \bar{D}^m_{ij} \right)}{\sum_{\langle i,j \rangle} \hat{a}_{ij}(L - L^0_{ij} + M - M^0_{ij})}, \tag{4.39}$$

$$\hat{\xi}^2_1 = \frac{\sum_{\langle i,j \rangle} \hat{a}_{ij} \left(\sum_{l: D^l_{ij} > 0} (D^l_{ij} - \hat{\chi}_1)^2 + \sum_{m: \bar{D}^m_{ij} > 0} (\bar{D}^m_{ij} - \hat{\chi}_1)^2 \right)}{\sum_{\langle i,j \rangle} \hat{a}_{ij}(L - L^0_{ij} + M - M^0_{ij})}. \tag{4.40}$$

Once again, the parameter updates for $\{\rho_0, \chi_0, \xi^2_0\}$ are trivially obtained from these expressions by replacing $\hat{a}_{ij}$ with $(1 - \hat{a}_{ij})$.

Similar to the previous algorithm, we update $\hat{\eta}$ and $\hat{\epsilon}$ using a Newton's method iteration. We omit the expressions for the first and second derivatives, as they do not provide additional insight into the algorithm.

4.2.3 Implementation Details

In this section we describe the optimization choices in our implementation of the variational EM algorithm.

Initialization

Like many hill-climbing methods, the quality of our results depends on proper initialization. For the variational algorithm, it suffices to initialize the model parameters $\Theta =$

$\{\pi, \eta, \epsilon, \mu, \sigma^2, \rho, \chi, \xi^2\}$ and the marginal posterior statistics for R, $E_{\hat{Q}^r}[R_i]$ and $E_{\hat{Q}^r}[R_i R_j]$. The algorithm proceeds by computing the joint posterior distribution $\hat{Q}^c(\cdot)$ in the E-step and alternatively updates $\hat{Q}^r(R)$ and $\hat{Q}^c(\cdot)$ until convergence. We then estimate the model parameters in the M-step and iterate.

We initialize the prior parameters π^a, π^f, the probability of not detecting white matter fibers ρ, and the Gaussian variances σ^2, ξ^2 using statistics of the data. We also set the initial value of the latent noise parameter $\epsilon = 0.01$, which encourages consistency between the region labels and the observed connectivity data. Perturbations in these values do not seem to impact our final solution. We uniformly sample the initial values for the Bernoulli region prior π^r and for the graph parameter η from the interval $[0.2, 0.5]$. Larger values of π^r and η encourage the algorithm to select more foci during the first iteration.

The initial values of the Gaussian means $\{\mu, \chi\}$ largely determine the initial latent connectivity assignments; hence, they have the biggest influence on the final solution. Empirically, our model prefers sparse solutions for the region label vector R. If the initial connectivity data is too similar between the populations, then the algorithm will converge to a sub-optimal solution (with respect to the free energy) in which none of the regions are diseased. Therefore, we initialize $\{\mu, \chi\}$ to exaggerate the relevant functional connectivity differences. In particular, we uniformly sample each of these values from a specific interval such that: (1) the initial distribution of latent functional connectivity is roughly uniform, and (2) the initial graph of latent anatomical connectivity is fairly dense. These choices improve our chances of finding the global optimum. We emphasize that our initialization is still fairly naïve and that we do not place strong *a priori* assumptions on the model. Rather, we inject just enough flexibility to allow the algorithm to efficiently traverse the parameter space.

Finally, we initialize the posterior statistics by computing the mean fMRI correlation across subjects and clustering these values for each connection. We select regions with the highest number of connections with different cluster assignments between the two populations as the set of disease foci. We uniformly sample $E_{\hat{Q}^r}[R_i] \in [0.8, 1]$ for each selected focus region and $E_{\hat{Q}^r}[R_i] \in [0, 0.2]$ for non-foci. The pairwise statistics are computed as $E_{\hat{Q}^r}[R_i R_j] = E_{\hat{Q}^r}[R_i] \cdot E_{\hat{Q}^r}[R_j]$.

Empirically, we find that the final region posterior distribution is fairly stable within the above parameter ranges. We run the algorithm five times for the functional model and ten times for the joint model to sample the solution space; we then select the solution with the lowest free energy.

Gibbs Sampling

In the E-step, we Gibbs sample the region indicator vector R from the posterior distribution $\hat{Q}^r(R)$. Specifically, for each region i, we sample the value $R_i \in \{0, 1\}$ while fixing the other region assignments. The regions are updated in random order. To speed up computation, we run Gibbs sampling simultaneously on four processors and combine the resulting samples. In each case, the first 500 iterations are used for burn-in; we collect 50 samples spaced 100 iterations apart. Here, one iteration refers to updating all elements of the vector R.

Convergence and Runtime

Convergence of our algorithms was based on a relative change in free energy of less than 10^{-4} between consecutive iterations. On average, both algorithms converge in less than 10 iterations (E-step/M-step updates). The algorithms are slow due to the Gibbs sampling procedure. The approximate runtime is 30 minutes for a single initialization using the functional model and 15 minutes for a single initialization using the joint model. However, this can be greatly improved using more parallel computation. All simulations were performed using MATLAB on a modern quad processor workstation.

4.3 Model Evaluation

4.3.1 Identifying Disease Foci

The marginal posterior distribution $\hat{Q}^r(R)$ informs us about the disease foci. Let $\hat{q}_i$ denote the marginal probability that region i is diseased. We estimate this quantity by averaging across Gibbs samples $\mathcal{S}$:

$$\hat{q}_i \triangleq \hat{Q}^r(R_i = 1) = \frac{1}{S} \sum_{s=1}^{S} R_i^s. \tag{4.41}$$

The joint distributions in our method are non-Gaussian due to multiplicative interactions between latent variables and the effects of unknown nonrandom parameters. Therefore, we evaluate the significance of our model through non-parametric permutation tests. To construct the null distribution for $\hat{q}_i$, we randomly permute the subject diagnoses (NC vs. SZ) 1,000 times. For each permutation, we fit the model and compute the statistic in Eq. (4.41). The significance of each region is equal to the proportion of permutations that yield a larger value of $\hat{q}_i$ than is obtained under the true labeling.

4.3.2 Graph of Abnormal Connectivity

The graph of connectivity differences T in Section 4.1 provides insight into the behavior of individual connections. Although we marginalize this random variable prior to inference, we can retroactively estimate T based on the maximum a posteriori (MAP) estimate of each R_i and the ML parameter estimates.

Given R, our models decouple by pairwise connection, so we can independently assign each T_{ij}. Recall that $T_{ij} = 0$ indicates a healthy edge and $T_{ij} = 1$ denotes a diseased connection. Based on our construction in Section 4.1, many of the values T_{ij} are deterministic. For example, Eq. (4.2) of the functional model implies that $T_{ij} = 0$ if $R_i = R_j = 0$ and $T_{ij} = 1$ if $R_i = R_j = 1$. For connections $\langle i, j \rangle$ such that $R_i \neq R_j$, we select the value $T_{ij} \in \{0, 1\}$ which optimizes:

$$\hat{T}_{ij} = \arg\max_{T_{ij}} E_{\hat{Q}^c}\left[\log P(T_{ij}|F_{ij}, \bar{F}_{ij}, \hat{R}; \hat{\Theta})\right] = \arg\max_{T_{ij}} E_{\hat{Q}^c}\left[\log P(T_{ij}, F_{ij}, \bar{F}_{ij}|\hat{R}; \hat{\Theta})\right]$$

Eq. (4.2) further specifies that if the region labels differ, the prior on T_{ij} is Bernoulli with parameter η. Additionally, if $T_{ij} = 0$ (the edge $\langle i, j \rangle$ is healthy), then the functional connectivity is the same in both populations with probability $1 - \epsilon$, and it differs with probability ϵ. Likewise, if $T_{ij} = 1$, then the functional connectivity differs between the populations with probability $1 - \epsilon$ and is the same with probability ϵ. After some algebraic manipulations we arrive at the decision rule for the functional model

$$\log(\hat{\eta}) + \hat{p}_{ij}\log(\hat{\epsilon}) + (1 - \hat{p}_{ij})\log\left(\frac{1-\hat{\epsilon}}{2}\right) \underset{T_{ij}=1}{\overset{T_{ij}=0}{\gtrless}} \log(1 - \hat{\eta}) + \hat{p}_{ij}\log(1 - \hat{\epsilon}) + (1 - \hat{p}_{ij})\log\left(\frac{\hat{\epsilon}}{2}\right), \quad (4.42)$$

where $\hat{p}_{ij}$ is defined in Eq. (4.24) and $\hat{\eta}, \hat{\epsilon}$ are estimated via Eq. (4.32).

The joint decision rule is similarly derived by incorporating the anatomical constraints in Eq. (4.7) and Eq. (4.8).

4.3.3 Varying the Region Prior π^r

Although our framework enables us to estimate all unknown parameters, we further explore the solution space by specifying the expected number of diseased regions via the prior π^r. In particular, the evolution of disease foci across a range of prior π^r (in this work $\pi^r \in [0, 0.5]$) illustrates the stability of our model in explaining the data. Moreover, tuning π^r is an intuitive way to inject clinical knowledge into our framework and may be useful in certain applications. Fixing π^r does not affect the update equations in Section 4.2.

Table 4.2: Likelihood parameterizations used to generate synthetic data.

	μ_{-1}	μ_0	μ_1	σ^2_{-1}	σ^2_0	σ^2_1
`Good Data`	-0.35	0	0.35	0.05	0.05	0.05
`Noisy Data`	-0.18	0	0.36	0.050	0.058	0.072

	ρ_0	ρ_1	χ_0	χ_1	ξ^2_0	ξ^2_1
`Good Data`	0.7	0.10	0.45	0.35	0.005	0.005
`Noisy Data`	0.67	0.10	0.41	0.34	0.005	0.0026

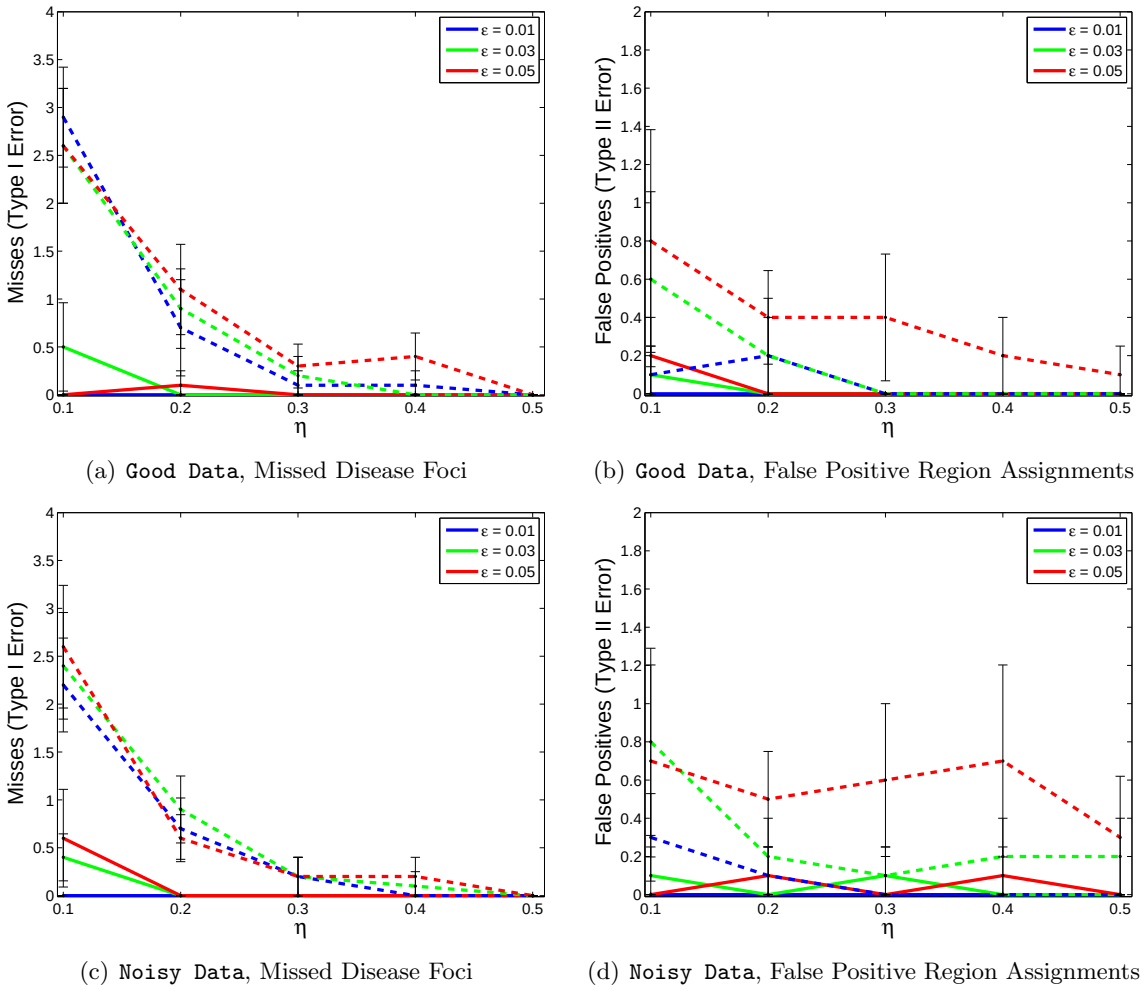
4.4 Experimental Results - Synthetic Data

We first evaluate the robustness and sensitivity of our algorithms using synthetic data. Our primary focus is the effect of the parameters η and ϵ on identifying the disease foci. We expect the performance to improve with increasing η and worsen with increasing ϵ. This is because higher values of η raise the number of functional connectivity differences associated with each disease foci. Consequently, the algorithms can better detect these regions. In contrast, larger values of ϵ increase the number of functional differences involving healthy regions, which negatively impacts the final solution.

We sweep the parameter values across the ranges $\eta \in [0.1, 0.5]$ and $\epsilon \in [0, 0.05]$; for each (η, ϵ) pair, we generate the latent connectivity templates and observed data according to Figs. 4-1(b) and 4-2(b). We fit the data using the algorithms presented in Section 4.2 and compute the false-negative (Type I) and false-positive (Type II) errors based on the MAP estimate $\hat{R}_i \in \{0, 1\}$ for each region i.

We mimic the organization of our clinical dataset by specifying a template with 78 regions (39 per hemisphere) and with two disease foci in each hemisphere. Throughout this section, we fix the functional prior π^f to the value inferred from the clinical experiments. We also sample the latent anatomical connectivity A such that the intra- and inter-hemisphere statistics match those of our clinical data.

We consider two likelihood parameterizations for $\{\mu, \sigma^s, \rho, \chi, \xi^2\}$, as shown in Table 4.2. The `Good Data` parameterization assumes a clear separation between the data distributions for different latent connectivity values. In this case, we can accurately infer the connectivity templates $\{A, F, \bar{F}\}$, which are then used for the region assignments R. The `Noisy Data` parameterization uses the ML parameter estimates $\hat{\Theta}$ from the clinical experiments to generate the observed synthetic measurements. In this case, there is a significant overlap in the data distributions; hence, we observe the effects of noise on the latent connectivity and region assignments.



(a) `Good Data`, Missed Disease Foci

(b) `Good Data`, False Positive Region Assignments

(c) `Noisy Data`, Missed Disease Foci

(d) `Noisy Data`, False Positive Region Assignments

Figure 4-3: Average number of mis-labeled region assignments when sampling from the functional model. The solid lines correspond to fitting the functional model, and the dashed lines represent the joint model results. The error bars denote one standard deviation. Type I error corresponds to the number of disease foci that were missed by our algorithm. Type II error denotes the number of healthy regions that were incorrectly identified as diseased.

4.4.1 Sampling from the Functional Model

Given the region labels R, we sample the graph structure T, the latent functional templates $F, \bar{F}$ and the observed fMRI correlations $\{B, \bar{B}\}$ according to Eqs. (4.2-4.5). In order to fit the joint model, we independently generate the latent anatomical connectivity A and the observed DWI measures $\{D, \bar{D}\}$ via Eq. (4.6) and Eq. (4.9), respectively. We re-sample the latent connectivity templates and observed data 10 times to collect error statistics.

Fig. 4-3 depicts the error in determining the region labels across 10 samples of the latent connectivity templates and corresponding observed data. Unsurprisingly, the functional model achieves uniformly lower Type I and Type II error. This is because the functional model exploits all pairwise connectivity information when determining the region labels, whereas the joint model must rely on a random subset of connections, specified by A. Nonetheless, the detection accuracy of the joint model improves significantly for larger values of η. The parameter η controls the density of non-zero edges in the variable T. Hence, as η increases, we are more likely to observe functional connectivity differences along the randomly generated anatomical template A.

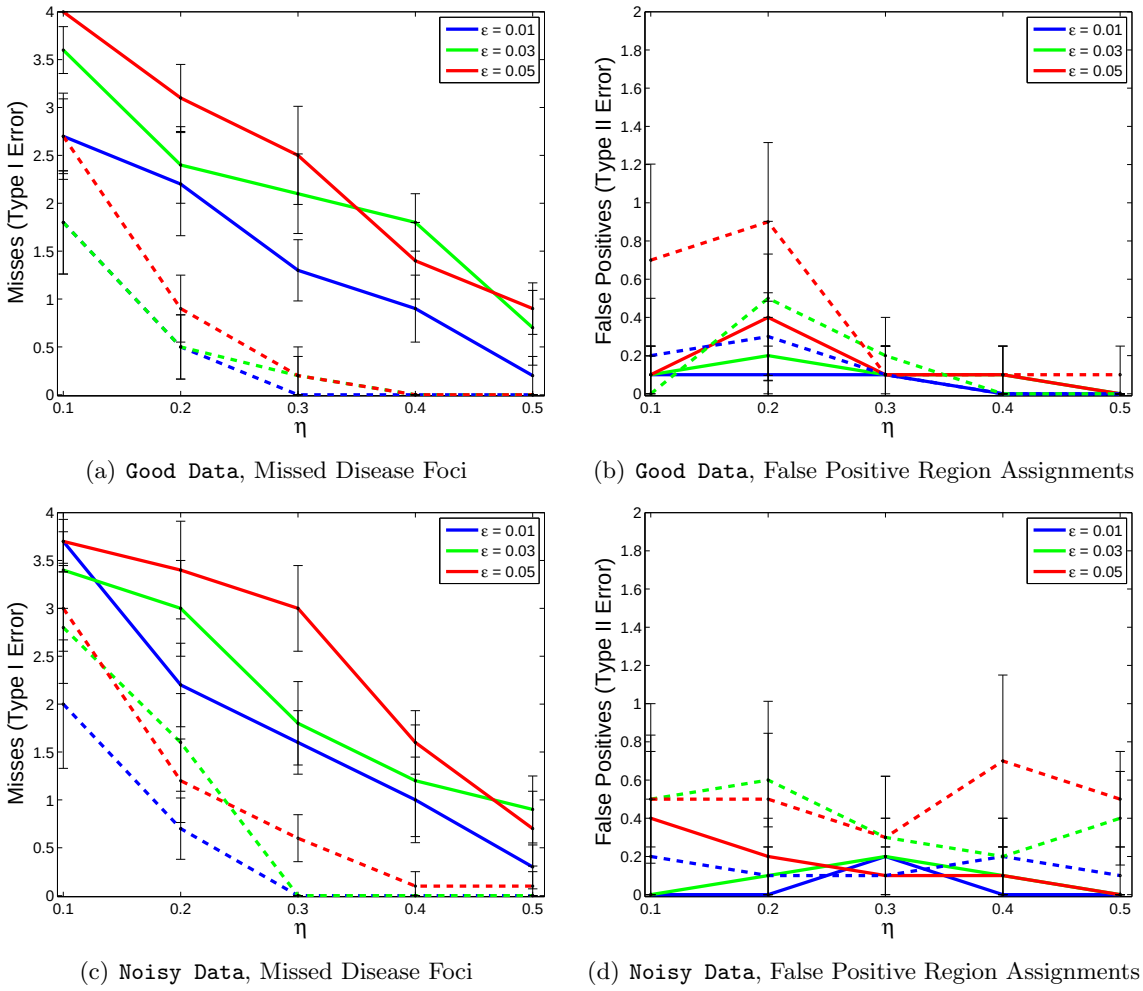
The parameter ϵ influences the rate of false-positive assignments, particularly for the joint model. Intuitively, higher values of ϵ produce a greater number of (spurious) functional connectivity differences involving healthy regions. Therefore, the algorithm is more likely to incorrectly label one of these regions as diseased.

Despite the large variability in Type II error in Fig. 4-3(b,d), on average less than two out of 74 healthy regions are labeled as disease foci. This behavior suggests an implicit regularization in our framework. Specifically, labeling a region as diseased benefits the free energy optimization by permitting the associated functional connections to differ between groups. However, connections to all other foci are automatically diseased, which may increase the free energy. Our algorithm balances these competing influences by identifying a sparse set of disease foci.

Finally, we observe that the error rates are similar for both the `Good Data` and the `Noisy Data` likelihood parameterizations. This indicates that errors in region assignments are primarily due to functional differences that are inconsistent with the underlying disease foci rather than to noisy data observations.

4.4.2 Sampling from the Joint Model

We now evaluate the model in a situation when the functional effects of a disease are restricted to direct anatomical pathways. Given the region labels R, we generate the control template F, the latent anatomical connectivity A and the graph structure T according to Fig. 4-2(b). However, we modify the construction of the clinical template $\bar{F}$. Since the joint model does not impose any correspondence between the values F_{ij} and $\bar{F}_{ij}$ in the absence of an anatomical connection, the latent templates differ dramatically when $A_{ij} = 0$. The functional model assumes all connections are equally important. Consequently, it cannot detect the true disease foci amid the overwhelming number of unrelated connectivity



(a) `Good Data`, Missed Disease Foci

(b) `Good Data`, False Positive Region Assignments

(c) `Noisy Data`, Missed Disease Foci

(d) `Noisy Data`, False Positive Region Assignments

Figure 4-4: Average number of mis-labeled region assignments when sampling from the joint model. The solid lines are obtained when fitting the functional model, and the dashed lines correspond to the joint model results. The error bars denote one standard deviation. Type I error corresponds to the number of disease foci that were missed by our algorithm. Type II error denotes the number of healthy regions that were incorrectly identified as diseased.

differences. For this reason, we sample $\bar{F}$ using Eq. (4.4), repeated below for convenience:

$$P(\bar{F}_{ij}|F_{ij},T_{ij};\epsilon) = \left[(1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\right]^{(1-T_{ij})}\left[\epsilon^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{1-\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\right]^{T_{ij}}. \quad (4.43)$$

Since $T_{ij} = 0$ if $A_{ij} = 0$, we omit the multinomial prior when there is no underlying anatomical connection. Instead, we encourage the latent functional connectivity templates to be the same in the control and clinical populations. Although not fully consistent with the

joint model, Eq. (4.43) allows us to fit the functional model with some degree of accuracy. The observed data $\{B, \bar{B}, D, \bar{D}\}$ is generated according to Eq. (4.5) and Eq. (4.9). We repeat the experiment 10 times to collect error statistics.

Fig. 4-4 illustrates the error in region assignments across 10 instantiations of the latent connectivity templates and observed data measures. Despite modifying the sampling procedure to accommodate the functional model, it exhibits significantly worse detection accuracy than the joint model for nearly all (η, ϵ) values. The performance reduction can be attributed to the anatomical constraint, which reduces the effective number of connections, and subsequently the number functional differences, associated with each region by 40-60%. Since the functional model treats all connections equally, the reduced number of functional differences is insufficient to pinpoint the disease foci. In contrast, the joint model adjusts the number of connectivity differences associated with a given region by the number of anatomical connections. Hence, the algorithm can isolate the diseased regions based on fewer differences. Despite the reduced detection performance, the functional model demonstrates lower Type II error. This suggests that it produces sparser estimates of the disease foci than the joint model.

We also observe similarities between our synthetic results in Figs. 4-3 and 4-4. As expected, the detection accuracy improves with increasing η, as it results in a greater number of functional differences associated with each diseased region. In addition, the Type II error variance is high, but on average, relatively few healthy regions are mis-labeled. Finally, the error rates are similar for both likelihood parameterizations. Once again, this suggests that noise in the latent structure has a greater impact than observation noise.

In summary, each model can robustly identify diseased regions if the data is sampled accordingly. In Fig. 4-3, the joint model exhibits slightly worse detection accuracy than the functional model; however, Fig. 4-4 reports a considerable drop in performance of the functional model when applied to the joint data. Both models exhibit an intrinsic regularization and infer sparse sets of foci with few false positive assignments.

4.5 Experimental Results - Clinical Data

Once again, we demonstrate our model on the clinical study of schizophrenia (see Section 2.7). We compute the DWI connectivity D^l_{ij} between regions i and j in subject l by averaging FA along all fibers that connect regions i and j. If no tracts are found, D^l_{ij} is set to zero. We extract the fMRI connectivity B^l_{ij} as the Pearson correlation coefficient between the mean time courses of regions i and j in subject l.

Table 4.3: Parameters of the functional model in Fig. 4-1(b) and the joint model in Fig. 4-2(b) estimated from the clinical data.

	π^r	π^f_{-1}	π^f_0	π^f_1	η	ϵ	μ_{-1}	μ_0	μ_1
Functional	0.039	0.33	0.46	0.21	0.16	0.030	-0.18	0	0.36
Joint	0.040	0.36	0.45	0.19	0.17	0.025	-0.18	0	0.36

	σ^2_{-1}	σ^2_0	σ^2_1	π^a	ρ_0	ρ_1	χ_0	χ_1	ξ^2_0	ξ^2_1
Functional	0.050	0.059	0.073	—	—	—	—	—	—	—
Joint	0.050	0.058	0.073	0.34	0.67	0.10	0.41	0.34	0.0050	0.0026

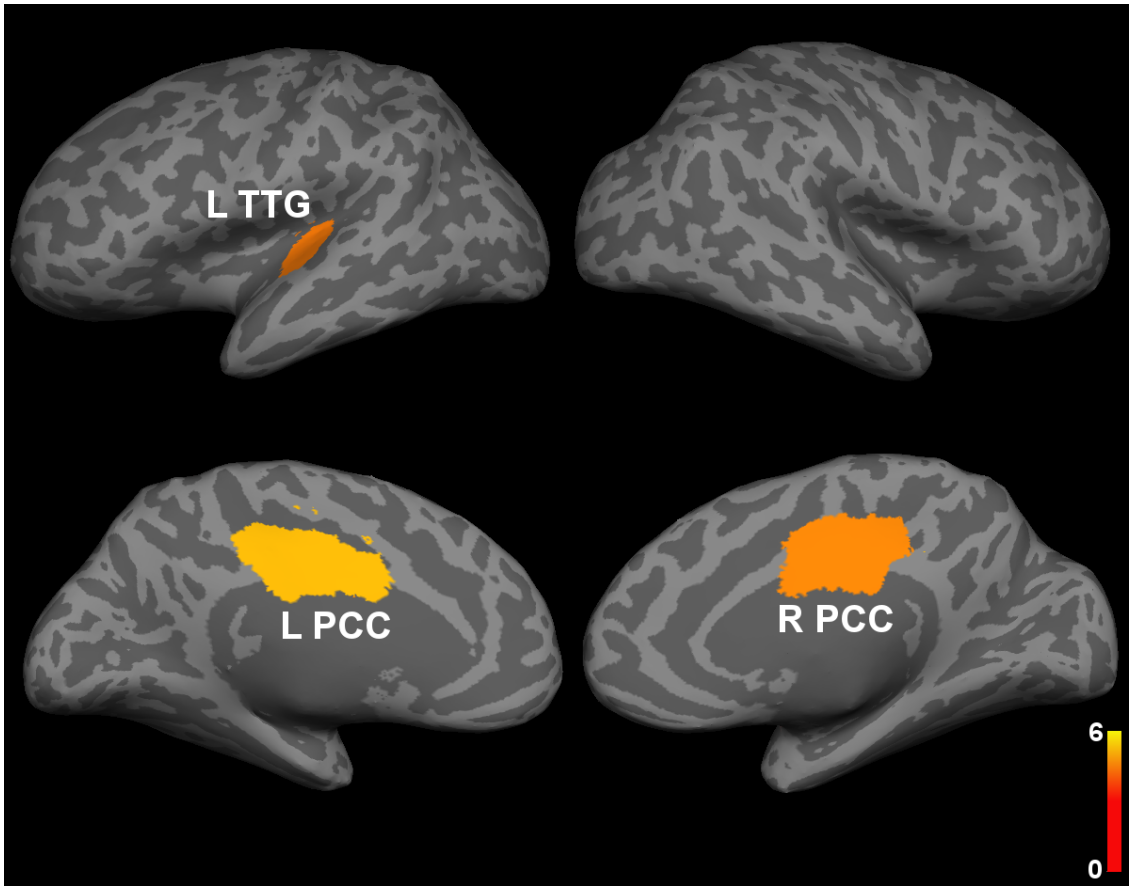


Figure 4-5: Significant regions based on permutation tests ($q_i > 0.5$, uncorrected $p < 0.021$) identified by the functional model. The colorbar corresponds to the negative log p-value. We present the lateral and medial viewpoints for each hemisphere. The highlighted regions are the posterior cingulate (L PCC & R PCC) and the transverse temporal gyrus (L TTG).

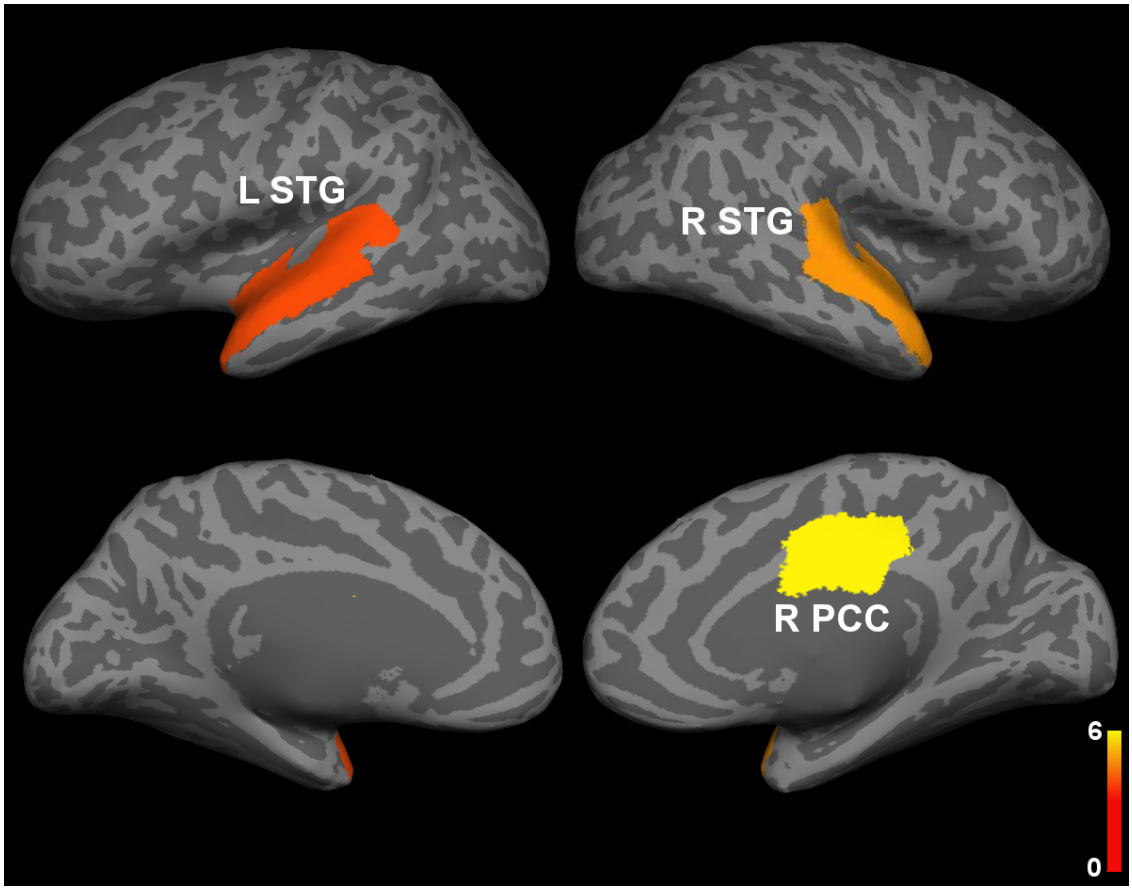
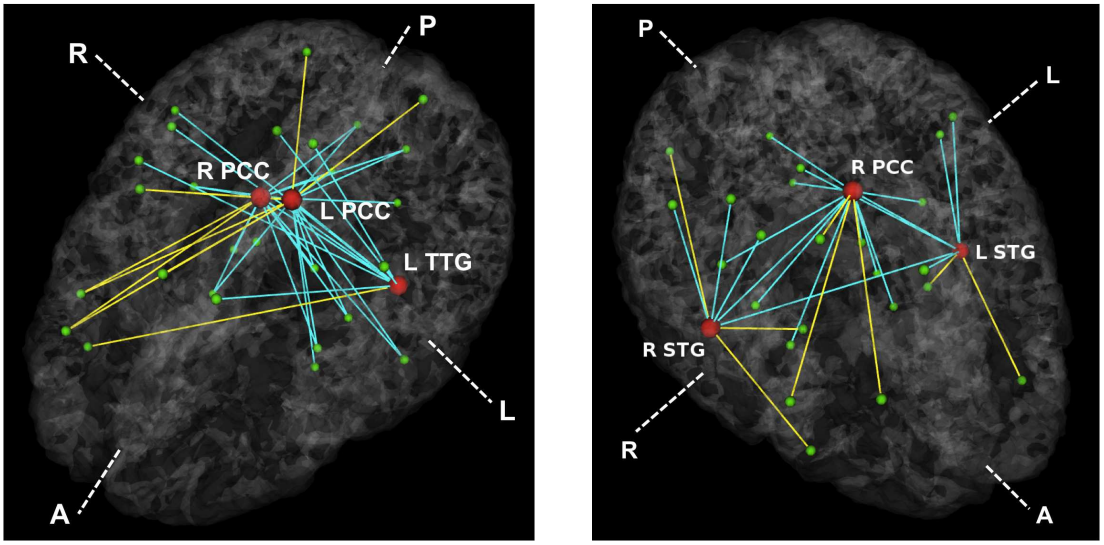


Figure 4-6: Significant regions based on permutation tests ($q_i > 0.5$, uncorrected $p < 0.044$) identified by the joint model. The colorbar corresponds to the negative log p-value. We present the lateral and medial viewpoints for each hemisphere. The highlighted regions are the posterior cingulate (R PCC) and the superior temporal gyrus (L STG & R STG).

4.5.1 Significant Regions

Fig. 4-5 and Fig. 4-6 illustrate the detected disease foci ($\hat{q}_i > 0.5$) for the functional and joint models, respectively. We color each region according to $-\log$(p-value) such that red corresponds to low significance and yellow indicates high significance. Each method identified three disease foci, all of which are significant. The functional model implicated the left posterior cingulate ($\hat{q}_i = 1, p < 0.008$), the right posterior cingulate ($\hat{q}_i = 1, p < 0.017$) and the left transverse temporal gyrus (Heschl's gyrus) ($\hat{q}_i = 1, p < 0.021$). The joint mode implicates a different subset of regions, namely, the right posterior cingulate ($\hat{q}_i = 1, p < 0.004$), the right superior temporal gyrus ($\hat{q}_i = 1, p < 0.014$), and the left superior temporal gyrus ($\hat{q}_i = 1, p < 0.044$).



(a) Functional Model (b) Joint Model

Figure 4-7: Estimated graph of functional connectivity differences. The red nodes indicate the disease foci. Blue lines indicate reduced functional connectivity and yellow lines indicate increased functional connectivity in the schizophrenia population.

Both models identify significant foci in the default network and in the temporal lobes of the brain. All regions are among the 8 expert selected brain structures in Chapter 3. Interestingly, we observe symmetry in region assignments across the hemispheres, as evident for the posterior cingulate (PCC) and the superior temporal gyri (STG). This phenomenon may arise from the well-documented symmetry found in resting-state fMRI correlations [90]. We analyze the differences between Fig. 4-5 and Fig. 4-6 in Section 4.6.

Table 4.3 reports the parameters inferred by our algorithms. We notice that the fMRI likelihood parameters are almost identical for both algorithms. This suggests that the difference between the regions in Fig. 4-5 and Fig. 4-6 is driven by the hierarchical structure from connections to region assignments rather than by the inference of latent functional connectivity from the data. Additionally, we observe consistency in parameter estimates across random subject re-labelings in the permutation procedure (not shown). This implies that the main effects of permuting the subject diagnoses are reflected in the latent assignments rather than in the data likelihood.

4.5.2 Differences in Functional Connectivity

Fig. 4-7 displays the estimated graph of anomalous functional connectivity for each model. The functional model identifies abnormal connections distributed throughout the brain.

For the joint model, abnormalities that originate in the posterior cingulate project to the midbrain and frontal lobe, whereas abnormalities stemming from the right and left superior temporal gyri tend to span their respective hemispheres. This difference in organization is explained by the constraint in Fig. 4-2(a) that functional connectivity differences occur along anatomical pathways.

Both models detect an overall reduction in functional connectivity for schizophrenia patients. Of notable exception are connections to the frontal lobe. This phenomenon has been reported in prior studies of schizophrenia [37] and is believed to interfere with perception by misdirecting attentional resources.

4.5.3 Effect of Region Prior

Fig. 4-8 and Fig. 4-9 illustrate the results of varying the prior π^r of the region indicator vector R for the functional and joint models, respectively. We color each of the selected regions according to the smallest value of π^r such that the marginal posterior of the region i being a focus is greater than 0.2 (i.e, $\hat{q}_i > 0.2$). The yellow regions are always identified as foci, whereas the orange and red regions are only selected for larger prior values.

We observe that the functional model identifies a stable set of disease foci with an additional region for large values of π^r. In contrast, the sets of affected regions in the joint model form a nested substructure as π^r increases. It suggests an initial set of disease foci, identical to the significant regions in Fig. 4-6. For increasing π^r, the algorithm progressively includes regions that exhibit some functional abnormalities but are not as strongly implicated by the data. This extended set of regions is a superset of those identified by the functional model. We elaborate on the differences between Fig. 4-8 and Fig. 4-9 in the following section.

4.6 Discussion

We present a unified approach to infer regions associated with a disorder based on population differences in connectivity. Our first model operates on the complete graph of pairwise functional connections. Our second model incorporates anatomical constraints into this basic framework. We derive a variational EM algorithm for maximum likelihood estimation of the model parameters. The algorithm simultaneously infers the posterior distribution over the region labels and the set of abnormal functional connections.

Fig. 4-5 and Fig. 4-6 presents the diseased regions implicated by each model. The main difference between the two results is that the functional model labels the transverse temporal gyrus as a disease focus, whereas the joint model pinpoints the superior temporal gyrus as relevant for schizophrenia. This discrepancy is partially explained by the size

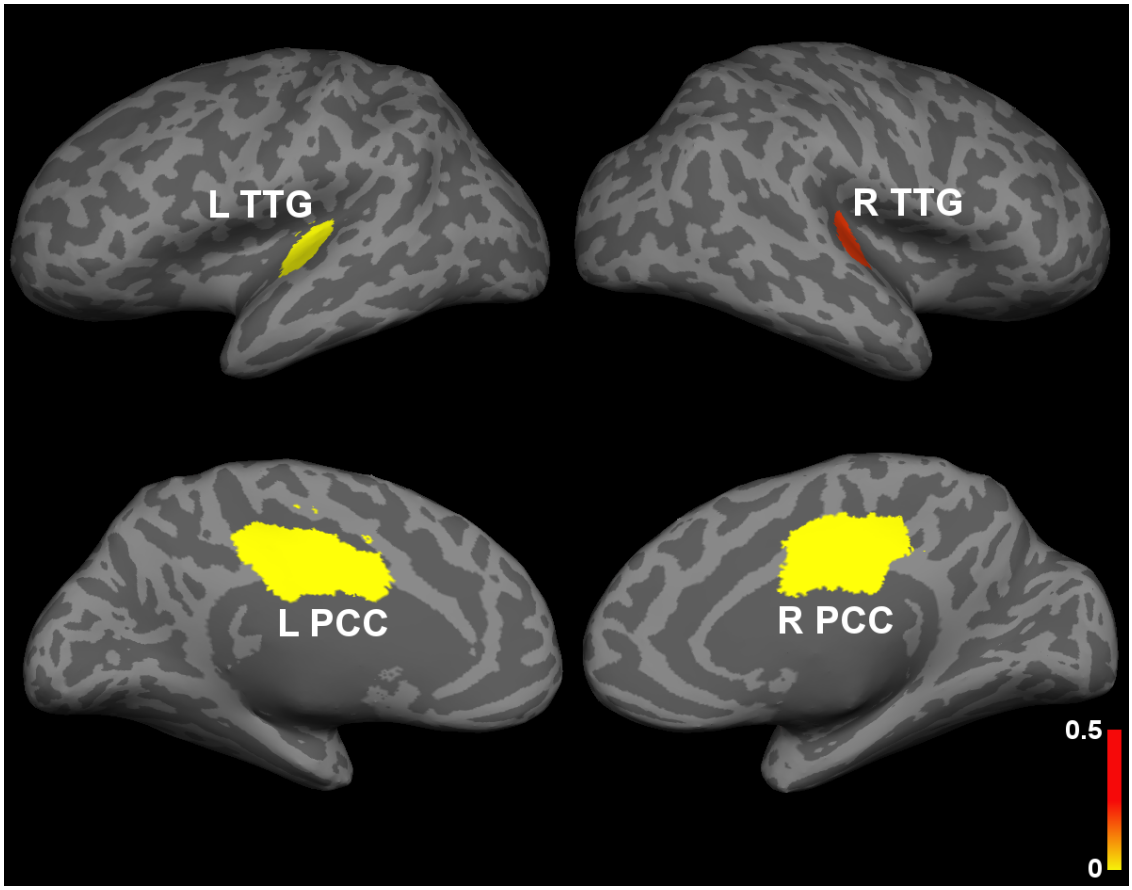


Figure 4-8: Evolution of the disease foci when varying the region prior π^r for the functional model. The colorbar corresponds to the smallest value of π^r such that $q_i > 0.2$. The highlighted regions correspond to the posterior cingulate (L PCC & R PCC) and the transverse temporal gyrus (L TTG & R TTG).

difference between these regions. As seen in Fig. 4-10, we identify significantly more white matter tracts involving the (large) STG than for the TTG. Hence, we are more likely to detect functional abnormalities associated with the STG that occur along direct anatomical pathways. This is reflected in Fig. 4-7(a). The majority of abnormal functional connections emanating from the TTG are inter-hemispheric, and hence, do not coincide with latent anatomical connections. Fig. 4-10 suggests that the quality of the joint model is largely dependent on the detection power of tractography. This underscores the need for advanced tractography algorithms that reliably identify long-range connections.

The TTG, also known as Heschl Gyrus, plays crucial role in auditory perception and language processing. Its volume reductions, especially on the left, have been long asso-

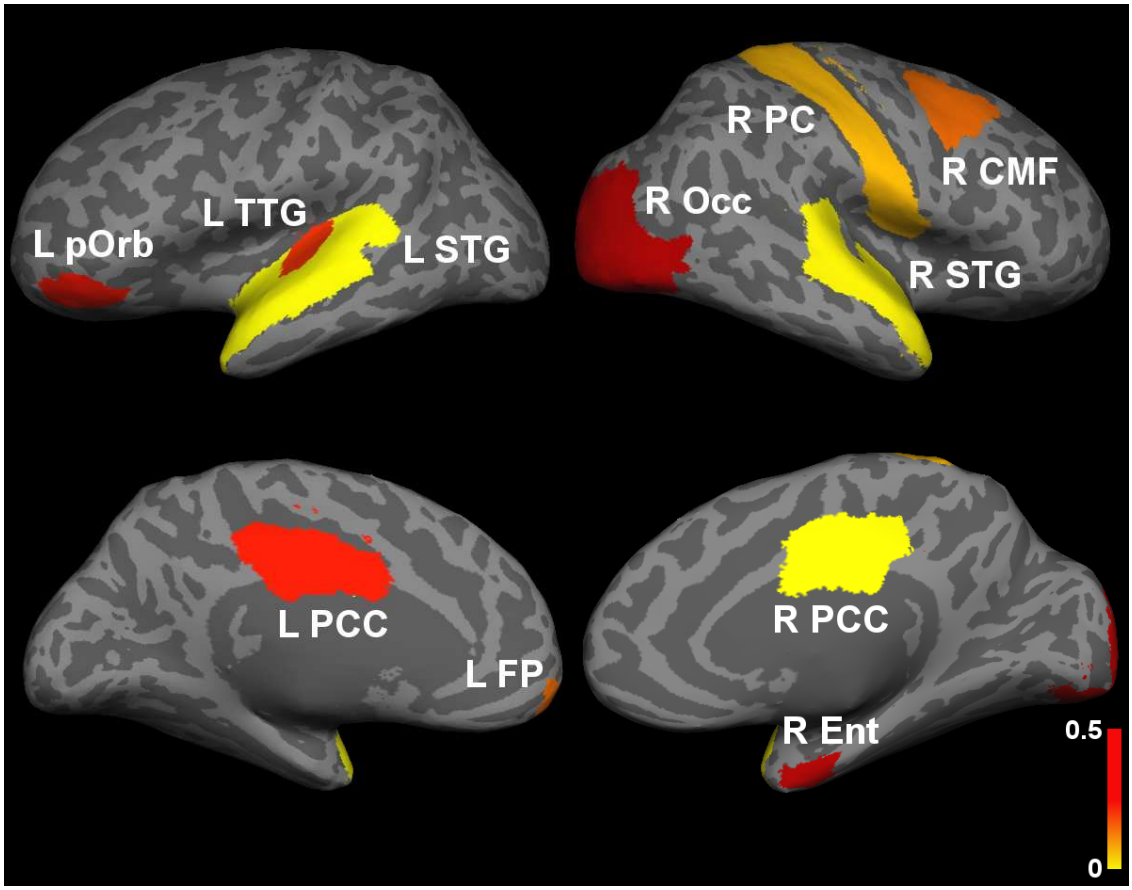
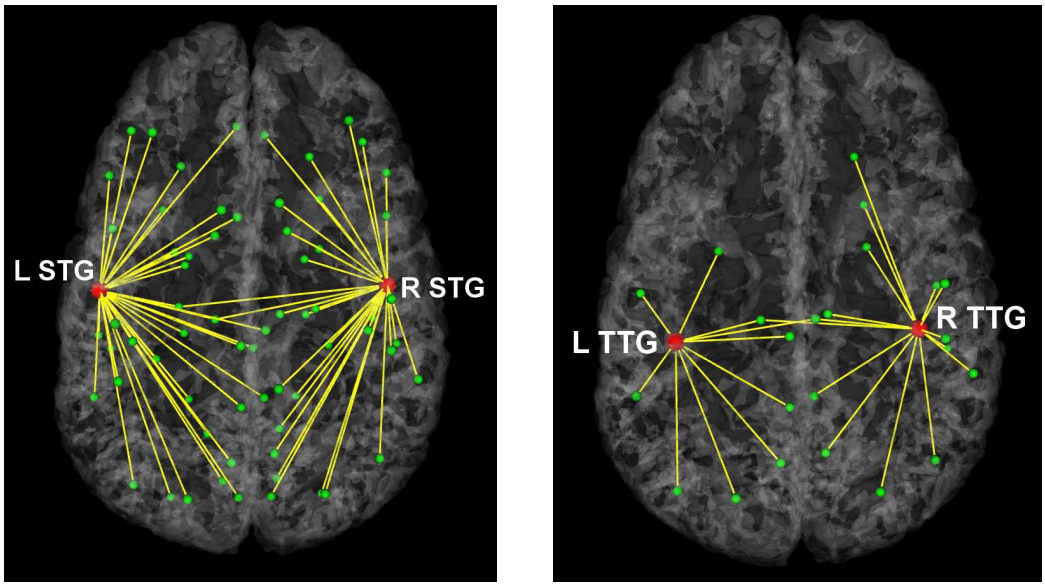


Figure 4-9: Evolution of the disease foci when varying the region prior π^r for the joint model. The colorbar corresponds to the smallest value of π^r such that $q_i > 0.2$. The highlighted regions correspond to the posterior cingulate (L PCC & R PCC), the superior temporal gyrus (L STG & R STG), the postcentral gyrus (R PC), the frontal pole (L FP), the caudal middle frontal gyrus (R CMF), the transverse temporal gyrus (L TTG), the pars orbitalis (L pOrb), the entorhinal cortex (R Ent) and the lateral occipital cortex (R LOcc).

ciated with hallmark schizophrenia symptoms, such as auditory hallucinations, delusions and thought disorder [81]. Heschl's gyrus has also been linked to disease progression [59], suggesting its crucial role in schizophrenia pathophysiology.

The STG connects with heteromodal neocortical regions and temporolimbic areas. Electrophysiology and PET/fMRI studies in humans emphasized the STG's role in the interpretation, production and self-monitoring of language. There is also evidence for structural and functional abnormalities of the STG in schizophrenia, which may be associated with formal thought disorder and auditory hallucinations [64, 78].

The PCC is one of the key structures in the default mode network. Recent functional



(a) Superior Temporal Gyri (b) Transverse Temporal Gyri

Figure 4-10: Latent anatomical connections estimated by the joint model.

schizophrenia studies [38] reported altered temporal frequency and spatial location of the default mode network. This suggests that the default network may be under- or overmodulated by key regions, including the anterior and the posterior cingulate cortex. Our results confirm this hypothesis, further illustrating how such modulation can affect functional connectivity (decreased connectivity between PCC and posterior parietal and temporal regions and increased connectivity between PCC and occipital and frontal lobes in Fig. 4-7). Reduced connectivity in the posterior cingulate has been shown to correlate with both positive and negative symptoms of schizophrenia [9].

The role of anatomy is also evident in the graphs of aberrant functional connectivity depicted in Fig. 4-7. The functional results are distributed across the brain with little high-level organization. In contrast, the connections identified by the joint model are largely separated by hemisphere and seem consistent with estimated white matter tracts. Despite their differences, both models detect a similar global pattern, which may reveal underlying neurological changes induced by schizophrenia. Specifically, we observe increased functional connectivity to the frontal lobe and reduced functional connectivity between the parietal/posterior cingulate region and the temporal lobe in the clinical population.

Increased connectivity between the default network and the medial frontal lobe, both at rest and during task, has been reported in schizophrenia [37, 104]. It is believed to interfere with perception of the external world by misdirecting attentional resources. Interestingly,

decreased connectivity within the default network has been described as well [9, 103]. The later study reported decreased functional connectivity between the posterior cingulate gyrus and the hippocampus, which is consistent with our findings. The relationship between disruptions in functional connectivity and the integrity of the fornix has also been suggested. Along with prior findings, our results suggest an inverse relationship between connectivity in the temporal and frontal parts of the default network. Such "anticorrelations" have been previously described between the default and task-related networks. Two connections along white matter tracts in Fig. 4-7(b) have been implicated in schizophrenia [73]: the connection between the left and the right STG, provided by corpus callosum, and the connection between posterior and anterior CG, provided by cingulum bundle. These two white matter tracts suggest a direct, causative relationship between anatomical and functional connectivity disruptions in schizophrenia.

Tuning the region prior parameter π^r enables us to explore the solution space. Once again, we observe differences between the two models. The functional results are consistent across a large range of prior values. In contrast, the joint model localizes nested subsets of disease foci as π^r increases. This suggests that the anatomical constraint increases the sensitivity of the joint model. Specifically, the effective number of connections to each region is reduced to the number of direct anatomical pathways. Hence, the joint model selects diseased regions based on fewer functional connectivity differences. Since many regions are weakly implicated by the data (i.e., associated with a few abnormal connections), biasing the algorithm through the region prior π^r causes them to be selected as foci.

The question remains: which model should we use? Presently, there is no standard technique to integrate anatomical and functional connectivity in order to pinpoint region impairments. Therefore, we argue that this is a largely a philosophical issue based on a set of assumptions one makes about the brain. This work presents two different viewpoints. Clearly, if we assume that impairments of a neurological disorder equally affect functional synchrony between any two brain regions, then Fig. 4-3 suggests that we should fit the functional model. Similarly, if we assume that the most salient effects of a disorder occur along direct anatomical connections, then Fig. 4-4 encourages us to choose the joint model. If we are unsure, then our synthetic results suggest that, on average, we are better off using the joint model. This is because the joint model achieves higher detection accuracy on data sampled from the functional model than *vice versa*. In the absence of latent anatomical connectivity, the joint model compares aggregate statistics of the templates F and $\bar{F}$. Therefore, data sampled according to the functional model in Fig. 4-1(b) is fairly consistent with the assumptions of the joint model. In contrast, the functional model cannot

accurately be fit to data sampled from the joint model in Fig. 4-2(b). A future extension of this work may consider all two-stage anatomical pathways as being relevant for disease localization. This can be achieved by incorporating the pairwise terms $A_{ik}A_{kj}$ into the distribution for the functional template $\bar{F}$ of the clinical population in Eq. (4.8).

Encouragingly, both models in our current formulations localize similar disease foci. In fact, the joint model recovers both posterior cingulate regions as well as the left transverse temporal gyrus when we vary the region prior parameter π^r. The increased sensitivity of the joint model may prove beneficial, as it identifies a larger set of candidate regions (Fig. 4-9). The effects of a complex disorder like schizophrenia are often subtle. Hence, the functional model, which only identifies the strongest connectivity differences, may not locate all relevant disease foci.

Our results may also be influenced by our selection of regions. If the regions are too small, then the variability in DWI tractography across subjects makes it difficult to infer the template anatomical connectivity and group-level parameters [48]. However, larger regions smooth out important functional connectivity information. In this work, we rely on Brodmann regions identified by Freesurfer [29]. Brodmann areas provide anatomically meaningful correspondences across subjects that roughly correspond to functional divisions within the brain. Moreover, these regions are large enough to ensure stable tractography results. We emphasize that our framework applies readily to any set of ROIs that are defined consistently across subjects.

Finally, it is worth noting that our current framework may not apply to all degenerative conditions. Specifically, if the abnormal connectivity does not localize to a sparse subset of regions (i.e., aging and Alzheimer's), then region-based models are not appropriate. Rather, one should use our connection-based model in Chapter 3 for these applications. In future, we can design statistical tests and rely on clinical knowledge to determine the correct model.

Nonetheless, the generative models presented in this chapter provide a starting point for relating connections and regions. However, there is plenty of room for improvement. For example, our joint model considers only direct anatomical connections and places a binary constraint on the graph of functional aberrations; our functional model ignores all anatomical information. In reality, the interaction between anatomy and function is likely a hybrid of these competing viewpoints. Additionally, we assume a single set of disease foci that share mutually abnormal connectivity. However, neurological disorders can arise from several impairments in the brain that do not directly interact.

Similar to our joint model in Chapter 3, these choices are intentional. Despite advancements in the field, the effects of schizophrenia (and other clinical disorders) on brain

connectivity is neither well understood nor well characterized. Therefore, we formulate a simple relationship between region assignments and latent connectivity. Furthermore, given the potentially large amounts of inter-subject variability and external noise, we formulate a small set of model parameters to avoid over-fitting. Once again, these limitations provide ample opportunities for future work.

Chapter 5

Extension to Multi-class Networks

Chapter 4 describes a generative framework that integrates population differences in functional connectivity to isolate foci of a neurological disorder. The model assumes a single collection of foci and only specifies the interaction between healthy and diseased nodes. This binary partition of regions cannot account for multiple sub-networks that together are responsible for the abnormal connectivity patterns in the brain.

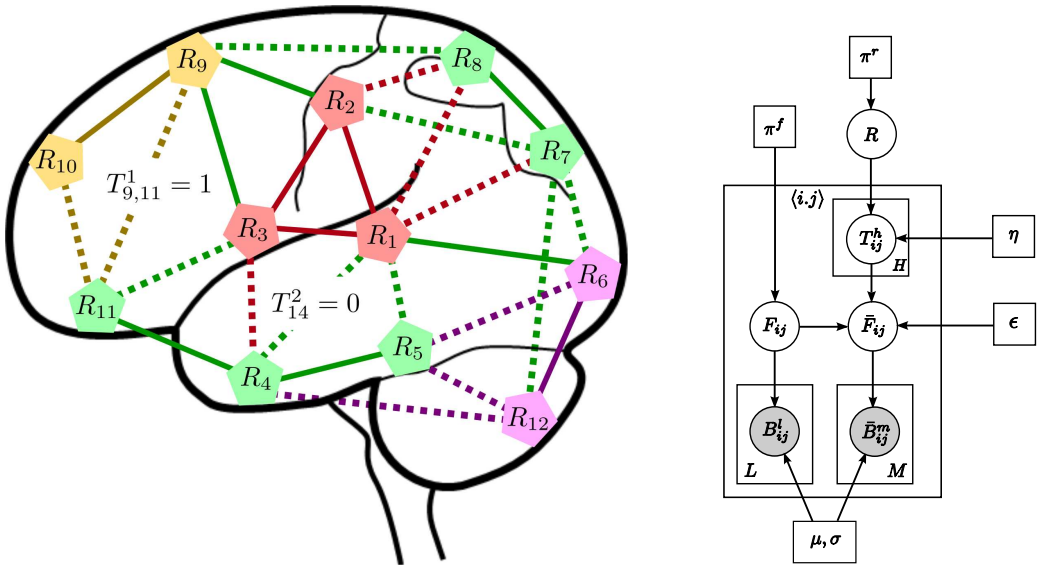
In this chapter we extend the formulation to account for several *disease clusters*. Here, each cluster is responsible for an independent network of functional differences. This is accomplished via a straightforward modification to the probabilistic model in Chapter 4. The result is analogous to a clustering problem where each region i is assigned a multi-class label that designates it either as healthy or as belonging to a particular disease cluster h. We rely on a similar variational EM algorithm to fit the model to the observed data. Our algorithm jointly infers the disease clusters and the induced connectivity differences.

In this exploratory chapter, we focus on the complete graph of pairwise functional differences and do not incorporate anatomy. We presents results on both synthetic and real-world data.

The remainder of this chapter is organized as follows. We modify the latent variables for the functional model in Section 5.1 and derive the corresponding variational algorithm in Section 5.2. We discuss the non-trivial implementation details in Section 5.3. Section 5.4 and Section 5.5 report experimental results based on synthetic and clinical data, respectively. Finally, Section 5.6 discusses the results and the relevance of our model.

5.1 Generative Model

We assume that the effects of a disorder can be localized to a small subset of brain regions. These regions are partitioned into H groups, which we call *disease clusters*; each cluster is responsible for a sub-network of abnormal neural activity. Fig. 5-1 presents the latent network structure in the brain and the corresponding graphical model. The nodes in Fig. 5-1(a) correspond to regions in the brain, and the edges denote pairwise functional connections



(a) Network Model of Brain Connectivity (b) Graphical Model

Figure 5-1: (a) Latent organization of the disorder. The nodes correspond to regions in the brain, and the lines denote pairwise functional connections. Only a subset of edges is shown; the model is defined on the full graph of pairwise connections. The green nodes and edges correspond to the healthy regions and connections, respectively. The red, yellow and purple nodes represent three clusters of disease foci, and the colored edges specify pathways of abnormal functional connectivity. The solid lines are deterministic given the region labels; the dashed lines are probabilistic. (b) The corresponding graphical model. R specifies disease cluster labels. F_{ij} denotes the latent functional connectivity between regions i and j. B^l_{ij} is the observed fMRI measurements in the l^{th} subject. Variables associated with the clinical population are identified by an overbar.

between them. The green nodes/edges are healthy. The yellow, red and purple nodes/edges represent three disease clusters and the associated abnormal functional pathways.

Based on the region assignments, we construct binary graphs $\{T^h\}_{h=1}^H$ of aberrant functional connectivity using four simple rules: (1) a connection between regions i and j in disease cluster h is always abnormal ($T^h_{ij} = 1$, solid yellow/red/purple lines in Fig. 5-1(a)), (2) a connection between region i in disease cluster h and region j in disease cluster $h' \neq h$ is never abnormal ($T^h_{ij} = 0$), (3) a connection between two healthy regions is never abnormal, and (4) a connection between a healthy region i and a region j in disease cluster h is abnormal with probability η (dashed lines). These are similar to the rules proposed in Section 4.1, except that we now specify the interaction between nodes in different clusters.

Once again, the latent functional connectivity variables F_{ij} and $\bar{F}_{ij}$ model neural synchrony between the regions. Ideally, $F_{ij} \neq \bar{F}_{ij}$ for connections altered by any disease cluster,

and $F_{ij} = \bar{F}_{ij}$ otherwise. However, due to noise and inter-subject variability, we assume the templates can differ from the graphs $\{T^h\}$ with probability ϵ. The fMRI correlations B^l_{ij} are noisy observations of the latent network.

Region Labels Given N brain regions, the random variable R_i is a multi-class label for the state of region i ($i = 1, \ldots, N$). R_i is generated from the multinomial prior π^r. For convenience, we model R_i as a length-$(H+1)$ indicator vector such exactly that one of its elements $[R_{i0}\, R_{i1} \ldots R_{iH}]$ equals to one:

$$P(R_i; \pi^r) = \prod_{h=0}^{H} (\pi^r_h)^{R_{ih}}. \tag{5.1}$$

The state $R_i = 0$ ($R_{i0} = 1$) corresponds to region i being healthy.

Graphs of Abnormal Connectivity The graph T^h captures the abnormal functional connectivity associated with disease cluster h. Formally, we have

$$P(T^h_{ij} | R_i, R_j; \eta) = \begin{cases} \delta(T^h_{ij}), & R_i \neq h, R_j \neq h, \\ \delta(T^h_{ij}), & (R_i = h, R_j = h') \text{ or } (R_i = h', R_j = h), \\ 1 - \delta(T^h_{ij}), & R_i = R_j = h, \\ \eta^{T^h_{ij}} (1-\eta)^{1-T^h_{ij}}, & (R_i = h, R_j = 0) \text{ or } (R_i = 0, R_j = h), \end{cases} \tag{5.2}$$

where each edge T^h_{ij} is generated independently given the region labels R. The first and second conditions in Eq. (5.2) impose a strict separation between graphs $\{T^h\}$. Namely, the abnormal connections associated with different disease clusters do not overlap. This avoids an identifiability problem in which we cannot determine the cluster responsible for a particular functional difference.

Latent Functional Connectivity The latent functional connectivity F_{ij} of the control population is unchanged from Eq. (4.3).

The latent functional connectivity $\bar{F}_{ij}$ of the clinical population is based on F_{ij} and the graphs $\{T^h\}$. Let T denote the aggregate graph structure, which we construct as the union $T = \sum_{h=1}^{H} T^h$. Specifically, the aggregate edge $\langle i, j \rangle$ is abnormal ($T_{ij} = 1$) if and only if there exists $h \in 1, \ldots, H$ such that the individual graph edge T^h_{ij} is abnormal, i.e., $T^h_{ij} = 1$.

Given T, the conditional distribution of the clinical template $\bar{F}_{ij}$ is identical to Eq. (4.4).

We repeat the expression below for convenience:

$$P(\bar{F}_{ij}|F_{ij},T_{ij};\epsilon)=\left[(1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\right]^{1-T_{ij}}\left[\epsilon^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{1-\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\right]^{T_{ij}}. \quad (5.3)$$

The observed fMRI correlations are noisy measurements of latent functional connectivity. The likelihood distributions are the same as in Section 4.1.

5.2 Variational Inference

Our primary quantity of interest is the posterior probability of the region labels R. Hence, we marginalize out the graphs $\{T^h\}$ to simplify the relationship between R and the observed data. This implies the following conditional distribution of the clinical template $\bar{F}_{ij}$:

$$P(\bar{F}_{ij}|F_{ij},R_i,R_j;\eta,\epsilon)=\begin{cases}(1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & R_i=R_j=0,\\ (1-\epsilon)^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & R_i\neq R_j, R_i\neq 0, R_j\neq 0,\\ \epsilon^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{1-\epsilon}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & R_i=R_j\neq 0,\\ \epsilon_1^{F_{ij}^{\mathrm{T}}\bar{F}_{ij}}\left(\frac{1-\epsilon_1}{2}\right)^{1-F_{ij}^{\mathrm{T}}\bar{F}_{ij}}, & (R_i=0,R_j\neq 0)\ or\ (R_i\neq 0,R_j=0),\end{cases} \quad (5.4)$$

where $\epsilon_1=\eta\epsilon+(1-\eta)(1-\epsilon)$ as in Section 4.2. Regardless of the number of clusters H, the conditions in Eq. (5.4) depend on simple relationships between the regions. Namely, is either region healthy? If not, do they belong to the same or to different disease clusters? These relationships are derived by aggregating the conditional distributions of $\{T^h\}$ in Eq. (5.2).

Let $Y=\{B,\bar{B}\}$ and $\Theta=\{\pi,\eta,\epsilon,\mu,\sigma^2\}$ denote the observed fMRI measurements and the set of model parameters, respectively. We employ a maximum likelihood (ML) framework to fit the model to the data. Once again, the coupling induced by R forces us to adopt a variational approximation [57] for the posterior probability distribution of the latent variables given the observed data:

$$Q(R,F,\bar{F})=Q^r(R)\cdot Q^c(F,\bar{F})=Q^r(R)\prod_{<i,j>}Q_{ij}^c(F_{ij},\bar{F}_{ij}), \quad (5.5)$$

where $Q^r(\cdot)$ is a joint distribution over the N length-$(H+1)$ multinomial vectors $\{R_i\}$ and $Q_{ij}^c(\cdot)$ is an 9-state multinomial distribution corresponding to all configurations of latent functional connectivity. Again, this factorization yields a tractable inference algorithm while preserving the dependency between F_{ij}, and $\bar{F}_{ij}$ given the region indicator vector R.

Similar to Section 4.2, we employ a variational EM formulation [23] to obtain the pos-

terior distribution $Q(\cdot)$ and model parameters Θ that minimize the variational free energy

$$\mathcal{FE} = -E_Q\left[\log P(R, F, \bar{F}, Y; \Theta)\right] - \mathcal{H}(Q), \tag{5.6}$$

where the joint log-likelihood of all hidden and observed variables is given by:

$$\begin{aligned}
\log P(R, F, \bar{F}, Y; \Theta) = &\sum_{i=1}^{N}\sum_{h=0}^{H} R_{ih}\log(\pi_s^r) + \sum_{\langle i,j\rangle}\sum_{k=-1}^{1} F_{ijk}\log\left(\pi_k^f\right) \\
&+ \sum_{\langle i,j\rangle}(R_{i0}R_{j0})\left(F_{ij}^{\mathrm{T}}\bar{F}_{ij}\log(1-\epsilon) + (1-F_{ij}^{\mathrm{T}}\bar{F}_{ij})\log\left(\frac{\epsilon}{2}\right)\right) \\
&+ \sum_{\langle i,j\rangle}\left((1-R_{i0})(1-R_{j0}) - \sum_{h=1}^{H} R_{ih}R_{js}\right)\left(F_{ij}^{\mathrm{T}}\bar{F}_{ij}\log(1-\epsilon) + (1-F_{ij}^{\mathrm{T}}\bar{F}_{ij})\log\left(\frac{\epsilon}{2}\right)\right) \\
&+ \sum_{\langle i,j\rangle}\left(\sum_{h=1}^{H} R_{ih}R_{jh}\right)\left(F_{ij}^{\mathrm{T}}\bar{F}_{ij}\log(\epsilon) + (1-F_{ij}^{\mathrm{T}}\bar{F}_{ij})\log\left(\frac{1-\epsilon}{2}\right)\right) \\
&+ \sum_{\langle i,j\rangle}(R_{i0}(1-R_{j0}) + (1-R_{i0})R_{j0})\left(F_{ij}^{\mathrm{T}}\bar{F}_{ij}\log(\epsilon_1) + (1-F_{ij}^{\mathrm{T}}\bar{F}_{ij})\log\left(\frac{1-\epsilon_1}{2}\right)\right) \\
&+ \sum_{\langle i,j\rangle}\sum_{k=-1}^{1}\left[F_{ijk}\sum_{l=1}^{L}\log\mathcal{N}\left(B_{ij}^l;\mu_k,\sigma_k^2\right) + \bar{F}_{ijk}\sum_{m=1}^{M}\log\mathcal{N}\left(\bar{B}_{ij}^m;\mu_k,\sigma_k^2\right)\right].
\end{aligned} \tag{5.7}$$

The product terms involving R are based on the conditions in Eq. (5.4).

E-Step: For a fixed setting of model parameters $\hat{\Theta}$, we showed in Section 4.2 that the variational posterior $\hat{Q}(\cdot)$ that minimizes Eq. 5.6 satisfies the following fixed point equations:

$$\hat{Q}^c(F, \bar{F}) = \tilde{P}(F, \bar{F}; \hat{\Theta}) \propto \exp\left\{E_{\hat{Q}^r}\left[\log P(R, F, \bar{F}, Y; \hat{\Theta})\right]\right\}, \tag{5.8}$$

$$\hat{Q}^r(R) = \tilde{P}(R; \hat{\Theta}) \propto \exp\left\{E_{\hat{Q}^c}\left[\log P(R, F, \bar{F}, Y; \hat{\Theta})\right]\right\}. \tag{5.9}$$

In the E-step, we alternatively update $\hat{Q}^r(R)$ and $\hat{Q}^c(F, \bar{F})$ until convergence. To update $\hat{Q}^c(\cdot)$, we evaluate the right-hand side of Eq. (5.8) for each of the nine configurations $F_{ij} = k$, $\bar{F}_{ij} = k'$ $(k, k' \in \{-1, 0, 1\})$ and normalize to obtain a valid probability distribution. Once again, the right-hand side of Eq. (5.9) can be expressed in terms of $E_{\hat{Q}^c}[F_{ij}^{\mathrm{T}}\bar{F}_{ij}]$. Since F_{ij} and $\bar{F}_{ij}$ are indicator variables, this quantity can be evaluated as

$$\hat{p}_{ij} \triangleq E_{\hat{Q}^c}[F_{ij}^{\mathrm{T}}\bar{F}_{ij}] = \sum_{k=-1}^{1}\hat{Q}_{ij}^c(F_{ij} = k, \bar{F}_{ij} = k). \tag{5.10}$$

Eq. (5.10) represents the posterior probability that the latent connectivity of edge $\langle i, j \rangle$ is the same for both populations.

We use Gibbs sampling to obtain samples $\mathcal{S} = \{R^s\}$ from $\hat{Q}^r(R)$. Based on the joint log-likelihood in Eq. (5.7), the right-hand side of Eq. (5.8) is given in terms of

$$\hat{q}_i^0 \triangleq E_{\hat{Q}^r}[R_{i0}], \tag{5.11}$$

$$\hat{q}_{ij}^{00} \triangleq E_{\hat{Q}^r}[R_{i0}R_{j0}], \tag{5.12}$$

$$\hat{q}_{i=j} \triangleq E_{\hat{Q}^r}\left[\sum_{h=1}^{H} R_{ih}R_{jh}\right]. \tag{5.13}$$

Eqs. (5.11-5.13) represent the posterior probabilities that (1) an individual region is healthy, (2) a pair of regions is healthy, and (3) two regions belong to the same disease cluster. We approximate these quantities using averages over the elements of $\mathcal{S}$.

As in Section 4.2, the model parameter estimates $\hat{\Theta}$ in the M-Step rely on marginal probabilities of $\hat{Q}^c(F, \bar{F})$. We compute these quantities after convergence of the variational posterior distribution:

$$\hat{s}_{ijk} = \hat{P}(F_{ij} = k|Y; \Theta) = \sum_{\bar{F}_{ij}} \hat{Q}_{ij}^c(F_{ijk} = 1, \bar{F}_{ij}), \tag{5.14}$$

$$\hat{u}_{ijk} = \hat{P}(\bar{F}_{ij} = k|Y; \Theta) = \sum_{F_{ij}} \hat{Q}_{ij}^c(F_{ij}, \bar{F}_{ijk} = 1). \tag{5.15}$$

M-Step: We fix the posterior probability estimates $\hat{Q}(R, F, \bar{F})$ and update the model parameter estimates $\hat{\Theta}$ by differentiating Eq. (5.6) with respect to each element of Θ and setting the gradient equal to zero. Many of these expressions are identical to Section 4.2; however, we repeat them below for completeness.

The update for π^r involves averaging the proportion of diseased regions across Gibbs samples:

$$\hat{\pi}_h^r = \frac{1}{NS} \sum_{i=1}^{N} \sum_{s=1}^{S} R_{ih}^s. \tag{5.16}$$

The multinomial prior π^f reduces to an average over the marginal posterior distribution:

$$\hat{\pi}_k^f = \frac{1}{C} \sum_{\langle i,j \rangle} \hat{s}_{ijk}, \tag{5.17}$$

where C is the total number of pairwise connections. The fMRI likelihood parameter

estimates are computed as weighted statistics of the data:

$$\mu_k = \begin{cases} \frac{\sum_{\langle i,j \rangle} \left[\hat{s}_{ijk} \sum_l B^l_{ij} + \hat{u}_{ijk} \sum_m \bar{B}^m_{ij} \right]}{\sum_{\langle i,j \rangle} \left[L \cdot \hat{s}_{ijk} + M \cdot \hat{u}_{ijk} \right]}, & k = \pm 1, \\ 0, & k = 0, \end{cases} \tag{5.18}$$

$$\sigma^2_k = \frac{\sum_{\langle i,j \rangle} \left[\hat{s}_{ijk} \sum_l (B^l_{ij} - \hat{\mu}_k)^2 + \hat{u}_{ijk} \sum_m (\bar{B}^m_{ij} - \hat{\mu}_k)^2 \right]}{\sum_{\langle i,j \rangle} \left[L \cdot \hat{s}_{ijk} + M \cdot \hat{u}_{ijk} \right]}, \tag{5.19}$$

where we have fixed $\mu_0 = 0$ for the component that represents zero functional synchrony to center the parameter estimates and regularize the model.

We use Newton's method to jointly update $\hat{\eta}$ and $\hat{\epsilon}$. We omit the details, which are very similar to the derivation in Section 4.2.

Identifying Disease Clusters The marginal posterior probability $\hat{Q}^r(R_i = h)$ expresses the likelihood that region i is healthy ($h = 0$) or that it belongs to disease cluster h. We estimate this quantity by averaging across Gibbs samples $\mathcal{S}$:

$$\hat{q}_{ih} \triangleq \hat{Q}^r(R_i = h) = \frac{1}{S} \sum_{s=1}^{S} R^s_{ih}. \tag{5.20}$$

5.3 Implementation Details

The variational EM implementation is similar to that of the functional model in Chapter 4. The main differences arise from the multi-class region assignments. We describe the related optimization choices in this section.

Initialization of Marginal Statistics Along with the model parameters $\Theta = \{\pi, \eta, \epsilon, \mu, \sigma^2\}$, it suffices to initialize the marginal posterior statistics $\hat{q}^0_i$, $\hat{q}^{00}_{ij}$, $\hat{q}_{i=j}$ described in Eqs. (5.11-5.13). This is done by computing the mean fMRI correlation across subjects and clustering these values (via the K-Means algorithm) to estimate the latent functional connectivity in each population. We then select a subset of regions with the greatest number of connectivity differences as diseased. The remaining regions are healthy.

We group the selected regions into disease clusters using Normalized Cut Spectral Clustering [82]. Entries W_{ij} of the affinity matrix are related to the difference in functional connectivity along the corresponding edge $\langle i, j \rangle$. This is because our formulation specifies a disease cluster to be regions that are *abnormally* connected to each other but *normally* connected to nodes in other clusters.

Table 5.1: Likelihood parameterizations used to generate synthetic data.

	μ_{-1}	μ_0	μ_1	σ^2_{-1}	σ^2_0	σ^2_1
`Good Data`	-0.35	0	0.35	0.05	0.05	0.05
`Noisy Data`	-0.18	0	0.36	0.050	0.058	0.073

As described in Section 4.2.3, the algorithm proceeds by computing the joint posterior distribution $\hat{Q}^c(\cdot)$ in the E-step and alternatively updates $\hat{Q}^r(R)$ and $\hat{Q}^c(\cdot)$ until convergence. The algorithm then estimate the model parameters in the M-step and iterate. We initialize the algorithm ten times to sample the solution space and select the solution with the lowest free energy.

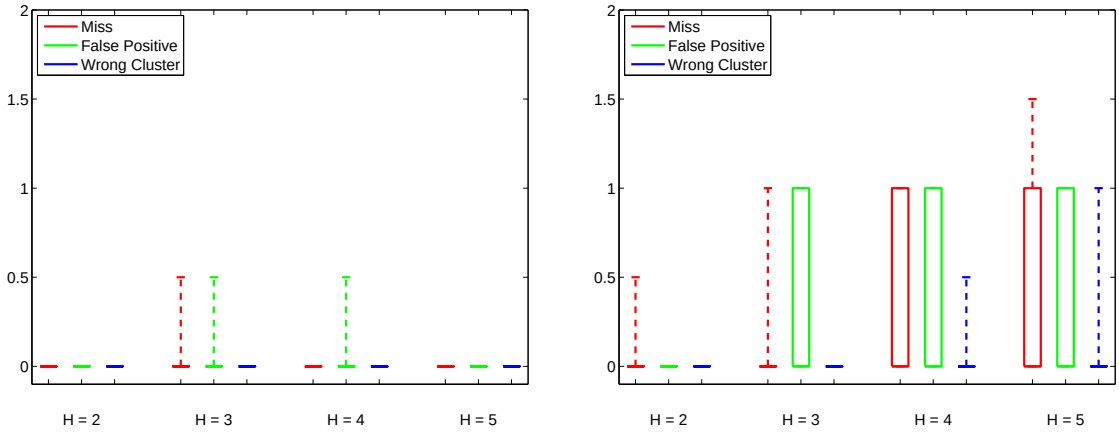
Cluster Alignment Like most clustering algorithms, the cluster indices $h = 1, \ldots, H$ are arbitrarily assigned. Empirically, the Gibbs samples $\mathcal{S}$ concentrate around modes in $\hat{Q}^r(R)$ that have the same region grouping but with permuted cluster indices. Since these permutations are qualitatively identical, we align the Gibbs samples to maximize the number of consistently labeled regions. This results in a peakier distribution $\hat{Q}^r(R)$ and reduces the number of samples needed for stable posterior estimation. We emphasize that permuting the cluster indices does not affect Eqs. (5.11-5.13), which depends only on the relative grouping between regions. Hence, our estimates of latent functional connectivity and the non-random model parameters are unchanged.

5.4 Results: Synthetic Data

We first evaluate our algorithms on synthetic data sampled from the model. We mimic the organization of our clinical dataset by specifying a latent template with 78 regions; there are two regions in each disease cluster, and the remaining are healthy. We fix η, ϵ, π^f to the values inferred from the clinical experiments.

We consider two likelihood parameterizations for $\{\mu, \sigma^2\}$ as shown in Table 5.1; these are the same as in Section 4.4. The `Good Data` parameterization assumes a clear separation between the Gaussian distributions for different values of latent functional connectivity. In this case, we can accurately infer the templates $\{F, \bar{F}\}$, which are subsequently used to cluster regions. The `Noisy Data` parameterization uses the ML estimates from the clinical dataset, which results in a significant overlap in the data distributions.

Given the region labels R, we sample the graph structures $\{T^h\}$, the latent functional templates $F, \bar{F}$ and the observed data $\{B, \bar{B}\}$ according to Fig. 5-1(b). We fit the model using the variational EM algorithm presented in Section 5.2 and extract three error metrics:



(a) `Good Data` (b) `Noisy Data`

Figure 5-2: Region assignment errors for the synthetic experiment. The median error is zero in all cases. The box and error bars denote the 75^{th} and 90^{th} percentiles, respectively. The statistics are computed over 10 re-samplings of the latent templates and observed data.

the number of diseased regions that are incorrectly labeled as healthy (*miss*), the number of healthy regions that are labeled as diseased (*false-positive*), and the number of regions that are assigned to the wrong disease cluster (*wrong cluster*). We re-sample the latent template and observed data 10 times to collect statistics.

Fig. 5-2 illustrates the error in region assignments when varying the number of disease clusters ($H = 2, 3, 4, 5$). Unlike the results (functional model) in Fig. 4-3, the errors are significantly higher for the noisy parameterization. Intuitively, the algorithm relies more heavily on the latent connectivity assignments to cluster the diseased regions. Poor inference of latent connectivity from the fMRI data produces incorrect region labels.

Similarly, the error in Fig. 5-2(b) increases with the number of disease clusters H. Since connections across disease clusters should be normal, higher values of H increase the number of *healthy* connections that must be observed to correctly partition the diseased regions. Moreover, the complexity of the problem increases with H. It is reasonable that incorrect latent connectivity assignments (based on noisy data) have a greater impact if they do not conform with the rules introduced in Section 5.1.

Finally, the maximum error in Fig. 5-2 is less than two for any error type. This suggest that our algorithm can accurately infer the ground truth region assignments given (noisy) observations whose statistics match our clinical dataset.

Table 5.2: Parameters of the multi-class model in Fig. 5-1(b) estimated from the clinical data. We vary the number of disease clusters H.

	π^f_{-1}	π^f_0	π^f_1	η	ϵ	μ_{-1}	μ_0	μ_1	σ^2_{-1}	σ^2_0	σ^2_1
$H=2$	0.33	0.46	0.21	0.15	0.026	-0.18	0	0.36	0.050	0.058	0.073
$H=3$	0.33	0.46	0.21	0.15	0.024	-0.18	0	0.36	0.050	0.058	0.073
$H=4$	0.33	0.46	0.21	0.15	0.022	-0.18	0	0.36	0.050	0.058	0.073
$H=5$	0.33	0.46	0.21	0.15	0.017	-0.18	0	0.36	0.050	0.058	0.073

5.5 Results: Clinical Data

Once again, we demonstrate our model on the clinical study of schizophrenia (see Section 2.7). The fMRI connectivity B^l_{ij} is computed as the Pearson correlation coefficient between the mean time courses of regions i and j in subject l.

Table 5.2 reports the parameter values estimated by our algorithm when varying the number of disease clusters. Encouragingly, the results are extremely stable across different values of H. Only the latent noise parameter ϵ changes between trials; its value decreases as H increases. This is because our model can explain more of the functional differences between the populations with a larger number of disease clusters. Consistency in the other parameter values suggests that the model infers similar latent connectivity patterns for each value of H.

Fig. 5-3 depicts the maximum *a posteriori* disease cluster assignments in the brain for $H = 2, 3, 4, 5$. In all cases $\hat{q}_{ih} > 0.8$. Fig. 5-4 presents the estimated graphs of abnormal connectivity for each cluster; they are computed using a similar procedure as in Section 4.3.2. As seen, the results exhibit a nesting property. Specifically, as the number of disease clusters increases, the model progressively adds subsets of regions while largely preserving the old clustering organization. The only exception are regions the left transverse temporal and superior temporal gyri. We also observe symmetry in region assignments across the hemispheres, as evident for the posterior cingulate (PCC) and the paracentral gyri (pC). Once again, this may be attributed to the well-documented symmetry found in resting-state fMRI correlations [90]. We elaborate on the model results in the following section.

5.6 Discussion

We present an extension to the functional model in Chapter 4 that allows for multiple clusters of diseased regions. We describe a variational EM algorithm to fit the revised model to the observed fMRI correlations. Finally, we demonstrate preliminary results both on synthetic and on clinical data.

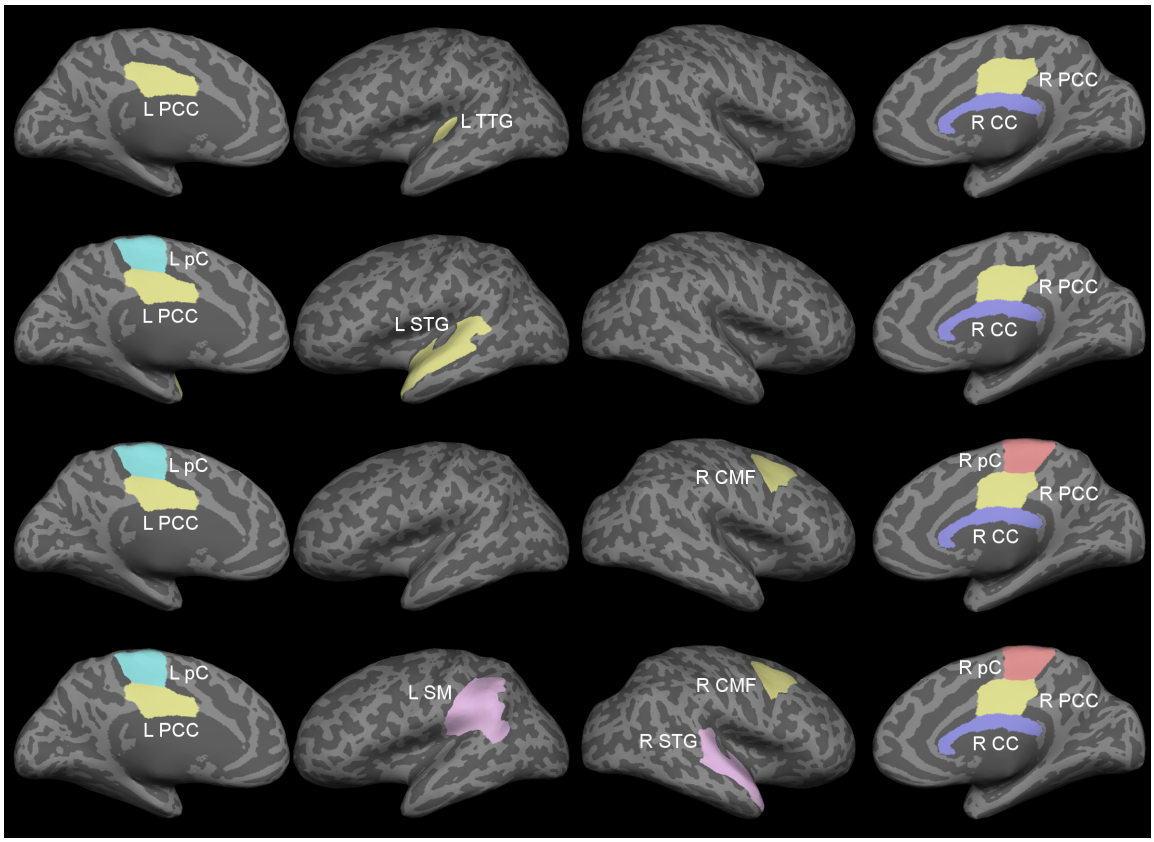
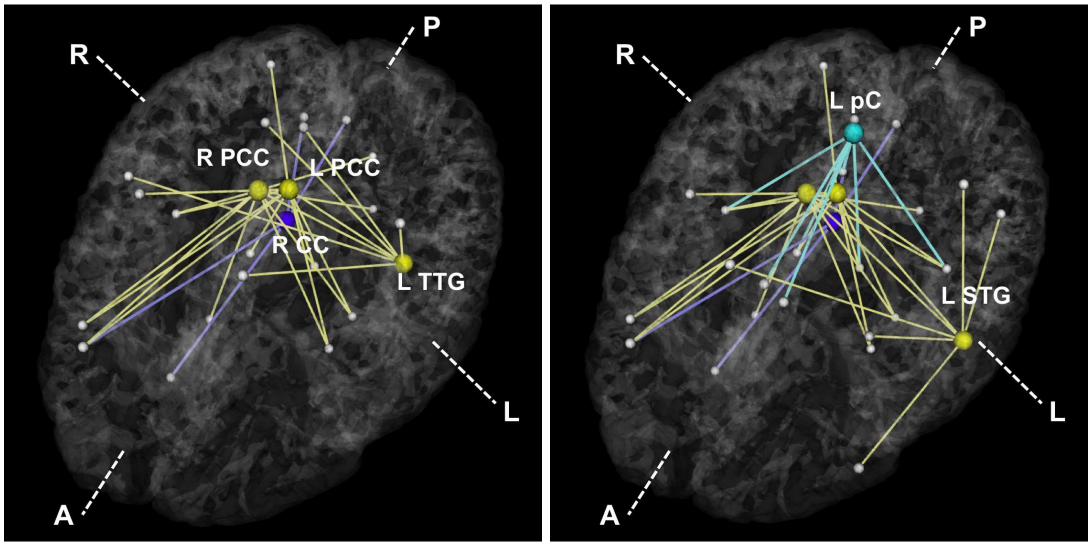
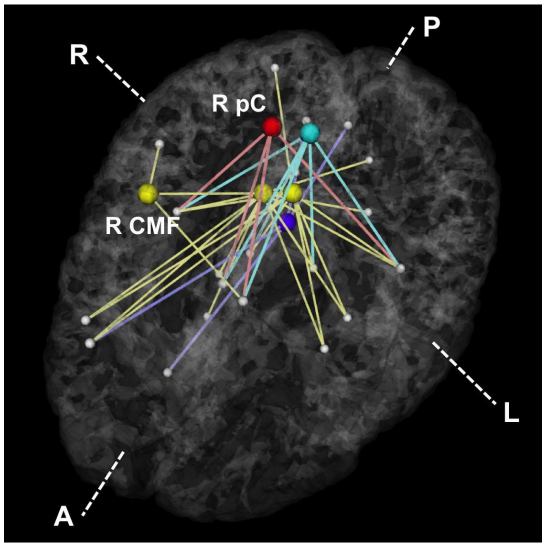


Figure 5-3: Disease clusters for $H = 2, 3, 4, 5$. Clusters are delineated by the yellow, dark blue, light blue, red and pink regions, respectively. To facilitate comparison, we align the clusters across values of H. The regions correspond to the posterior cingulate (L & R PCC), the corpus callosum (R CC), the transverse temporal gyrus (L TTG), the paracentral gyrus (L & R pC), the superior temporal gyrus (L & R STG), the caudal middle frontal gyrus (R CMF) and the supramarginal gyrus (L SM).

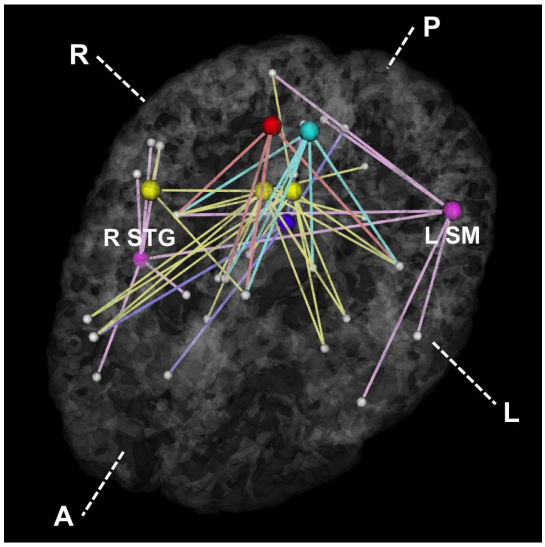


(a) $H = 2$

(b) $H = 3$



(c) $H = 4$



(d) $H = 5$

Figure 5-4: Estimated graph of functional connectivity differences. We use the color scheme in Fig. 5-3 to identify disease clusters. For clarity of presentation, we only label new diseased regions in each figure.

Fig. 5-3 illustrates the regions implicated by our model when varying the number of disease clusters H. These results correspond well to the disease foci presented in Section 4.5. For example, the green cluster for $H = 2$ (top row of Fig. 5-3) is identical to the significant regions identified by the functional model in Fig. 4-5. Although the transverse temporal gyrus is not included in the disease clusters for $H > 2$ (bottom three rows), the right and left posterior cingulate remain grouped together. In addition, the right and left superior temporal gyri and the right caudal middle frontal gyrus are all identified by the joint model in Fig. 4-9 when varying the region prior.

The nested structure observed in both Fig. 5-3 and Fig. 5-4 is a highly desirable trait. It suggests that the model learns a stable set of functional differences. As we increase the number of clusters, the model explains more of the overall pattern by adding diseased regions. This is further evidenced by the consistent parameter estimates in Table 5.2.

Although not shown, we observe a fair amount of variability in the disease cluster assignments based on initialization. Specifically, we recover several solutions that identify the same subset of diseased regions, but with slightly different groupings into clusters. In some cases, the Gibbs samples are drawn from 2-3 *modes* with different clusterings, so the marginal posterior distributions are not as sharp. These solutions have nearly identical free energy values. Hence, the alternate groupings may signify complex relationships between the diseased region that cannot be explained by our model.

We emphasize that the results presented in this chapter are preliminary. Further analysis must be done to fully characterize the behavior of the multi-class model. Such analysis may explain the inconsistencies in the left-temporal lobe of the brain in Fig. 5-3 or the cluster modes described above. However, the stability across clusters and the correspondence to the results in Section 4.5 reveal this framework to be a promising direction for future work.

Chapter 6

Conclusion

In this thesis we introduced a powerful generative framework to analyze brain connectivity in clinical applications. Our models treat resting-state fMRI and DWI data as imperfect observations of the brain. We abstract the "true" anatomy and function into a collection of latent variables. Differences attributed to a neuropsychiatric disorder are captured by these latent templates.

Chapter 3 introduces our basic assumptions. Here, we define the latent anatomical and functional connectivity variables and present an intuitive data likelihood. Our initial model identifies connections that are significantly affected by the disease. Recognizing that distributed connectivity results are difficult to interpret, we focus on an alternative question in Chapter 4: do certain regions act as centers of abnormal activity? In response, we add region labels and describe how to construct the graph of functional connectivity differences. Our results localize a stable subset of *disease foci* in schizophrenia. Finally, Chapter 5 expands our region-based model to incorporate multiple *disease clusters*. Our preliminary results suggest a nested structure, which progressively includes regions as the number of clusters increases.

We deliberately focus on a generative, rather than a discriminative, framework to analyze the effects of a disorder. This reflects our belief that advancements in the understanding and treatment of illnesses will depend on our ability to *explain* (and not merely *identify*) how and why various phenomena arise. To this end, our models have allowed us to explore new and exciting clinical questions for which no other unified method has been proposed. We have identified abnormal connectivity patterns using whole-brain information with minimal *a priori* knowledge. In fact, the only user-specified parameter in this entire thesis is the number of disease clusters H in Chapter 5.

Despite the benefits and possibilities of generative methods, there is one major drawback: the quality of any such model depends on how closely our assumptions reflect the true underlying process. The simple interactions formulated in this thesis may not capture the full complexity of the brain. These limitations suggest potential directions for future work.

6.1 Future Directions

In this section we sketch out three avenues of research that promise to enhance our understanding of and ability to model the brain.

1. **Incorporating Temporal Information:** Traditional functional connectivity analysis extracts patterns from the fMRI data based on the entire time course at each spatial location. These techniques effectively assume temporal stationarity of resting-state fMRI, i.e., there is a single connectivity network that persists across time. However, recent work demonstrates that functional relationships are significantly affected by cognitive processes, fatigue and conscious awareness [25, 27, 49]. This suggests that modeling the time-varying nature of functional interactions is crucial to understanding brain connectivity.

   Multivariate auto-regressive (MAR) models are gaining popularity for resting-state fMRI analysis [18, 45, 67]. Here, the observation of a network at a given time depends on its past history. As such, the MAR framework offers a computationally tractable method for incorporating temporal information. It has been used in clinical applications [44] and in multi-modal analysis of fMRI and DWI data [21]. Therefore, using a MAR model to account for the fMRI dynamics may strengthen our generative approach in this thesis.

2. **Sophisticated Modeling of Brain Interactions:** The graphical models presented in this thesis distill the complexities of the brain into a set of Gaussian and multinomial distributions. Although this simplification allows us to infer aggregate properties of the signals, we might be losing vital information. In the past there has been some effort to characterize the relationship between fMRI time courses and neuronal activity [46]. Task-based fMRI experiments typically assume a linear system whereby each neuronal impulse is filtered by the hemodynamic response function [35]. However, the problem is more challenging for resting-state fMRI, which lacks a guiding protocol. Recently, nonlinear relationships have been proposed [10, 68]. Incorporating these results into the likelihood function for the observed variables is a difficult but worthwhile task.

   Similarly, many studies have analyzed the relationship between anatomical and functional connectivity [47,48,85]. Future iterations of our model should account for multi-stage anatomical connections as well as excitatory versus inhibitory neural pathways.

3. **Data-Driven Region Definitions:** Our work depends on a set of pre-defined regions

that are consistent across subjects. This thesis relies on the standard Freesurfer parcellation into Brodmann areas [29]. Although roughly based on the underlying cytoarchitecture of the brain, Brodmann areas do not always represent functionally or anatomically coherent regions. Specifically, different areas within a single region may exhibit distinct connectivity patterns. Therefore, we would like to develop a method to automatically extract regions across subjects. Such a model will draw inspiration from landmark-based and region growing approaches [69, 88].

The proposed generative formulation is a powerful tool for brain connectivity analysis and will provide a new direction in the field. In the immediate future, we seek to apply our method to other clinical populations, such as epilepsy and Huntington's Disease. We are also interested in the effects of aging, which require a more gradual evolution of our latent templates. As seen, our basic framework can be adapted to a wide variety of clinical and neuroscientific problems. We are confident that it will further our understanding and exploration of the brain.

Appendix A

Robust Feature Selection via Random Forests

This appendix details our work on robust feature selection, as applied to a population study of resting-state functional connectivity. Our method identifies a sparse set of functional correlations that captures much of the information about schizophrenia. Furthermore, by incorporating minimal *a priori* knowledge, we can predict the clinical diagnosis of a test subject with substantially higher accuracy. Encouragingly, many of the selected connections are also found using our joint generative model in Chapter 3. These results further validate our probabilistic framework. The work in this appendix was published in [96].

Traditional population analysis relies on univariate tests to identify significant differences [40, 66, 100, 103], which ignores *networks* of abnormal connectivity within the brain. Due to the limited number of subjects, univariate tests are often done once using the entire dataset; stability of the method and of the results is rarely assessed. We address these limitations through ensemble learning. The Random Forest is an ensemble of decision tree classifiers that incorporates multiple levels of randomization [11]. Each tree is grown using a random subset of the training data; each node is constructed by searching over a random subset of features. The Random Forest derives a score for each feature, known as the Gini Importance (GI), which summarizes its discriminative power and can be used as an alternative to univariate statistics.

Our approach to feature selection confers several advantages. The randomization over subjects is designed to improve generalization accuracy, especially given a small number of training examples relative to the number of features. The randomization over features increases the likelihood of identifying *all* functional connections useful for group discrimination (rather than an uncorrelated subset). Finally, due to the ensemble-based learning, the Random Forest produces nonlinear decision boundaries. Hence, it can capture significant *patterns* of functional connectivity across distributed networks in the brain. We demonstrate that the significant functional connections based on univariate tests vary sub-

stantially across different subsets of the data and have poor predictive power. In contrast, GI is a stable metric that selects good features for classification.

A.1 Methods

We begin with an overview of the Random Forest algorithm and construction of the Gini Importance measure. We then review the standard two-sample t-test used for comparison, and conclude with a description of our empirical validation procedure. In this application, we treat functional correlations between two brain regions as features.

A.1.1 Random Forest and Gini Importance

The Random Forest is an ensemble of decision-tree classifiers. At each decision node, the algorithm selects a feature and threshold that maximize the separation between classes [86]. Mathematically, let ν represent a decision node of a single tree. We define n_ν to be the total number of samples assigned to ν, such that n_ν^1 is the number of samples in the first class and n_ν^2 is the number of samples belonging to the second class ($n_\nu = n_\nu^1 + n_\nu^2$). The *Gini Impurity* $G(\nu)$ estimates the probability that two random observations, drawn from the same class distribution as the initial n_ν samples, will have different labels:

$$G(\nu) = \frac{n_\nu^1}{n_\nu}\left(1 - \frac{n_\nu^1}{n_\nu}\right) + \frac{n_\nu^2}{n_\nu}\left(1 - \frac{n_\nu^2}{n_\nu}\right). \tag{A.1}$$

Given a feature f and a threshold η, we construct the two child nodes ν_1 and ν_2 of ν by partitioning the dataset along f according to η. As a result, $n_{\nu_1}(f,\eta)$ of the initial samples are assigned to child node ν_1, and the remaining $n_{\nu_2}(f,\eta)$ samples are assigned to child node ν_2. We can now compute the change in Gini Impurity between the node ν and its children:

$$\Delta G(\nu; f, \eta) = G(\nu) - \frac{n_{\nu_1}(f,\eta)}{n_\nu} G(\nu_1) - \frac{n_{\nu_2}(f,\eta)}{n_\nu} G(\nu_2). \tag{A.2}$$

During training, the Random Forest selects the feature $f^*(\nu)$ and the corresponding threshold $\eta^*(\nu)$ that together maximize Equation (A.2) at node ν. This process is continued recursively for all child nodes until each leaf of the tree defines a unique class. The final classification is obtained by a majority vote among all the random trees.

The Gini Importance (GI) of a feature f is found by integrating the reduction in Gini Impurity throughout the entire forest:

$$GI(f) = \sum_{trees} \sum_{\{\nu: f = f^*(\nu)\}} \Delta G(\nu; f, \eta^*(\nu)). \tag{A.3}$$

Thus, GI can be viewed as the aggregate amount of separation between the two classes gained by selecting a particular feature and corresponding threshold. We use this quantitative measure to rank the features according to their predictive power.

A.1.2 Baseline Univariate Tests

Univariate tests are one of the standard tools used in the clinical analysis of functional connectivity [40, 66, 100, 103]. The two-sample t-test evaluates the null hypothesis that the population means of a (normally distributed) feature are equal. Mathematically, let $\bar{f}_C$ and $\bar{f}_S$ be the means of feature f for the control and schizophrenia populations, respectively, and let $\bar{\sigma}_C^2$ and $\bar{\sigma}_S^2$ denote the corresponding empirical variances. The t-score for f is defined

$$t_f = \frac{\left|\bar{f}_C - \bar{f}_S\right|}{\sqrt{\frac{(N_C-1)\bar{\sigma}_C^2+(N_S-1)\bar{\sigma}_S^2}{N_C+N_S-2} \cdot \left[\frac{1}{N_C} + \frac{1}{N_S}\right]}}, \tag{A.4}$$

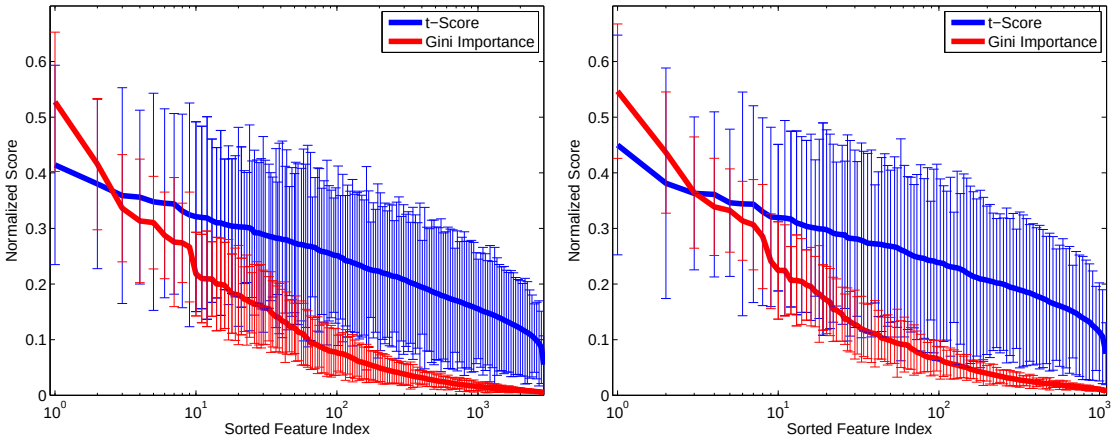
where N_C and N_S denote the number of subjects in each group. The significance, or p-value, represents the probability of obtaining a statistic greater in magnitude than t_f under the null hypothesis.

A.1.3 Validation

We use ten-fold cross-validation to quantify the performance of each method. The dataset is randomly divided into 10 subsets, each with an equal number of controls and schizophrenia patients. We then compute the Gini Importance values and t-scores using 9 of these subsets and reserve one for testing. This process is repeated for each of the 10 sub-groups. Additionally, we repeat this re-sampling process 10 times to collect stable statistics.

Cross-validation allows us to evaluate several aspects of each feature selection methods. For example, we assess the rate of decay of the GI values and t-scores. A rapid decay is indicative of a sparse representation for the population differences. Additionally, we investigate the variability of the scores and the stability of the feature rankings. Little or no fluctuation in the scores and rank-order implies a robust representation across different subsets of the data. Finally, we examine the prediction accuracy for various set sizes K.

During testing, we rank the functional correlations either by GI value or by t-score magnitude. Our assumption is that the significant differences between the control and clinical populations are contained in the first K features. We assess this hypothesis by training both a Random Forest classifier and a Radial Basis Function Support Vector Machine (RBF-SVM) [20] using just these K functional correlations, and evaluating the classifica-



(a) Full Dataset (b) Selected Features

Figure A-1: Stability of the GI values and t-scores on a log scale. For visualization, the values are normalized by the maximum GI and maximum t-score, respectively. Thick lines represent mean values, and the error bars correspond to standard deviations over the 100 cross-validation runs.

tion accuracy on the held-out group. Utilizing multiple classifiers ensures a fair comparison between GI and univariate tests.

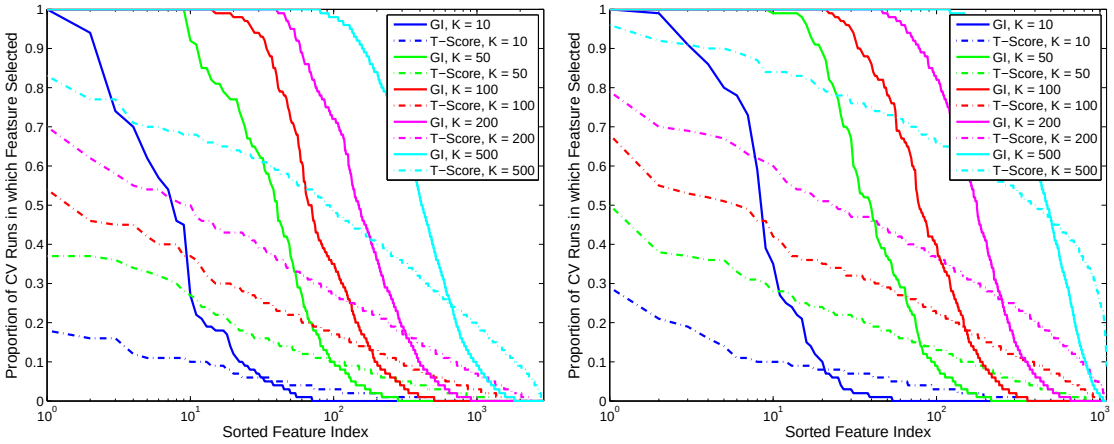
A.2 Experimental Results

We demonstrate our model on the clinical study of schizophrenia (see Section 2.7). We extract the fMRI connectivity B^l_{ij} between regions i and j in subject l by computing Pearson correlation coefficients of the time courses between every pair of voxels in the two regions, applying the Fisher-r-to-z transform to each correlation (to enforce normality), and averaging these values. Since our anatomical regions are large, the correlation between the mean time courses of two regions shows poor correspondence with the distribution of voxel-wise correlations between them. Therefore, we believe our measure is more appropriate for fMRI correlations across subjects than the standard correlation of mean time courses.

To inject prior clinical knowledge, we pre-select 8 brain structures (corresponding to 16 regions) that are believed to play a role in schizophrenia: the superior temporal gyrus, rostral middle frontal gyrus, hippocampus, amygdala, posterior cingulate, rostral anterior cingulate, parahippocampal gyrus, and transverse temporal gyrus. We model only the 1096 $\left(16 \times 76 - \binom{16}{2}\right)$ unique pairwise connections between these ROIs and all other regions in the brain.

Fig. A-1 depicts the stability of GI values and t-scores for each functional correlation

across all 100 cross-validation runs. As seen, the t-scores exhibit far greater variability that the Gini Importance values. Additionally, the variance in GI is concentrated among the top features, whereas less-informative features are always assigned values near zero. Hence, although the top functional correlations may be ranked differently during each cross-validation run, the Random Forest isolates a consistent set of predictive features. In contrast, the t-scores vary uniformly over all features, regardless of significance. Thus, the set of predictive features can vary drastically over cross-validation runs.

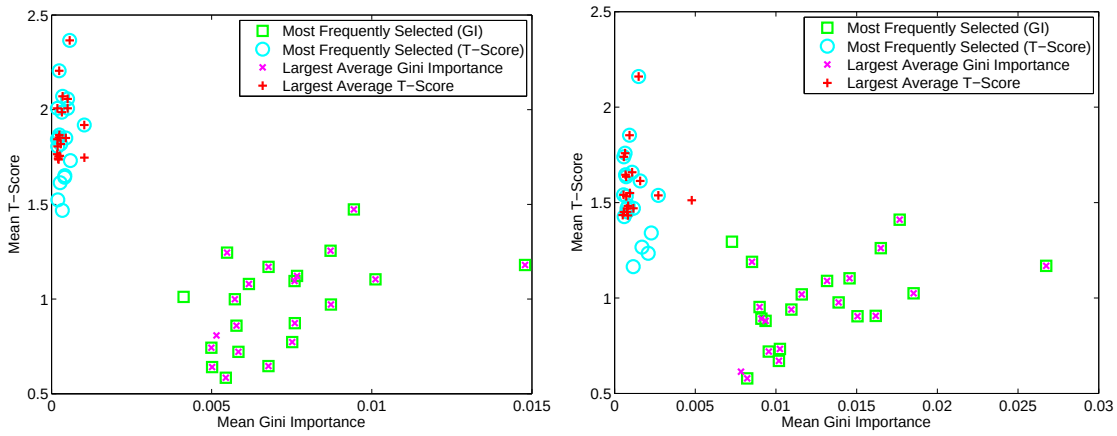


(a) Full Dataset (b) Selected Features

Figure A-2: Proportion of the 100 cross-validation runs during which the feature is selected. The solid lines denote performance based on GI values for various K. The dashed lines represent the corresponding metric using t-scores.

Fig. A-2 shows the proportion of cross-validation runs during which a particular functional correlation is ranked among the top K features, as measured by GI value or t-score. We observe that the decay in the proportion of iterations based on GI is relatively sharp from one to zero. Hence, if a feature is relevant for group discrimination, it tends to be ranked among the top; otherwise, it is almost always ignored. In contrast, feature selection based on t-scores is inconsistent and depends on the dataset. It is worth noting that *none* of the functional correlations are ranked in the top 500 by t-score for all 100 cross-validation iterations, even when we *a priori* specify the regions of interest.

Fig. A-3 compares the average Gini Importance and average t-scores of the top 20 functional correlations as specified by the average score and the frequency of selection for each method, respectively. Notice that the highest average scores are well correlated with the most often selected. However, features that are ranked highly by one method are scored poorly by the other. This may be attributed to the variability of t-scores over



(a) Full Dataset (b) Selected Features

Figure A-3: Relationship between the average GI and the average t-score for the top 20 functional correlations using each feature selection algorithm. The green boxes and blue circles denote features that were most frequently included in the top 50 features for each method.

Table A.1: Classification accuracy based on the entire dataset.

K	GI, RF Classify	GI, SVM Classify	t-score, RF Classify	t-score, SVM Classify
25	0.59 ± 0.047	0.60 ± 0.040	0.50 ± 0.10	0.51 ± 0.053
150	0.58 ± 0.026	0.56 ± 0.037	0.54 ± 0.059	0.53 ± 0.038
300	0.57 ± 0.043	0.55 ± 0.040	0.57 ± 0.073	0.55 ± 0.031

Table A.2: Classification accuracy based on the expert-selected regions.

K	GI, RF Classify	GI, SVM Classify	t-score, RF Classify	t-score, SVM Classify
10	0.75 ± 0.034	0.66 ± 0.033	0.53 ± 0.053	0.54 ± 0.058
50	0.66 ± 0.048	0.60 ± 0.043	0.57 ± 0.056	0.57 ± 0.050
100	0.63 ± 0.029	0.59 ± 0.032	0.57 ± 0.034	0.58 ± 0.058

the cross-validation iterations. It also suggests that the differences between a control and schizophrenia population are captured in a complex pattern of functional connectivity, which cannot be detected by univariate tests.

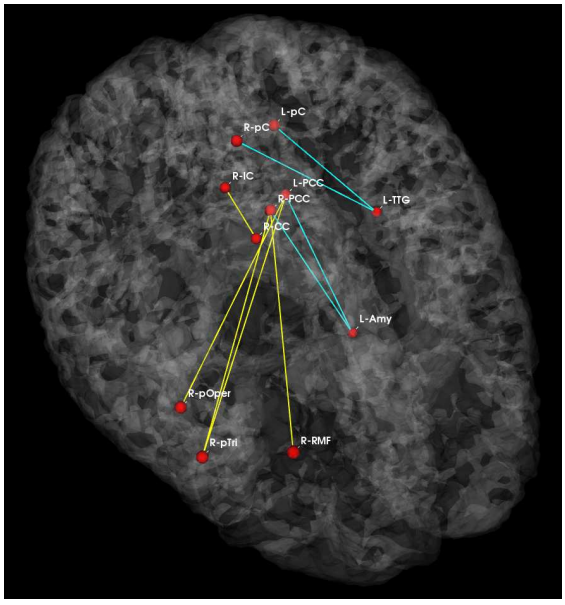
Tables A.1 and A.2 report the classification accuracy for each feature selection/classifier pair based on the entire dataset and on the expert-selected ROIs, respectively. The three values of K roughly correspond to thresholding the mean p-value of the K^{th} feature to 0.01, 0.05 and 0.10, For small values of the feature count K, the classification accuracy based on univariate statistics is near chance. This indicates that functional connectivity selection

based on large t-scores has *no predictive power*. In contrast, we achieve as high as 75% prediction accuracy using GI values.

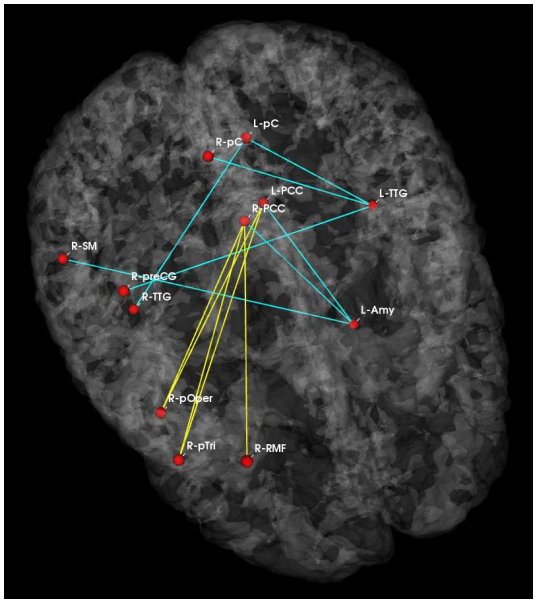
As K increases, all classifiers converge towards the base accuracy obtained by incorporating all of the features. However, the GI-based classifiers approach this baseline from above, whereas the univariate classifiers approach from below. This behavior is reflected in Tables A.1 and A.2. In particular, the classification accuracy decreases with K in the first two columns (GI) and increases with K in the last two columns (univariate). It is worth noting that while the classification accuracy improves with K for the univariate classifiers, the average p-value is rapidly decreasing. Therefore, one would never report these connections as being significant.

The above results demonstrate that the Gini Importance is a more robust feature selection criterion for clinical data than the univariate t-test. Fig. A-4 and Table A.3 report the features (connections) selected during at least half of the cross-validation iterations. For GI, we depict results for $K = 15$, which yields the best classification accuracy. For t-score, we used $K = 150$ for the full dataset and $K = 50$ for the selected features. This roughly corresponds to p-values less than 0.05. We observe that many of the significant functional correlations are consistent between Fig. A-4(a) and Fig. A-4(b). This confirms the clinical hypotheses about brain regions that play a role in schizophrenia. In contrast, Fig. A-4(c-d) scarcely exhibit any consistent connections.

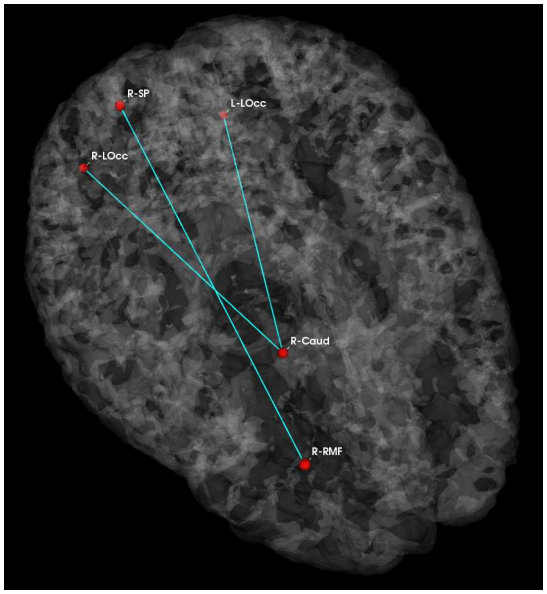
As mentioned earlier , over half of the connections in Table A.3 correspond to the significant differences reported in Table 3.3 and Table 3.4 from Chapter 3. We observe that the joint model (blue/purple) has a much greater overlap with the GI-based features than the functional model (green). The correspondence between our generative and discriminative methods suggests that the anatomical information in DWI data somehow stabilizes the estimates of latent functional connectivity. This, in turn, allows the joint model to pinpoint discriminative connections. The GI method also detects similar connectivity patterns in Fig. A-4(a-b) as in Fig. 3-9. Namely, schizophrenia patients exhibit increased functional connectivity between the parietal/posterior cingulate region and the frontal lobe and reduced functional connectivity between the parietal/posterior cingulate region and the temporal lobe. These results confirm the hypotheses of widespread functional connectivity changes in schizophrenia and of functional abnormalities involving the default network.



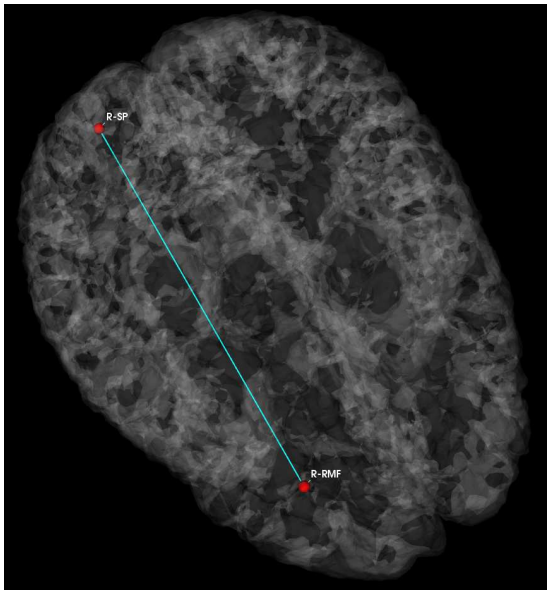
(a) GI, Full Dataset



(b) GI, Selected Features



(c) t-score, Full Dataset



(d) t-score, Selected Features

Figure A-4: Connections selected during at least half of the cross-validation runs. Blue lines indicate higher connectivity in the control group; yellow lines indicate higher connectivity in the schizophrenia population.

Table A.3: Connections selected during at least half of the cross-validation runs. For GI, we used $K = 15$, which gives the best classification accuracy. For t-score, we used $K = 150$ for the full dataset and $K = 50$ for the selected features. This roughly corresponds to p-values less than 0.05. The blue connections were identified by the joint generative model in Chapter 3, and the green connection overlaps with the fMRI-only model.

Region 1	Region 2	Prop Sel
GI, Full Dataset		
L Posterior Cingulate (L-PCC)	L Amygdala (L-Amy)	1.00
R Paracentral Gyrus (R-pC)	L Transverse Temporal (L-TTG)	1.00
R Posterior Cingulate (R-PCC)	R Pars Triangularis (R-pTri)	0.89
L Transverse Temporal (L-TTG)	L Paracentral Gyrus (L-pC)	0.84
R Posterior Cingulate (R-PCC)	L Amygdala (L-Amy)	0.83
R Pars Triangularis (R-pTri)	L Posterior Cingulate (L-PCC)	0.78
R Pars Opercularis (R-pOper)	L Posterior Cingulate (L-PCC)	0.72
R Isthmus Cingulate (R-IC)	R Posterior Cingulate (R-PCC)	0.59
R Rostral Middle Frontal (R-RMF)	R Corpus Callosum (R-CC)	0.57
GI, Selected Dataset		
L Posterior Cingulate (L-PCC)	L Amygdala (L-Amy)	1.00
R Paracentral Gyrus (R-pC)	L Transverse Temporal (L-TTG)	1.00
R Posterior Cingulate (R-PCC)	R Pars Triangularis (R-pTri)	1.00
R Pars Triangularis (R-pTri)	L Posterior Cingulate (L-PCC)	0.97
R Posterior Cingulate (R-PCC)	L Amygdala (L-Amy)	0.96
L Transverse Temporal (L-TTG)	L Paracentral Gyrus (L-pC)	0.95
R Pars Opercularis (R-pOper)	L Posterior Cingulate (L-PCC)	0.93
R Rostral Middle Frontal (R-RMF)	R Posterior Cingulate (R-PCC)	0.76
R Posterior Cingulate (R-PCC)	R Pars Opercularis (R-pOper)	0.62
R Transverse Temporal (R-TTG)	L Paracentral Gyrus (L-pC)	0.55
R Supramarginal Gyrus (R-SM)	L Amygdala (L-Amy)	0.51
R Precentral Gyrus (R-preCG)	L Transverse Temporal (L-TTG)	0.51
t-score, Full Dataset		
R Superiorparietal Gyrus (R-SP)	R Rostral Middle Frontal (R-RMF)	0.64
R Lateral Occipital Cortex (R-LOcc)	R Caudate Nucleus (R-Caud)	0.55
L Lateral Occipital Cortex (R-LOcc)	R Caudate Nucleus (R-Caud)	0.53
t-score, Selected Dataset		
R Superiorparietal Gyrus (R-SP)	R Rostral Middle Frontal (R-RMF)	0.50

Bibliography

[1] D.C. Alexander. Multiple-fiber reconstruction algorithms for diffusion mri. *Annals of the New York Academy of Sciences*, 1046(1):113–133, 2005.

[2] M.J. McKeown an S. Makeig, G.G. Brown, T-P. Jung, S.S. Kindermann, A.J. Bell, and T.J. Sejnowski. Analysis of fmri data by blind separation into spatial independent components. *Human Brain Mapping*, 6:160–188, 1998.

[3] J. Ashburner and K.J. Friston. Nonlinear spatial normalization using basis functions. *Human Brain Mapping*, 7:254–266, 1999.

[4] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber-tractography in human brain using diffusion tensor mri (dt-mri) data. *Magnetic Resonance in Medicine*, 44(4):625–632, 2000.

[5] P.J. Basser and C. Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of Magnetic Resonance*, 111:209–219, 1996.

[6] J. Bezdek. Review of mr image segmentation techniques using pattern recognition. *Medical Physics*, 20(4):1033–1048, 1993.

[7] D. Le Bihan, J-F. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.

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

[9] R.L. Bluhm, J. Miller, R.A. Lanius, E.A. Osuch, K. Boksman, R.W.J. Neufeld, J. Théberge, B. Schaefer, and P. Williamson. Spontaneous low-frequency fluctuations in the bold signal in schizophrenic patients: Abnormalities in the default network. *Schizophrenia Bulletin*, pages 1–9, 2007.

[10] M. Breakspear, J.R. Terry, and K.J. Friston. Modulation of excitatory synaptic coupling facilitates sychronization and complex dynamics in a biophysical model of neuronal dynamics. *Network¿ Computational Neural Systems*, 14:703–732, 2003.

[11] L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

[12] R.L. Buckner, J.R. Andrews-Hanna, and D.L. Schacter. The brain's default network anatomy, function, and relevence to disease. *Annals of the N.Y. Academy of Sciences*, 1124:1–38, 2008.

[13] R.L. Buckner, J. Sepulcre, T. Talukdar, F.M. Krienen, H. Liu, T. Hedden, J.R. Andrews-Hanna, R.A. Sperling, and K.A. Johnson. Cortical hubs revealed by intrinsic functional connectivity: Mapping, assessment of stability, and relation to alzheimer's disease. *The Journal of Neuroscience*, 29(6):1860–1873, 2009.

[14] R.L. Buckner and J.L. Vincent. Unrest at rest: Default activity and spontaneous network correlations. *NeuroImage*, 37(4):1091–1096, 2007.

[15] E. Bullmore and O. Sporns. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews*, 10:186–198, 2009.

[16] J. Burns, D. Job, M.E. Bastin, H. Whalley, T. MacGillivray, E.C. Johnstone, and S.M. Lawrie. Structural disconnectivity in schizophrenia: A diffusion tensor magnetic resonance imaging study. *British Journal of Psychiatry*, 182:439–443, 2003.

[17] V.D. Calhoun and T. Adali. Unmixing fmri with independent component analysis. *IEEE Engineering in Medicine and Biology Magazine*, 25(2):79–90, 2006.

[18] K. Chang and G. Glover. Time-frequency dynamics of resting-state brain connectivity measured with fmri. *NeuroImage*, 50(1):81–98, 2010.

[19] D. Cordes, V.M. Haughton, K. Arfanakis, J.D. Carew, and K. Maravilla. Hierarchical clustering to measure connectivity in fmri resting-state data. *Magnetic Resonance in Medicine*, 20(4):305–317, 2002.

[20] C. Cortes and V. Vapnik. Support vector networks. *Machine Learning*, 20(3):273–297, 1995.

[21] F. Deligianni, G. Varoquaux, B. Thirion, E. Robinson, D.J. Sharp, A.D. Edwards, and D. Rueckert. A probabilistic framework to infer brain functional connectivity from anatomical connections. In *IPMI: Information Processing in Medical Imaging*, pages 296–307. Springer, 2011.

[22] O. Demirci, V.P. Clark, V.A. Magnotta, N.C. Audreasen, J. Lauriello, K.A. Keihl, G.D. Pearlson, and V.D. Calhoun. A review of challenges in the use of fmri for disease classification/characterization and a projection pursuit application from a multi-site fmri schizophrenia study. *Brain Imaging and Behavior*, 2:207–226, 2008.

[23] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977.

[24] M. Descoteaux, R. Deriche, T. Knoesche, and A. Anwander. Deterministic and probabilistic tractography based on complex fiber orientation distributions. *IEEE Transactions on Medical Imaging*, 28(2):269–286, 2009.

[25] G. Desphande, S. LaConte, S. Peltier, and X. Hu. Directed transfer function analysis of fmri data to investigate network dynamics. In *IEEE Engineering in Medicine and Biology Society*, pages 671–674, 2006.

[26] R.F. Dougherty, M. Ben-Shachar, R. Bammer, A.A. Brewer, and B.A. Wandell. Functional organization of occipital-callosal fiber tracts. *PNAS*, 102(20):7350–7355, 2005.

[27] F. Esposito, A. Bertolino, T. Scarabino, V. Latorre, G. Blasi, T. Popolizio, G. Tedeschi, S. Cirillo, R. Goebel, and F. Di Salle. Independent component model of the default-mode brain function: Assessing the impact of active thinking. *Brain Research Bulletin*, 70:263–269, 2006.

[28] Y. Fan, R.E. Gur, R.C. Gur, X. Wu, D. Shen, M.E. Calkins, and C. Davatzikos. Unaffected family members and schizopphrenia patients share brain structure patterns: A high-dimensional pattern classification study. *Biological Psychiatry*, 63(1):118–124, 2008.

[29] B. Fischl, D.H. Salat, A.J.W. van der Kouwe, N. Makris, F. Ségonne, B.T. Quinn, and A.M. Dale. Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, 23:69–84, 2004.

[30] B. Fischl, M. Sereno, R. Tootell, and A. Dale. High-resolution intersubject averaging and a coordinate system for cortical surfaces. *Human Brain Mapping*, 8(4):272–284, 1999.

[31] J. Ford, H. Farid, F. Makedon, L.A. Flashman, T.W. McAllister, V. Megalooikonomou, and A.J. Saykin. Patient classification of fmri activation maps. In *MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention*, volume 6, pages 58–65, 2003.

[32] M.D. Fox and M.E. Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature*, 8:700–711, 2007.

[33] O. Friman, G. Farneback, and C-F. Westin. A bayesian approach for stochastic white matter tractography. *IEEE Transactions on Medical Imaging*, 25(8):965–978, 2006.

[34] K.J. Friston and C.D. Frith. Schizophrenia: A disconnection syndrome? *Clinical Neuroscience*, 3(2):89–97, 1995.

[35] K.J. Friston, A.P. Holmes, K.J. Worsley, J-P. Poline, C.D. Frith, and R.S.J Frackowiak. Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2:189–210, 1995.

[36] K.J. Friston, P. Jezzard, and R. Turner. Analysis of functional mri time-series. *Human Brain Mapping*, 1:151–174, 1994.

[37] S. Gabrieli-Whitfield, H.W. Thermenos, Z. Milanovic, M.T. Tsuang, S.V. Faraone, R.W. McCarley, M.E. Shenton, A.I. Green, A. Nieto-Castanon, P. LaViolette, J. Wojcik, J.D.E. Gabrieli, and L.J. Seidman. Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. *PNAS*, pages 1279–1284, 2009.

[38] A.G. Garrity, G.D. Pearlson, K. McKiernan, D. Lloyd, K.A. Kiehl, and V.D. Calhoun. Aberrant 'default mode' functional connectivity in schizophrenia. *American Journal of Psychiatry*, 164(3):450–457, 2007.

[39] P. Golland, Y. Golland, and R. Malach. Detection of spatial activation patterns as unsupervised segmentation of fmri data. In *MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 110–118. LNCS 479, 2007.

[40] M.D. Greicius, B.H. Flores, V. Menon, G.H. Glover, H.B. Solvason, H. Kenna, A.L. Reiss, and A.F. Schatzberg. Resting-state functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biological Psychiatry*, 62:429–437, 2007.

[41] M.D. Greicius, K. Supekar, V. Menon, and R.F. Dougherty. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral Cortex*, 19(1):72–78, 2008.

[42] M. Guye, G.J.M. Parker, M. Symms, P. Boulby, C.A.M. Wheeler-Kingshott, A. Salek-Haddadi, G.J. Barker, and J.S. Duncan. Combined functional mri and tractography to demonstrate the connectivity of the human primary motor cortex in vivo. *NeuroImage*, 19:1349–1360, 2003.

[43] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C.J. Honey, V.J. Weeden, and O. Sporns. Mapping the structural core of human cerebral cortex. *PLOS Biology*, 6(7):1479–1493, 2008.

[44] J.P. Hamilton, G. Chen, M.E. Thomason, M.E. Schwartz, and I.H. Gotlib. Investigating neural primacy in major depressive disorder: Multivariate granger causality analysis of resting-state fmri time-series data. *Molecular Psychiatry*, 16:763–772, 2011.

[45] L. Harrison, W.D. Penny, and K. Friston. Multivariate autoregressive modeling of fmri time series. *NeuroImage*, 19:1477–1491, 2003.

[46] D.J. Heeger and D. Ress. What does fmri tell us about neuronal activity? *Nature Reviews*, 3:142–151, 2002.

[47] C.J. Honey, R. Kotter, M. Breakspear, and O. Sporns. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *PNAS*, 104(24):10240–10245, 2007.

[48] C.J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J.P. Thiran, R. Meuli, and P. Hagmann. Predicting human resting-state functional connectivity from structural connectivity. *PNAS*, 106(6):2035–2040, 2009.

[49] S.G. Horovitz, A.R. Braun, W.S. Carr, D. Picchioni, T.J. Balkin, M. Fukunaga, and J.H. Duyn. Decoupling of the brain's default mode network during deep sleep. *PNAS*, 106:11376–11381, 2009.

[50] M.A. Horsfield and D.K. Jones. Applications of diffusion-weighted and diffusion tensor mri to white matter diseases - a review. *NMR in Biomedicine*, 15:570–577, 2002.

[51] A. Hyvarinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4-6):411–430, 2000.

[52] M.J. Jafri and V.D. Calhoun. Functional classification of schizophrenia using feed forward neural networks. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6631–6634, 2006.

[53] M.J. Jafri, G.D. Pearlson, M. Stevens, and V.D. Calhoun. A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage*, 39:1666–81, 2008.

[54] P. Jezzard, A.S. Barnett, and C. Pierpaoli. Characterization of and correction for eddy current artifacts in echo planar diffusion imaging. *Magnetic Resonance in Medicine*, 39(5):801–812, 1998.

[55] H. Johansen-Berg, T.E.J. Behrens, M.D. Robson, I. Drobnjak, M.F.S. Rushworth, J.M. Brady, S. M. Smith, D.J. Higham, and P.M. Matthews. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *PNAS*, 101(36):13335–13340, 2004.

[56] M.I. Jordan. An introduction to probabilistic graphical models. Technical report, University of California, Berkeley, 2003.

[57] M.I. Jordan, Z. Ghahramani, T.S. Jaakkola, and L.K. Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37:183–233, 1999.

[58] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23:151–160, 2004.

[59] K. Kasai, M.E. Shenton, D.F. Salisbury, Y. Hirayasu, T. Onitsuka, M.H. Spencer, D.A. Yurgelun-Todd, R. Kikinis, F.A. Jolesz, and R.W. McCarley. Progressive decrease of left heschl gyrus and planum temporale gray matter volume in first-episode schizophrenia, a longitudinal magnetic resonance imaging study. *Archives of General Psychiatry*, 60:766–775, 2003.

[60] M. Ke, X. Huang, H. Shen, Z. Zhou, X. Chen, and D. Hu. Combined analysis for resting state fmri and dti data reveals abnormal development of function-structure in early-onset schizophrenia. *LNAI*, 5009:628–635, 2008.

[61] M.A. Koch, D.G. Norris, and M. Hund-Georgiadis. An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage*, 16:241–250, 2002.

[62] A. Konrad and G. Winterer. Disturbed structural connectivity in schizophrenia - primary factor in pathology or epiphenomenon. *Schizophrenia Bulletin*, 34(1):72–92, 2008.

[63] M. Kubicki, R. McCarley, C-F Westin, H-J Park, S. Maier, R. Kikinis, F.A. Jolesz, and M.E. Shenton. A review of diffusion tensor imaging studies in schizophrenia. *Journal of Psychiatric Research*, 41:15–30, 2007.

[64] K. Lee, T. Yoshida, M. Kubicki, S. Bouix, C-F. Westin, G. Kindlmann, M. Niznikiewicz, A. Cohen, R.W. McCarley, and M.E. Shenton. Increased diffusivity in superior temporal gyrus in patients with schizophrenia: A diffusion tensor imaging study. *Schizophrenia Research*, 108(1-3):33–40, 2009.

[65] T-W. Lee, M. Girolami, and T.J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, 11:417–441, 1999.

[66] M. Liang, Y. Zhou, T. Jiang, Z. Liu, L. Tian, H. Liu, and Y. Hao. Widespread functional disconnectivity in schizophrenia with resting-state functional magnetic resonance imaging. *NeuroReport Brain Imaging*, 17(2):209–213, 2006.

[67] W. Liao, D. Mantini, Z. Zhang, Z. Pan, J. Ding, Q. Gong, Y. Yang, and H. Chen. Evaluating the effective connectivity of resting state networks using conditional granger causality. *Biological Cybernetics*, 102(1):57–69, 2010.

[68] N.K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann. Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412:150–157, 2001.

[69] Y. Lu, T. Jiang, and Y. Zang. Region growing method for the analysis of functional mri data. *NeuroImage*, 20(1):455–465, 2003.

[70] J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2:1–36, 1998.

[71] J.G. Malcolm, O. Michailovich, S. Bouix, C-F. Westin, M.E. Shenton, and Y. Rathi. A filtered approach to neural tractography using the watson directional function. *NeuroImage*, 14(1):58–69, 2010.

[72] A. Mechelli, C.J. Price, K.J. Friston, and J. Ashburner. Voxel-based morphometry of the human brain: Methods and applications. *Current Medical Imaging Reviews*, 1(2):105–113, 2005.

[73] E.D. Melonakos, M.E. Shenton, Y. Rathi, D.P. Terry, S. Bouix, and M. Kubicki. Voxel-based morphometry (vbm) studies in schizophrenia - can white matter changes be reliably detected with vbm. *Psychiatry Research*, 193(2):65–70, 2011.

[74] R.L.C. Mitchell, R. Elliott, and P.W.R. Woodruff. fmri and cognitive dysfunction in schizophrenia. *TRENDS in Cognitive Sciences*, 5(2):71–81, 2001.

[75] V. Mohan, G. Sundaramoorhi, M. Kubicki, T. Douglas, and A.R. Tannenbaum. Population analysis of the cingulum bundle using the tubular surface model for schizophrenia detection. In *Medical Imaging 2010: Computer-Aided Diagnosis*, volume 7624, 2010.

[76] B. Nadler, S. Lafon, R.R. Coifman, and I.G. Kevrekidis. Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. *Applied and Computational Harmonic Analysis: Special Issue on Diffusion Maps and Wavelets*, 21:113–127, 2006.

[77] S. Ogawa, T.M. Lee, A.R. Kay, and D.W. Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *PNAS*, 87:9868–72, 1990.

[78] G.D. Pearlson, P.E. Barta, R.E. Powers, R.R. Menon, S.S. Richards, E.H. Aylward, E.B. Federman, G.A. Chase, R.G. Petty, and A.Y. Tien. Median and superior temporal gyral volumes and cerebral asymmetry in schizophrenia versus bipolar disorder. *Biological Psychiatry*, 41(1):1–14, 1997.

[79] H.W.R. Powell, G.J.M. Parker, D.C. Alexander, M.R. Symms, P.A. Boulby, C.A.M. Wheeler-Kingshott, G.J. Barker, U. Noppeney, M.J. Koepp, and J.S. Duncan. Hemispheric asymmetries in language-related pathways: A combined functional mri and tractography study. *NeuroImage*, 32:388–399, 2006.

[80] E. Rykhlevskaia, G. Gratton, and M. Fabiani. Combining structural and functional neuroimaging data for studying brain connectivity: A review. *Psychophysiology*, 45:173–187, 2008.

[81] M.E. Shenton, C.C. Dickey, M. Frumin, and R.W. McCarley. A review of mri findings in schizophrenia. *Schizophrenia Research*, 49:1–52, 2001.

[82] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[83] S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Bern, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, R.K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J.M. Brady, and P.M. Matthews. Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23(S1):208–219, 2004.

[84] O. Sporns, D.R. Chialvo, M. Kaiser, and C.C. Hilgetag. Organization, development and function of complex brain networks. *TRENDS in Cognitive Science*, 8(9):418–425, 2004.

[85] O. Sporns, G. Tononi, and G.M. Edelman. Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex*, 10:127–41, 2000.

[86] C. Strobl, A-L. Boulesteix, and T. Augustin. Unbiased split selection for classification trees based on the gini index. *Comp. Stat. and Data Anal.*, 52:483–501, 2007.

[87] R. Tandon, M.S. Keshavan, and H.A. Nasrallah. Schizophrenia, 'just the facts': What we know in 2008, part 1: Overview. *Schizophrenia Research*, 100:4–19, 2008.

[88] B. Thirion, G. Varoquaux, and J-B. Poline. Accurate definition of brain regions position through the functional landmark approach. In *MICCAI: Medical Image Computing and Computer Assisted Intervention*, pages 241–248. LNCS, 2010.

[89] A.T. Toosy, , O. Ciccarelli, G.J.M. Parker, C.A.M. Wheeler-Kingshott, D.H. Miller, and A.J. Thompson. Characterizing function-structure relationships in the human visual system with functional mri and diffusion tensor imaging. *NeuroImage*, 21:1452–1463, 2004.

[90] K.R.A. van Dijk, T. Hedden, A. Venkataraman, K.C. Evans, S.W. Lazar, and R.L. Buckner. Intrinsic functional connectivity as a tool for human connectomics: Theory, properties and optimization. *Journal of Neurophysiology*, 103:297–321, 2010.

[91] G. Varoquaux, F. Baronnet, A. Kleinschmidt, P. Fillard, and B. Thirion. Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling. In *MICCAI: Medical Image Computing and Computer Assisted Intervention*, pages 200–208. LCNS, 2010.

[92] G. Varoquaux, A. Gramfort, F. Pedregosa, V. Michel, and B. Thirion. Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. In *IPMI: Information Processing in Medical Imaging*, pages 562–573. Springer, 2011.

[93] G. Varoquaux, A. Gramfort, J.B. Poline, and B. Thirion. Brain covariance selection: Better individual functional connectivity models using population prior. In *NIPS: Advances in Neural Information Processing Systems*, pages 2334–2342, 2010.

[94] A. Venkataraman, K.R.A. Van Dijk, R.L. Buckner, and P. Golland. Exploring functional connectivity in fmri via clustering. In *ICASSP: IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 441–444, 2009.

[95] A. Venkataraman, M. Kubicki, and P. Golland. From brain connectivity models to identifying foci of a neurological disorder. In *Accepted to MICCAI: Medical Image Computing and Computer Assisted Intervention*, 2012.

[96] A. Venkataraman, M. Kubicki, C-F. Westin, and P. Golland. Robust feature selection in resting-state fmri connectivity based on population studies. In *MMBIA: IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 63–70, 2010.

[97] A. Venkataraman, Y. Rathi, M. Kubicki, C-F. Westin, and P. Golland. Joint modeling of anatomical and functional connectivity for population studies. *IEEE Transactions on Medical Imaging*, 31(2):164–182, 2012.

[98] U. von Luxburg. A tutorial on spectral clustering. Technical Report TR-149, Max Planck Institute for Biological Cybernetics, 2006.

[99] M.J. Wainwright and M.I. Jordan. Graphical models, exponential families and variational inference. *Foundations and Trends in Machine Learning*, 1(1-2):1–305, 2008.

[100] L. Wang, Y. Zang, Y. He, M. Liang, X. Zhang, L. Tian, T. Wu, T. Jiang, and K. Li. Changes in hippocampal connectivity in the early stages of alzheimer's disease: Evidence from resting-state fmri. *NeuroImage*, 31:496–504, 2006.

[101] S. Warach, J. Gaa, B. Siewert, P. Wielopolski, and R.R. Edelman. Acute human stroke studied by whole brain echo planar diffusion-weighted magnetic resonance imaging. *Annals of Neurology*, 37(2):231–241, 1995.

[102] B.T. Yeo, F.M. Krienen, J. Sepulcre, M.R. Sabuncu, D. Lashkari, M. Hollinshead, J.L. Roffman, J.W. Smoller, L. Zollei, J.R. Polimeni, B. Fischl, H. Liu, and R.L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106:1125–1165, 2011.

[103] Y. Zhou, M. Liang, T. Jiang, L. Tian, Y. Liu, Z. Liu, H. Liu, and F. Kuang. Functional dysconnectivity of the dorsolateral prefrontal cortex in first-episode schizophrenia using resting-state fmri. *Neuro Letters*, 417:297–302, 2007.

[104] Y. Zhou, M. Liang, T. Jiang, L. Tian, Y. Liu, Z. Liu, H. Liu, and F. Kuang. Functional dysconnectivity of the dorsolateral prefrontal cortex in first-episode schizophrenia using resting-state fmri. *Neuroscience Letters*, 417:297–302, 2007.

[105] Y. Zhou, N. Shu, Y. Liu, M. Song, Y. Hao, H. Liu, C. Yu, Z. Liu, and T. Jiang. Altered resting-state functional connectivity and anatomical connectivity of hippocampus in schizophrenia. *Schizophrenia Research*, 100:120–132, 2008.

[106] L. Zollei, E. Learned-Miller, E. Grimson, and W. Wells. Efficient population registration in 3d data. In *MICCAI: Medical Image Computing and Computer Assisted Intervention*, pages 367–374. LNCS, 2007.